

Często stawiane, o prowokacyjnym charakterze pytanie: od czego uzależniona jest ilość dokonywanych przestępstw pospolitych w Polsce? było bezpośrednią przyczyną opracowania modelu ekonometrycznego obejmującego lata 1991 - 2002. Niniejsze opracowanie jest próbą udzielenia uproszczonej odpowiedzi na następujące pytania:

1. Czy i w jakim stopniu czynniki o charakterze demograficznym, takie jak:
  - a) wiek kobiet i mężczyzn,
  - b) ilość i poziom wykształconych bezrobotnych,
  - c) poziom zamożności społeczeństwa,
  - d) skala zjawiska braku aktywności zawodowej ludności w wieku produkcyjnymmają wpływa na dokonywanie przestępstw:
  - a) przeciwko bezpieczeństwu powszechnemu oraz bezpieczeństwu w komunikacji,
  - b) przeciwko wolności i obyczajowości
  - c) przeciwko czci i nietykalności cielesnej
  - d) przeciwko rodzinie, opiece młodzieży,
  - e) przeciwko mieniu społecznemu i prywatnemu,
  - f) przestępstw gospodarczych i skarbowych?
  
2. Czy otrzymane wyniki potwierdzą rezultaty badań naukowych, prowadzonych w USA, dotyczących międzynarodowego terroryzmu i zorganizowanej działalności przestępczej, że najbardziej podatną grupą wiekową na wszelkiego typu patologie o charakterze kryminogennym są mężczyźni od 20 do 34 roku życia?
  
3. Czy przestępczość małoletnich jest istotną składową przestępczości w naszym kraju?

Brak możliwości zgromadzenia kompletnych danych, rzeczywisty brak niektórych danych statystycznych, spowodował, że zostały zastosowane nieliczne uproszczenia, zaznaczone w dalszej części opracowania, które nie mają wpływu na otrzymany wynik.

# **1. POTENCJALNE ZMIENNE MODELU**

**1.1. Zmienne objaśniające**

**1.2. Potencjalne zmienne objaśniające**

## 1.1. Zmienne objaśniane ( przestępstwa ) 1) dla modelu ekonometrycznego obejmującego lata 1991 - 2002

Rok	Przestępstwa ogółem	P-ZZ	P-WO	P-CNC	P-RiM	P-PM	P-GOSP	P-SKAR
1991	866094	20106	13069	2592	36643	620705	2041	25399
1992	881076	21529	16317	2797	36613	591250	3037	44465
1993	852507	25740	19268	2613	39381	561001	2564	2506
1994	906157	29307	24753	2568	47707	582539	1716	2163
1995	974941	31885	28516	2875	49254	715355	17529	2860
1996	897751	33349	29463	3024	49368	613900	22253	2474
1997	992373	35920	35611	3668	53074	676836	26773	2683
1998	1073042	34926	36032	3876	53981	747015	27000	2475
1999	1121545	32934	36085	3742	51498	803977	27598	1332
2000	1266910	35424	44494	5407	53709	925433	30567	4160
2001	1390089	34057	47556	5489	53032	904623	24378	5378
2002	1404229	33866	47816	5672	48026	890085	21677	4910

Źródło danych: Roczniki statystyczne 1991 - 2003

1) ilość przestępstw stwierdzonych w zakończonych postępowaniach przygotowawczych

<b>P-ZZ</b>	Przestępstwa przeciwko bezpieczeństwu powszechnemu oraz bezpieczeństwu w komunikacji
<b>P-WO</b>	Przestępstwa przeciwko wolności i obyczajowości
<b>P-CNC</b>	Przestępstwa przeciwko czci i nietykalności cielesnej
<b>P-RiM</b>	Przestępstwa przeciwko rodzinie, opiece i młodzieży
<b>P-PM</b>	Przestępstwa przeciwko mieniu społecznemu i prywatnemu
<b>P-GOSP</b>	Przestępstwa gospodarcze
<b>P-SKAR</b>	Przestępstwa skarbowe

## 1.2. Potencjalne zmienne objaśniające dla modelu ekonometrycznego

Rok	Ludność wg wieku <sup>1)</sup>					Mężczyźni wg wieku <sup>1)</sup>				
	ogółem	15-19	20-34	35-50	pow.50	ogółem	15-19	20-34	35-50	pow.50
1990	38183,0	2897,8	8327,3	7812,6	9622,8	18606,0	1481,1	4232,6	3888,9	4132,5
1991	38309,0	2998,6	8179,6	8090,0	9630,2	18661,3	1203,4	4159,2	4026,9	4128,6
1992	38418,0	3067,7	8057,4	8387,4	9628,5	18708,0	1567,4	4097,8	4176,3	4120,5
1993	38505,0	3139,0	7956,6	8675,5	9622,1	18746,3	1604,2	4047,0	4321,0	4112,3
1994	38581,0	3194,9	7919,1	8913,7	9630,4	18778,0	1632,0	4029,3	4440,9	4111,2
1995	38609,0	3236,8	8131,5	9114,2	9648,8	18796,7	1656,4	4077,6	4561,1	4177,0
1996	38639,0	3244,8	8012,7	9152,0	9783,7	18799,2 <sup>4)</sup>	1665,0 <sup>4)</sup>	4098,8 <sup>4)</sup>	4557,8 <sup>4)</sup>	4233,2 <sup>4)</sup>
1997	38660,0	3292,5	8101,3	9130,5	9966,2	18801,2	1680,0	4123,3	4550,7	4262,8
1998	38667,0	3343,1	8212,1	9060,8	10189,9	18798,3	1704,7	4179,9	4516,3	4368,4
1999	38654,0	3365,7	8345,2	8974,1	10411,0	18783,5	1715,8	4247,7	4474,7	4470,8
2000	38644,0	3356,9	8503,8	8835,9	10674,7	18773,0	1711,6	4327,8	4406,7	4595,9
2001	38632,0	3322,3	8659,0	8678,9	10952,9	18760,8	1693,6	4405,9	4331,9	4726,4
2002	38219,0	3231,1	8639,5	8371,4	11172,2	18506,7	1650,6	4385,4	4175,4	4811,0

Bezrobotni i bierni zawodowo <sup>1)</sup>				Bezrobotni wg wykształcenia i mężczyźni <sup>1)</sup>							
bezrobotni zarejestrowani ogółem	bierni zawodowo ogółem	bezrobotni zarejestrowani mężczyźni	bierni zawodowo mężczyźni	wyższe ogółem	wyższe mężczyźni	średnie ogólne ogółem	średnie ogólne mężczyźni	średnie zawodo we ogółem	średnie zawodo we mężczyźni	zasadnicze zawodo we ogółem	zasadnicze zawodo we mężczyźni
1126,1	4432,0 <sup>4)</sup>	552,4	1654,0 <sup>4)</sup>	34,2 <sup>4)</sup>	14,7 <sup>4)</sup>	80,4 <sup>4)</sup>	15,2 <sup>4)</sup>	252,9 <sup>4)</sup>	94,8 <sup>4)</sup>	420,9 <sup>4)</sup>	223,7 <sup>4)</sup>
2155,6	9001,0 <sup>4)</sup>	1021,5	3965,0 <sup>4)</sup>	64,1	31,4	155,5	33,7	482,5	177,8	805,6	443,7
2509,3	10862,0	1170,5	4054,0	56,6	25,7	177,9	34,4	527,9	175,2	964,6	525,2
2889,6	11013,0	1382,3	4096,0	52,1	23,9	193,2	35,9	581,5	193,5	1131,2	631,0
2838,0	11781,0	1343,0	4525,0	47,6	20,9	194,4	33,9	570,2	185,1	1118,3	619,4
2628,8	12102,0	1180,0	4626,0	38,9	15,1	188,2	31,4	531,7	160,8	1025,0	541,6
2359,5	12422,0	983,9	4785,0	31,4	11,7	151,7	22,1	471,0	126,0	907,7	442,5
1826,4	12675,0	723,2	4878,0	25,1	8,9 <sup>4)</sup>	115,2	16,5 <sup>4)</sup>	364,2	100,5 <sup>4)</sup>	700,7	348,4 <sup>4)</sup>
1831,4	12899,0	760,1	5022,0	30,2	10,5	111,2	17,6	369,0	102,2	698,0	337,5
2349,8	13174,0	1042,5	5171,0	47,7	17,0	142,0	25,7	483,0	151,3	898,2	462,8
2702,6	13371,0	1211,0	5222,0	69,4	24,3	168,7	32,9	561,9	183,0	998,8	522,3
3115,1	13664,0	1473,0	5381,0	100,5	37,5	192,1	42,2	666,0	237,6	1142,5	631,0
3217,0	14012,0	1571,2	5544,0	126,7	47,8	199,0	48,1	683,6	257,9	1163,6	658,9

Bezrobotni wg wykształcenia i mężczyźni <sup>1)</sup>		Wskaźnik inflacji, rok poprzedni = 100	Dochody rozp. na os.		Wydatki na osobę		Rok
podstawowe ogółem	podstawowe mężczyźni		ogółem w rodzinach	w rodzinie pracowniwej	ogółem w rodzinach	w rodzinie pracowniwej	
337,7 <sup>4)</sup>	163,4 <sup>4)</sup>	685,8	brak danych	591,55 <sup>2,4)</sup>	brak danych	488,70 <sup>2)</sup>	1990
663,3 <sup>4)</sup>	396,4 <sup>4)</sup>	170,3	brak danych	979,00 <sup>2)</sup>	brak danych	879,70 <sup>2)</sup>	1991
782,4	410,0	143	brak danych	1428,30 <sup>2)</sup>	brak danych	1378,80 <sup>2)</sup>	1992
931,5	498,0	135,3	173,62 <sup>3)</sup>	176,61 <sup>3)</sup>	169,32 <sup>3)</sup>	168,32 <sup>3)</sup>	1993
907,5	483,7	132,2	230,93 <sup>3)</sup>	234,69 <sup>3)</sup>	218,61 <sup>3)</sup>	216,92 <sup>3)</sup>	1994
845,0	431,3	127,8	300,56 <sup>3)</sup>	301,26 <sup>3)</sup>	276,32 <sup>3)</sup>	274,93 <sup>3)</sup>	1995
797,6	381,6	119,9	383,43 <sup>3)</sup>	395,25 <sup>3)</sup>	351,00 <sup>3)</sup>	355,29 <sup>3)</sup>	1996
621,3	289,9 <sup>4)</sup>	114,9	473,79 <sup>3)</sup>	486,41 <sup>3)</sup>	427,91 <sup>3)</sup>	435,61 <sup>3)</sup>	1997
622,8	292,3	111,8	522,93 <sup>3)</sup>	546,35 <sup>3)</sup>	503,03 <sup>3)</sup>	517,66 <sup>3)</sup>	1998
778,9	385,7	107,3	560,43 <sup>3)</sup>	brak danych	549,76 <sup>3)</sup>	brak danych	1999
903,8	448,4	110,1	610,51 <sup>3)</sup>	brak danych	599,49 <sup>3)</sup>	brak danych	2000
1014,0	524,7	105,5	644,48 <sup>3)</sup>	brak danych	609,72 <sup>3)</sup>	brak danych	2001
1044,1	558,5	101,9	664,21 <sup>3)</sup>	brak danych	624,99 <sup>3)</sup>	brak danych	2002

Źródło danych: Roczniki statystyczne 1991 - 2003

<sup>1)</sup> w tysiącach osób

<sup>2)</sup> w tysiącach złotych

<sup>3)</sup> w złotych

<sup>4)</sup> dane nie pochodzące z rocznika statystycznego

Rok	Ludność wg wieku					Bezrobotni ogółem	Bierni zawodowo ogółem	Bezrobotni wg wykształcenia i mężczyźni			
	L	L15-19	L20-34	L35-50	L>50	B	BIER	B-W	B-W-M	B-SO	B-SO-M
1991	38309,0	2998,6	8179,6	8090,0	9630,2	2155,6	9001,0	64,1	31,4	155,5	33,7
1992	38418,0	3067,7	8057,4	8387,4	9628,5	2509,3	10862,0	56,6	25,7	177,9	34,4
1993	38505,0	3139,0	7956,6	8675,5	9622,1	2889,6	11013,0	52,1	23,9	193,2	35,9
1994	38581,0	3194,9	7919,1	8913,7	9630,4	2838,0	11781,0	47,6	20,9	194,4	33,9
1995	38609,0	3236,8	8131,5	9114,2	9648,8	2628,8	12102,0	38,9	15,1	188,2	31,4
1996	38639,0	3244,8	8012,7	9152,0	9783,7	2359,5	12422,0	31,4	11,7	151,7	22,1
1997	38660,0	3292,5	8101,3	9130,5	9966,2	1826,4	12675,0	25,1	8,9	115,2	16,5
1998	38667,0	3343,1	8212,1	9060,8	10189,9	1831,4	12899,0	30,2	10,5	111,2	17,6
1999	38654,0	3365,7	8345,2	8974,1	10411,0	2349,8	13174,0	47,7	17,0	142,0	25,7
2000	38644,0	3356,9	8503,8	8835,9	10674,7	2702,6	13371,0	69,4	24,3	168,7	32,9
2001	38632,0	3322,3	8659,0	8678,9	10952,9	3115,1	13664,0	100,5	37,5	192,1	42,2
2002	38219,0	3231,1	8639,5	8371,4	11172,2	3217,0	14012,0	126,7	47,8	199,0	48,1

B-SZ	B-SZ-M	B-Z	B-Z-M	B-P	B-P-M	Doch. rozp./os.		Wydatki na osobę		Średnia	
						praco wnicze	ogółem	praco wnicze	ogółem	dochody	wydatki
482,5	177,8	805,6	443,7	663,3	396,4	979,00	brak danych	879,70	brak danych		
527,9	175,2	964,6	525,2	782,4	410,0	1428,30	brak danych	1378,80	brak danych		
581,5	193,5	1131,2	631,0	931,5	498,0	176,61	173,62	168,32	169,32	0,122	1,006
570,2	185,1	1118,3	619,4	907,5	483,7	234,69	230,93	216,92	218,61	1,308	1,008
531,7	160,8	1025,0	541,6	845,0	431,3	301,26	300,56	274,93	276,32	1,281	1,005
471,0	126,0	907,7	442,5	797,6	381,6	395,25	383,43	355,29	351,00	<b>średnia</b>	<b>średnia</b>
364,2	100,5	700,7	348,4	621,3	289,9	486,41	473,79	435,61	427,91	<b>0,903</b>	<b>1,006</b>
369,0	102,2	698,0	337,5	622,8	292,3	546,35	522,93	517,66	503,03		
483,0	151,3	898,2	462,8	778,9	385,7	brak danych	560,43	brak danych	549,76		
561,9	183,0	998,8	522,3	903,8	448,4	brak danych	610,51	brak danych	599,49		
666,0	237,6	1142,5	631,0	1014,0	524,7	brak danych	644,48	brak danych	609,72		
683,6	257,9	1163,6	658,9	1044,1	558,5	brak danych	664,21	brak danych	624,99		



Nominalne		Nominalne		Wskaźnik		Realne	
dochody	wydatki	dochody PLN	w wydatki PLN	inflacji, rok poprzedni =1	inflacji, rok 2002 = 1	dochody w PLN - ceny stałe z 2002.	wydatki w PLN - ceny stałe z 2002.
884,30	885,21	88,43	88,52	1,7030	0,1564	565,44	566,03
1290,13	1387,43	129,01	138,74	1,4300	0,2236	576,89	620,39
173,62	169,32	173,62	169,32	1,3530	0,3026	573,80	559,58
230,93	218,61	230,93	218,61	1,3220	0,4000	577,31	546,51
300,56	276,32	300,56	276,32	1,2780	0,5112	587,93	540,51
383,43	351,00	383,43	351,00	1,1990	0,6129	625,55	572,64
473,79	427,91	473,79	427,91	1,1490	0,7043	672,73	607,59
522,93	503,03	522,93	503,03	1,1180	0,7874	664,14	638,86
560,43	549,76	560,43	549,76	1,0730	0,8449	663,34	650,71
610,51	599,49	610,51	599,49	1,1010	0,9302	656,33	644,48
644,48	609,72	644,48	609,72	1,0550	0,9814	656,73	621,30
664,21	624,99	664,21	624,99	1,0190	1,0000	664,21	624,99

Zrezygnowano z wykorzystania wszystkich wskaźników wśród mężczyzn ponieważ w każdym z przypadków były one silnie skorelowane ( $r > 0,8$ ) z analogicznymi wskaźnikami w całej populacji społeczności.

Rok	L	L15-19	L20-34	L35-50	LB	LBZ	SB	WBZ
1991	38309,0	2998,6	8179,6	8090,0	2155,6	9001,0	12,2	37,8
1992	38418,0	3067,7	8057,4	8387,4	2509,3	10862,0	14,3	38,3
1993	38505,0	3139,0	7956,6	8675,5	2889,6	11013,0	16,4	38,8
1994	38581,0	3194,9	7919,1	8913,7	2838,0	11781,0	16,0	39,1
1995	38609,0	3236,8	8131,5	9114,2	2628,8	12102,0	14,9	41,6
1996	38639,0	3244,8	8012,7	9152,0	2359,5	12422,0	13,2	42,4
1997	38660,0	3292,5	8101,3	9130,5	1826,4	12675,0	10,3	42,6
1998	38667,0	3343,1	8212,1	9060,8	1831,4	12899,0	10,4	42,9
1999	38654,0	3365,7	8345,2	8974,1	2349,8	13174,0	13,1	43,3
2000	38644,0	3356,9	8503,8	8835,9	2702,6	13371,0	15,1	43,6
2001	38632,0	3322,3	8659,0	8678,9	3115,1	13664,0	17,5	44,2
2002	38219,0	3231,1	8639,5	8371,4	3217,0	14012,0	18,0	44,8

#### Ludność <sup>1)</sup>

<b>L</b>	ogółem
<b>L15-19</b>	od 15 do 20 roku życia
<b>L20-34</b>	od 20 do 34 roku życia
<b>L35-50</b>	od 35 do 50 roku życia

#### Bezrobotni <sup>1)</sup>

<b>LB</b>	bezrobotni ogółem
<b>LBZ</b>	liczba biernych zawodowo
<b>SB</b>	stopa bezrobocia
<b>WBZ <sup>5)</sup></b>	wskaźnik bierności zawodowej

B-W	B-SO	B-SZ	B-Z	B-P	DRG	WRG
64,1	155,5	482,5	805,6	663,3	618,64	566,03
56,6	177,9	527,9	964,6	782,4	631,16	620,39
52,1	193,2	581,5	1131,2	931,5	573,80	559,58
47,6	194,4	570,2	1118,3	907,5	577,31	546,51
38,9	188,2	531,7	1025,0	845,0	587,93	540,51
31,4	151,7	471,0	907,7	797,6	625,55	572,64
25,1	115,2	364,2	700,7	621,3	672,73	607,59
30,2	111,2	369,0	698,0	622,8	664,14	638,86
47,7	142,0	483,0	898,2	778,9	663,34	650,71
69,4	168,7	561,9	998,8	903,8	656,33	644,48
100,5	192,1	666,0	1142,5	1014,0	656,73	621,30
126,7	199,0	683,6	1163,6	1044,1	664,21	624,99

**Bezrobotni z wykształceniem <sup>1)</sup>**

**B\_W** wyższym  
**B\_SO** średnim ogólnym  
**B\_SZ** średnim zawodowym  
**B\_Z** zawodowym

**DRG <sup>3)</sup>**

**WRG <sup>3)</sup>**

dochód realny w rodzinie na osobę  
wydatki w rodzinie na osobę

<sup>1)</sup> w tysiącach osób

<sup>3)</sup> w złotych

<sup>5)</sup> **WBZ** = bierni zawodowo / ludność w wieku produkcyjnym

## **2. PRZESTĘPSTWA PRZECIWKO CZCI I NIETYKALNOŚCI CIELESNEJ**

- 2.1. Dobór zmiennych objaśniających**
- 2.2. Estymacja parametrów równania**
- 2.3. Weryfikacja równania**
- 2.4. Powtórna estymacja równania**
- 2.5. Weryfikacja poprawionego równania**
- 2.6. Prognoza zmiennych objaśniających**
- 2.7. Obliczenie prognozy dla przestępstw przeciwko czci i nietykalności cielesnej**

**2.1. Dobór zmiennych objaśniających ( egzogenicznych ) dla równania modelu ,  
w której zmienną objaśnianą stanowią przestępstwa przeciwko czci i nietykalności  
cielesnej ( P-CNC )**

**2.1.1. Baza danych niezbędna do wytypowania zmiennych równania**

P_CNC	L	L15-19	L20-34	L35-50	LB	LBZ	SB	WBZ	B_W	B_SO	B_SZ
2592	38309,0	2998,6	8179,6	8090,0	2155,6	9001,0	12,2	37,8	64,1	155,5	482,5
2797	38418,0	3067,7	8057,4	8387,4	2509,3	10862,0	14,3	38,3	56,6	177,9	527,9
2613	38505,0	3139,0	7956,6	8675,5	2889,6	11013,0	16,4	38,8	52,1	193,2	581,5
2568	38581,0	3194,9	7919,1	8913,7	2838,0	11781,0	16,0	39,1	47,6	194,4	570,2
2875	38609,0	3236,8	8131,5	9114,2	2628,8	12102,0	14,9	41,6	38,9	188,2	531,7
3024	38639,0	3244,8	8012,7	9152,0	2359,5	12422,0	13,2	42,4	31,4	151,7	471,0
3668	38660,0	3292,5	8101,3	9130,5	1826,4	12675,0	10,3	42,6	25,1	115,2	364,2
3876	38667,0	3343,1	8212,1	9060,8	1831,4	12899,0	10,4	42,9	30,2	111,2	369,0
3742	38654,0	3365,7	8345,2	8974,1	2349,8	13174,0	13,1	43,3	47,7	142,0	483,0
5407	38644,0	3356,9	8503,8	8835,9	2702,6	13371,0	15,1	43,6	69,4	168,7	561,9
5489	38632,0	3322,3	8659,0	8678,9	3115,1	13664,0	17,5	44,2	100,5	192,1	666,0
5672	38219,0	3231,1	8639,5	8371,4	3217,0	14012,0	18,0	44,8	126,7	199,0	683,6

B_Z	B_P	DRG	WRG	V
805,6	663,3	618,64	566,03	0
964,6	782,4	631,16	620,39	0
1131,2	931,5	573,80	559,58	0
1118,3	907,5	577,31	546,51	0
1025,0	845,0	587,93	540,51	0
907,7	797,6	625,55	572,64	0
700,7	621,3	672,73	607,59	0
698,0	622,8	664,14	638,86	0
898,2	778,9	663,34	650,71	0
998,8	903,8	656,33	644,48	1
1142,5	1014,0	656,73	621,30	1
1163,6	1044,1	664,21	624,99	1

## 2.1.2. Macierz korelacji zmiennych

	<i>P_CNC</i>	<i>L</i>	<i>L15-19</i>	<i>L20-34</i>	<i>L35-50</i>	<i>LB</i>	<i>LBZ</i>	<i>SB</i>	<i>WBZ</i>	<i>B_W</i>	<i>B_SO</i>
<i>P_CNC</i>	1,000	-0,017	0,610	0,935	-0,064	0,389	0,795	0,365	0,841	0,694	0,094
<i>L</i>	-0,017	1,000	0,699	-0,157	0,872	-0,350	0,369	-0,358	0,305	-0,622	-0,437
<i>L15-19</i>	0,610	0,699	1,000	0,465	0,703	-0,045	<b>0,885</b>	-0,071	<b>0,857</b>	-0,051	-0,326
<i>L20-34</i>	0,935	-0,157	0,465	1,000	-0,253	0,406	0,638	0,385	0,745	0,778	0,154
<i>L35-50</i>	-0,064	0,872	0,703	-0,253	1,000	-0,328	0,476	-0,338	0,398	-0,644	-0,409
<i>LB</i>	0,389	-0,350	-0,045	0,406	-0,328	1,000	0,272	0,999	0,141	0,760	0,935
<i>LBZ</i>	0,795	0,369	<b>0,885</b>	0,638	0,476	0,272	1,000	0,247	<b>0,940</b>	0,310	-0,034
<i>SB</i>	0,365	-0,358	-0,071	0,385	-0,338	0,999	0,247	1,000	0,111	0,751	0,944
<i>WBZ</i>	0,841	0,305	<b>0,857</b>	0,745	0,398	0,141	<b>0,940</b>	0,111	1,000	0,333	-0,170
<i>B_W</i>	0,694	-0,622	-0,051	0,778	-0,644	0,760	0,310	0,751	0,333	1,000	0,618
<i>B_SO</i>	0,094	-0,437	-0,326	0,154	-0,409	0,935	-0,034	0,944	-0,170	0,618	1,000
<i>B_SZ</i>	0,435	-0,445	-0,107	0,493	-0,461	0,985	0,223	0,983	0,131	0,844	0,921
<i>B_Z</i>	0,247	-0,295	-0,089	0,257	-0,262	0,985	0,199	0,988	0,034	0,646	0,957
<i>B_P</i>	0,474	-0,251	0,095	0,456	-0,215	0,987	0,401	0,982	0,269	0,742	0,877
<i>DRG</i>	0,707	0,104	0,557	0,665	0,045	-0,288	0,591	-0,312	0,713	0,237	-0,546
<i>WRG</i>	0,693	0,120	0,561	0,654	-0,026	-0,121	0,579	-0,141	0,606	0,281	-0,383

<i>B_SZ</i>	<i>B_Z</i>	<i>B_P</i>	<i>DRG</i>	<i>WRG</i>
0,435	0,247	0,474	0,707	0,693
-0,445	-0,295	-0,251	0,104	0,120
-0,107	-0,089	0,095	0,557	0,561
0,493	0,257	0,456	0,665	0,654
-0,461	-0,262	-0,215	0,045	-0,026
0,985	0,985	0,987	-0,288	-0,121
0,223	0,199	0,401	0,591	0,579
0,983	0,988	0,982	-0,312	-0,141
0,131	0,034	0,269	0,713	0,606
0,844	0,646	0,742	0,237	0,281
0,921	0,957	0,877	-0,546	-0,383
1,000	0,948	0,959	-0,215	-0,067
0,948	1,000	0,966	-0,425	-0,235
0,959	0,966	1,000	-0,209	-0,052
-0,215	-0,425	-0,209	1,000	<b>0,884</b>
-0,067	-0,235	-0,052	<b>0,884</b>	1,000

### 2.1.3. Zasady doboru zmiennych do równania

1. Wektor współczynników korelacji zmiennej objaśnianej z potencjalnymi zmiennymi objaśniającymi zaznaczono kolorem
2. Wybór najsilniej skorelowanej potencjalnej zmiennej egzogenicznej ze zmienną endogeniczną (objaśnianą) został zaznaczony kolorem . Zmienna została automatycznie zakwalifikowana do modelu. Ponadto przyjęto maksymalną wartość modułu współczynnika korelacji między zmiennymi egzogenicznymi na poziomie 0,8. Wyższe współczynniki pomiędzy potencjalnymi zmiennymi objaśniającymi zaznaczono kolorem . Do modelu należy zakwalifikować jedną z nich.
3. W następnej kolejności rozważano zasadność doboru drugiej (obok L\_20\_34) zmiennej objaśniającej. W celu wstępnej selekcji pozostałych zmiennych zastosowano następującą procedurę:
  - a) 

	>	
	<	

 Dobór zmiennej może posiadać uzasadnienie.
  - b) 

	>	
	<	

 Dobór zmiennej nie posiada uzasadnienia.

W świetle powyższych kryteriów można rozważyć dobór zmiennych oznaczonych kolorem . Należy zatem rozważyć następujące pary cech:

<p>A <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="background-color: #00ff00; width: 50px; height: 15px; text-align: center;">L20-34</td><td style="background-color: #ff0000; width: 50px; height: 15px; text-align: center;">L15-19</td></tr></table></p>	L20-34	L15-19	<p>B <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="background-color: #00ff00; width: 50px; height: 15px; text-align: center;">L20-34</td><td style="background-color: #ff0000; width: 50px; height: 15px; text-align: center;">LBZ</td></tr></table></p>	L20-34	LBZ	<p>C <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="background-color: #00ff00; width: 50px; height: 15px; text-align: center;">L20-34</td><td style="background-color: #ff0000; width: 50px; height: 15px; text-align: center;">WBZ</td></tr></table></p>	L20-34	WBZ
L20-34	L15-19							
L20-34	LBZ							
L20-34	WBZ							
<p>D <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="background-color: #00ff00; width: 50px; height: 15px; text-align: center;">L20-34</td><td style="background-color: #ff0000; width: 50px; height: 15px; text-align: center;">B_P</td></tr></table></p>	L20-34	B_P	<p>E <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="background-color: #00ff00; width: 50px; height: 15px; text-align: center;">L20-34</td><td style="background-color: #ff0000; width: 50px; height: 15px; text-align: center;">DRG</td></tr></table></p>	L20-34	DRG	<p>F <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="background-color: #00ff00; width: 50px; height: 15px; text-align: center;">L20-34</td><td style="background-color: #ff0000; width: 50px; height: 15px; text-align: center;">WRG</td></tr></table></p>	L20-34	WRG
L20-34	B_P							
L20-34	DRG							
L20-34	WRG							

W celu wyboru najkorzystniejszej kombinacji zastosowano metodę Hellwiga.

Wektor modułów współczynnika korelacji zmiennej objaśnianej ( endogenicznej ) z badanymi zmiennymi

	<i>P_CNC</i>
<b>L20-34</b>	0,935
<b>L15-19</b>	0,610
<b>LBZ</b>	0,795
<b>WBZ</b>	0,841
<b>B_P</b>	0,474
<b>DRG</b>	0,707
<b>WRG</b>	0,693



Macierz modułów współczynników korelacji między badanymi zmiennymi

	L20-34	L15-19	LBZ	WBZ	B_P	DRG	WRG
L20-34	1,000	0,465	0,638	0,745	0,456	0,665	0,654
L15-19	0,465	1,000	0,885	0,857	0,095	0,557	0,561
LBZ	0,638	0,885	1,000	0,940	0,401	0,591	0,579
WBZ	0,745	0,857	0,940	1,000	0,269	0,713	0,606
B_P	0,456	0,095	0,401	0,269	1,000	0,209	0,052
DRG	0,665	0,557	0,591	0,713	0,209	1,000	0,884
WRG	0,654	0,561	0,579	0,606	0,052	0,884	1,000

Obliczono wskaźniki pojemności indywidualnej nośnika informacji od A do F(...) w celu wyłonienia najkorzystniejszej kombinacji dwuelementowej:

**A**

h(A,L20-34)	0,597
h(A,L15-19)	0,254
H(A)	0,852

**B**

h(B,L20-34)	0,534
h(B,LBZ)	0,385
H(B)	0,920

**C**

h(C,L20-34)	0,501
h(C,WBZ)	0,405
H(C)	0,906

**D**

h(D,L20-34)	0,601021907
h(D, B_P)	0,154544395
H(D)	0,755566302

**E**

h(E,L20-34)	0,525266676
h(E, DRG)	0,29996469
H(E)	0,825231366

**F**








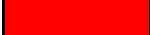
h(F,L20-34)	0,528893
h(F, WRG)	0,29004123
H(F)	0,81893423

**Wniosek:**

Spośród rozpatrywanych par najlepiej zmienną P\_CNC opisuje kombinacja B.

4. Nie ma jednak pewności, czy nie jest możliwa poprawa jakości równania poprzez włączenie trzeciej zmiennej.

Jako potencjalną trzecią zmienną rozważono każdą z cech oznaczonych wcześniej kolorem

	>		i		>		Dobór zmiennej może mieć uzasadnienie.
	<		lub		<		Dobór zmiennej nie ma uzasadnienia,

ponieważ w większym stopniu "powiela" ona jedną ze zmiennych egzogenicznych niż opisuje zmienną P\_WO

Odpowiednie wiersze i kolumny macierzy korelacji kształtują się następująco:

	<i>P_CNC</i>	<i>L20-34</i>	<i>LBZ</i>
<i>L15-19</i>	0,610	0,465	<b>0,885</b>
<i>WBZ</i>	0,841	0,745	<b>0,940</b>
<i>B_P</i>	<b>0,474</b>	<b>0,456</b>	<b>0,401</b>
<i>DRG</i>	<b>0,707</b>	<b>0,665</b>	<b>0,591</b>
<i>WRG</i>	<b>0,693</b>	<b>0,654</b>	<b>0,579</b>

Można zatem sprawdzić zasadność włączenia do równania jednej ze zmiennych oznaczonych na Pozostałe cechy są b.silnie skorelowane z LBZ.



Obliczono zatem wskaźnik pojemności indywidualnej nośnika informacji dla poniższych kombinacji:

**G**

<i>L20-34</i>	<i>LBZ</i>	<i>B_P</i>
---------------	------------	------------

**H**

<i>L20-34</i>	<i>LBZ</i>	<i>DRG</i>
---------------	------------	------------

**I**

<i>L20-34</i>	<i>LBZ</i>	<i>WRG</i>
---------------	------------	------------

**G**

h(G,L20-34)	0,418
h(G,LBZ)	0,310
h(G,B_P)	0,121
H(G)	0,849

**H**

h(H,L20-34)	0,380
h(H,LBZ)	0,221
h(H,DRG)	0,295
H(H)	0,896

**I**

h(I,L20-34)	0,382
h(I,LBZ)	0,285
h(I,WRG)	0,215
H(I)	0,881

**Wniosek:**

Dodanie jednej ze zmiennych B\_P, DRG oraz WRG nie poprawi jakości modelu, gdyż wskaźnik pojemności indywidualnej nośnika informacji H(B) dla kombinacji B jest wyższy od wskaźników H(G), H(H), H(I).  
Zatem do równania należy zakwalifikować L\_20-24 i LBZ.

## 2.2. Estymacja parametrów równania opisującego zmienną P\_CNC

### 2.2.1. Baza danych niezbędna do oszacowania równania

Rok	P_CNC	L20-34	LBZ
1991	2592	8179,6	9001
1992	2797	8057,4	10862
1993	2613	7956,6	11013
1994	2568	7919,1	11781
1995	2875	8131,5	12102
1996	3024	8012,7	12422
1997	3668	8101,3	12675
1998	3876	8212,1	12899
1999	3742	8345,2	13174
2000	5407	8503,8	13371
2001	5489	8659	13664
2002	5672	8639,5	14012

### 2.2.2. Klasyczną MNK za pomocą MC Excel oszacowano parametry strukturalne oraz parametry stochastyczne modelu

$$P\_CNC = 3,377[L20-34_t] + 0,281[LBZ_t] - 27534,07 + u_t$$

**(0,492)                      (0,089)                      (3457,545)**

### Interpretacja :

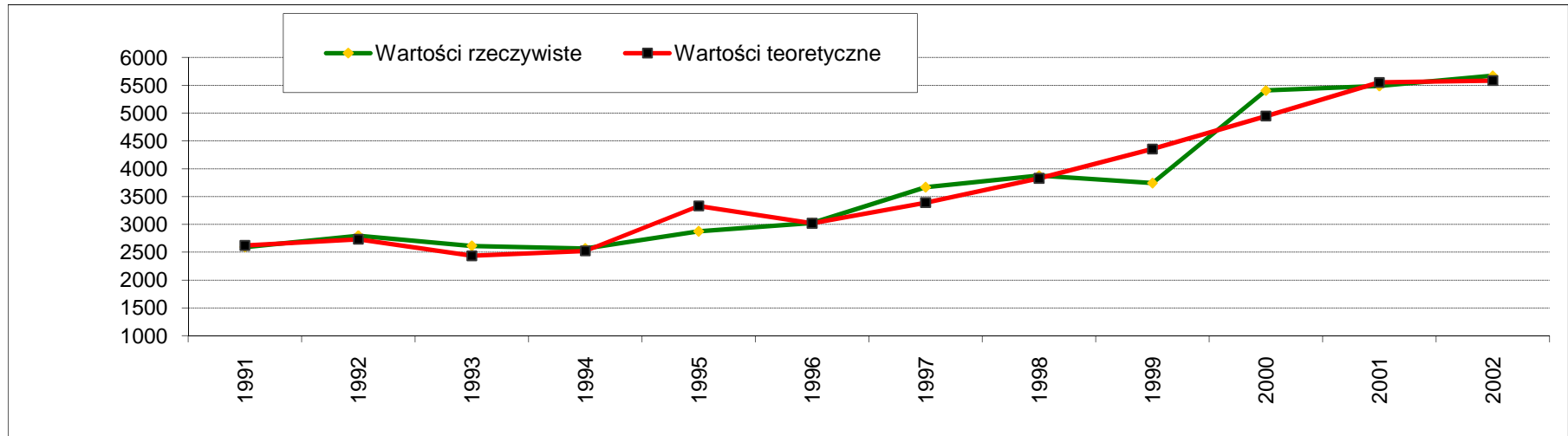
1. Jeżeli liczba ludności w wieku od 20 do 34 lat ( zmienna L20-34 ) wzrośnie o 1000 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) wzrośnie przeciętnie o 3,377 przypadków. przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) wzrośnie przeciętnie
2. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej L20-34 na podstawie 12-elementowych prób, pomylibyśmy się przeciętnie rzecz biorąc o 0,492.
3. Jeżeli liczba biernych zawodowo ( zmienna LBZ ) wzrośnie o 1000 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) wzrośnie przeciętnie o 0,281 przypadków .
4. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej LBZ na podstawie 12-elementowych prób, pomylibyśmy się przeciętnie rzecz biorąc o 0,089.

<i>Miary dopasowania</i>	
R	0,970
R <sup>2</sup>	0,941
dopas. R <sup>2</sup>	0,928
Su	320,823
Vu	8,69%

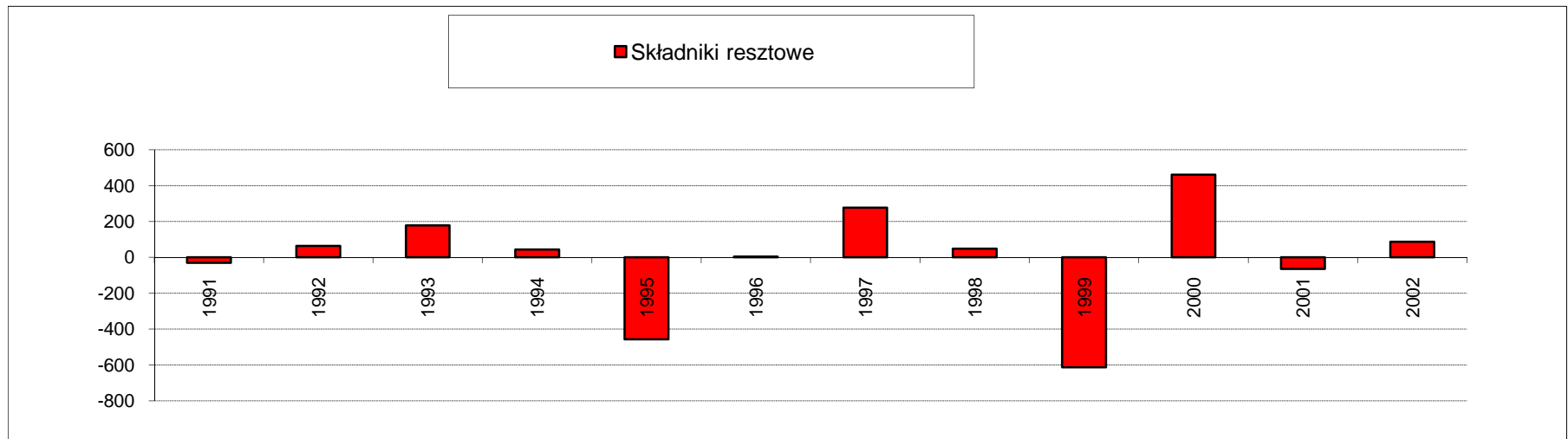
### Wnioski:

1. Ponad 94 % różnicowania liczby przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) jest wyjaśniane przez model.
2. Empiryczna liczba przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) odchyła się przeciętnie o 320,823 przypadków od wartości teoretycznej wyznaczonej na podstawie modelu.
3. Odchylenia te stanowią 8,69% przeciętnej liczby przestępstw przeciwko czci i nietykalności cielesnej.

Rys. nr 1 Empiryczne wartości zmiennej P\_CNC a wartości teoretyczne



Rys. nr 2 Składniki resztowe równania modelującego zmienną P\_CNC



## 2.3. Weryfikacja równania

### 2.3.1. Badanie istotności parametrów strukturalnych równania

<i>Współczynniki</i>		<i>Błąd stand.</i>	<i>t Stat</i>
Przecięcie	-27534,07	3457,545	-7,963
L20-34	3,377	0,492	6,867
LBZ	0,281	0,089	3,173

obserw.	12
zm. objaśnij.	2
poziom ist.	0,05
wartość $t^*$ <sup>1)</sup>	2,262

<sup>1)</sup> Wartość krytyczna rozkładu t - Studenta dla poziomu istotności 0,05 i dla 9 stopni swobody.

#### Wnioski:

1. Dla zmiennych ( L20-34 ); ( LBZ )  $|t| > t^*$  hipotezę o statycznej nieistotności parametrów strukturalnych należy odrzucić. Oznacza to, że z prawdopodobieństwem 95 % można twierdzić, że obie zmienne uwzględnione w modelu wywierają istotny wpływ na zmienną P\_CNC.

### 2.3.2. Badanie autokorelacji składnika losowego

W celu zbadania autokorelacji rzędu pierwszego posłużono się dwoma metodami:

1. Test Durбина - Watsona.
2. Wyznaczono wartość współczynnika autokorelacji rzędu pierwszego i przeprowadzono test t - Studenta na istotność współczynnika korelacji, gdzie statystyka testowa ma postać:

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-3}}} * \sqrt{n-3}$$

i rozkład t - Studenta o n - 3 stopniach swobody

Obserwacja	$\hat{P}_{CNC}$	reszty ( $u_t$ )	$u_t - u_{t-1}$
1991	2622,206	-30,206	-
1992	2732,797	64,203	94,408
1993	2434,822	178,178	113,975
1994	2524,129	43,871	-134,307
1995	3331,738	-456,738	-500,609
1996	3020,493	3,507	460,245
1997	3390,867	277,133	273,626
1998	3828,063	47,937	-229,196
1999	4354,915	-612,915	-660,852
2000	4945,955	461,045	1073,959
2001	5552,508	-63,508	-524,552
2002	5584,505	87,495	151,003
<b>suma kwadratów</b>		926344,5152	2517849,564

d	2,71804876
d'	1,28195124
obserwacji	12
poziom ist.	0,05
zm.objaśnaj.	2
$d_L$	0,812
$d_U$	1,579
$r_1$	-0,365235
t	-1,1770188
$t^*$	2,26215716

#### Wnioski:

1. Wartość statystyki d' Durбина - Watsona zawiera się w przedziale [  $d_L$  -  $d_U$  ].

Na podstawie testu Durбина - Watsona nie można wyciągnąć wniosków dotyczących autokorelacji.

Moduł statystyki t - Studenta nie przekracza wartości krytycznej, oznacza to, że można przyjąć hipotezę o braku autokorelacji rzędu pierwszego.

### 2.3.3. Badanie jednorodności wariancji składnika losowego test Goldfeldta - Quandta

Celem przeprowadzenia testu Goldfeldta - Quanta próbę podzielono na 2 połowy i oszacowano dla każdej z nich z osobna parametry strukturalne i stochastyczne modelu:

<i>Miary dopasowania</i>	
R	0,839
R <sup>2</sup>	0,704
dopas. R <sup>2</sup>	0,507
Su	129,256

<i>Miary dopasowania</i>	
R	0,927
R <sup>2</sup>	0,859
dopas. R <sup>2</sup>	0,765
Su	470,682

F	13,26036978
obserw. (1)	6
obserw. (2)	6
zm.objaśnaj.	2
poziom ist.	0,05
F* <sup>3)</sup>	9,277

<sup>3)</sup> Wartość krytyczna rozkładu F - Snedecora dla poziomu istotności 0,05 oraz dla 3 i 3 stopni swobody.

#### Wnioski:

1. Wartość statystyki F przekracza wartość krytyczną, nakazuje to odrzucenie hipotezy o jednorodności wariancji. Oznacza to, że klasyczna MNK nie jest właściwą metodą estymacji badanego równania. Zważywszy na niejednorodność wariancji można by szacować model uogólnioną MNK Aitkena. Jednakże warto zwrócić uwagę na radykalną zmianę liczby przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) w 2000 r. Był to rok, w którym nastąpiła nowelizacja kodeksu karnego. Z tej przyczyny wprowadzono zmienną zerojedynową V1, przyjmującą wartość zerową do 1999 r. a począwszy od 2000 r. - wartość równą jedności.



## 2.4. Powtórna estymacja parametrów równania

### 2.4.1. Baza danych niezbędna do oszacowania równania

Rok	P_CNC	L20-34	LBZ	V1
1991	2592	8179,6	9001	0
1992	2797	8057,4	10862	0
1993	2613	7956,6	11013	0
1994	2568	7919,1	11781	0
1995	2875	8131,5	12102	0
1996	3024	8012,7	12422	0
1997	3668	8101,3	12675	0
1998	3876	8212,1	12899	0
1999	3742	8345,2	13174	0
2000	5407	8503,8	13371	1
2001	5489	8659	13664	1
2002	5672	8639,5	14012	1

### 2.4.2. Oszacowanie parametrów strukturalnych oraz parametrów stochastycznych równania

$$P\_CNC = 1,898[L20-34_t] + 0,256[LBZ_t] + 1001,582V1_t - 15310,304 + u_t$$

**(0,614)                      (0,065)                      (337,216)                      (4830,493)**

### Interpretacja :

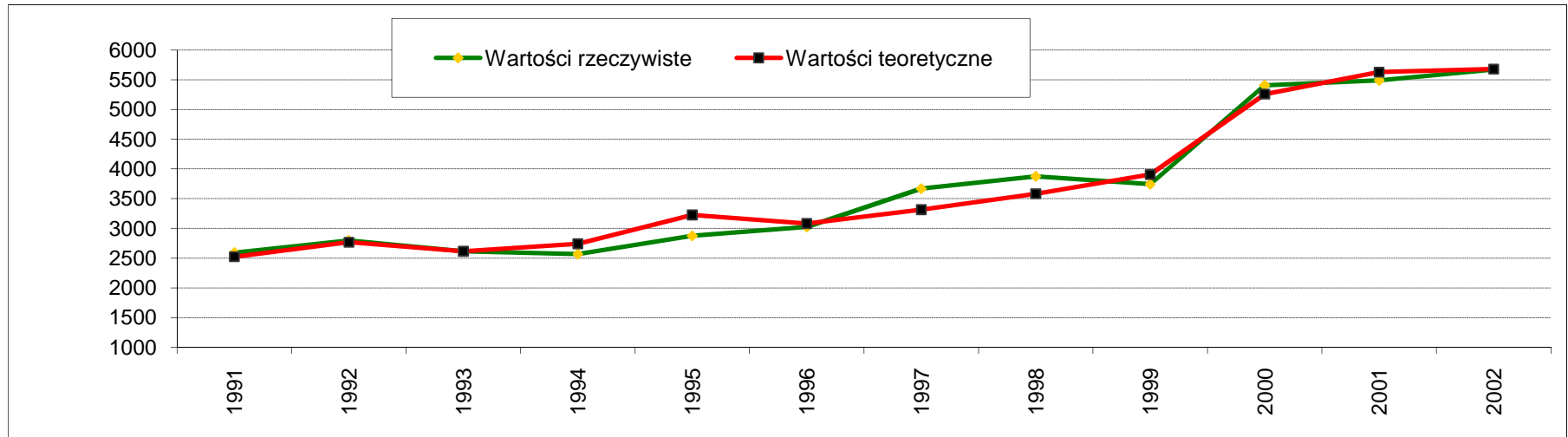
1. Jeżeli liczba ludności w wieku od 20 do 34 lat ( zmienna L20-34 ) wzrośnie o 1000 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko czci i nietykalności cielesnej (zmienna P\_CNC) wzrośnie przeciętnie o 1,898 przypadków.
2. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej L20-34 na podstawie 12-elementowych prób, pomyliliśmy się przeciętnie rzecz biorąc o 0,614.
3. Jeżeli liczba biernych zawodowo ( zmienna LBZ ) wzrośnie o 1000 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) wzrośnie przeciętnie o 0,256 przypadków .
4. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej LBZ na podstawie 12-elementowych prób, pomyliliśmy się przeciętnie rzecz biorąc o 0,065.
5. Nowelizacja kodeksu karnego przy niezmienności pozostałych czynników równania spowodowała wzrost przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) przeciętnie o 1001,493 przypadków .
6. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej V1 na podstawie 12-elementowych prób, pomyliliśmy się przeciętnie rzecz biorąc o 337,216

<i>Miary dopasowania</i>	
R	0,986
R <sup>2</sup>	0,972
dopas. R <sup>2</sup>	0,961
Su	234,666
Vu	6,35%

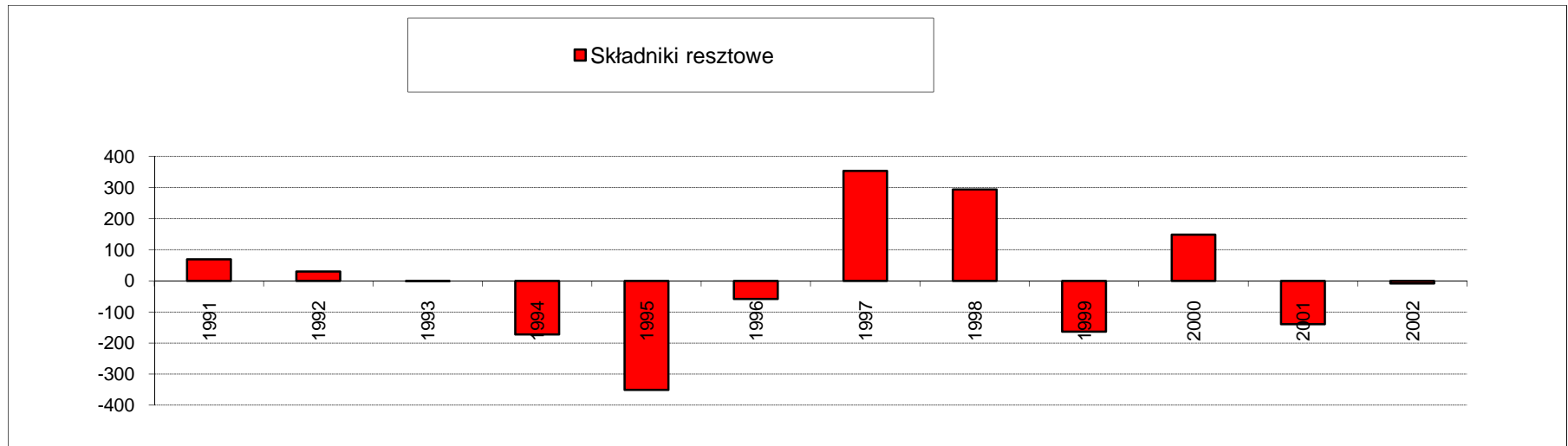
### Wnioski:

1. Ponad 97 % zróżnicowania liczby przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) jest wyjaśniane przez równania.
2. Empiryczna liczba przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) odchylna się przeciętnie o 234,666 przypadków od wartości teoretycznej wyznaczonej na podstawie równania.
3. Odchylenia te stanowią 6,35% przeciętnej liczby przestępstw przeciwko czci i nietykalności cielesnej.

Rys. nr 3 Empiryczne wartości zmiennej P\_CNC a wartości teoretyczne



Rys. nr 4 Składniki resztowe równania modelującego zmienną P\_CNC



## 2.5. Weryfikacja poprawionego równania

### 2.5.1. Badanie istotności parametrów strukturalnych równania

	<i>Współczynniki</i>	<i>Błąd stand.</i>	<i>t</i>
Przecięcie	-15310,304	4830,493	-3,170
L20-34	1,898	0,614	3,090
LBZ	0,256	0,065	3,919
V1	1001,582	337,216	2,970

obserwacje	12
zm. objaśn.	3
poziom ist.	0,050
wartość t*	2,306

#### Wnioski:

1. Dla zmiennych ( L20-34 ); ( LBZ ); ( V1 )  $|t| > t^*$  zatem hipotezę o statycznej nieistotności parametrów strukturalnych należy odrzucić. Oznacza to, że z prawdopodobieństwem 95 % można twierdzić, że zmienne uwzględnione w modelu wywierają istotny wpływ na zmienną P\_CNC.

## 2.5.2. Badanie autokorelacji składnika losowego

Obserwacja	$\hat{P}_{CNC}$	reszty ( $u_t$ )	$u_t - u_{t-1}$
1991	2522,479	69,521	-
1992	2767,199	29,801	-39,721
1993	2614,530	-1,530	-31,331
1994	2740,066	-172,066	-170,536
1995	3225,485	-350,485	-178,419
1996	3081,936	-57,936	292,549
1997	3314,929	353,071	411,007
1998	3582,636	293,364	-59,707
1999	3905,738	-163,738	-457,102
2000	5258,849	148,151	311,889
2001	5628,514	-139,514	-287,665
2002	5680,637	-8,637	130,877
suma kwadratów		440545,5498	727648,4905

d	1,65169865
obserwacji	12
poziom ist.	0,05
zm.objaśnaj.	3
$d_L$	0,812
$d_U$	1,579
$r_1$	0,1697388
t	0,51671439
$t^*$	2,26215716

### Uwaga:

1. W tablicach nie ma wartości krytycznych testu Durбина - Watsona dla 3 zmiennych objaśniających. W tabeli podano wartości krytyczne dla dwóch zmiennych i zaznaczono je kolorem . Wartość statystyki d nieznacznie przekracza  $d_U$  dla dwóch zmiennych. Jednakże wraz ze wzrostem liczby zmiennych objaśniających rośnie wartość  $d_L$  i maleje  $d_U$ . Nie wiadomo zatem, czy dla trzech zmiennych objaśniających wartość d byłaby większa od  $d_U$ . Nie można zatem na podstawie testu Durбина - Watsona wyciągnąć jednoznacznych wniosków. Wyniki testu t - Studenta pozwalają przyjąć hipotezę o braku autokorelacji.

### 2.5.3. Badanie jednorodności wariancji składnika losowego test Goldfeldta - Quandta

Celem przeprowadzenia testu Goldfeldta - Quandta próbę podzielono na 2 połowy i oszacowano dla każdej z nich z osobna parametry strukturalne i stochastyczne modelu:

<i>Miary dopasowania</i>	
R	0,839
R <sup>2</sup>	0,704
dopas. R <sup>2</sup>	0,507
Su	129,256

<i>Miary dopasowania</i>	
R	0,998
R <sup>2</sup>	0,995
dopas. R <sup>2</sup>	0,989
Su	103,981

Wyznaczenie wartości statystyki F - Snedecora

F	1,54522723
obserw. (1)	6
obserw. (2)	6
ilość zm. objaśniających (1) <sup>5)</sup>	2
ilość zm. objaśniających (2) <sup>5)</sup>	3
poziom ist.	0,05
F* <sup>4)</sup>	19,164

4) Wartość krytyczna rozkładu F - Snedecora dla poziomu istotności 0,05 dla dwóch i trzech swobody.

5) Różna liczba zmiennych objaśniających wynika z tego, że w pierwszych 6-ciu latach ( 1991 -1996 ) nie uwzględniono zmiennej V1, która przyjmuje wartość 0

#### Wniosek:

1. Wartość statystyki F nie przekracza wartości krytycznej F\* w związku z tym nie ma podstaw do odrzucenia hipotezy o jednorodności wariancji składnika losowego.

#### 2.5.4. Badanie symetrii reszt

m/n	0,416666667
t	-0,560611911
n	12
poziom ist.	0,05
t*	2,200985159

**Wnioski:**

1.  $|t| < t^*$  co pozwala przyjąć hipotezę o symetrii reszt.

#### 2.5.5. Badanie losowości reszt

K	6
liczba +	5
liczba -	7
poziom ist.	0,05
K* <sup>6)</sup>	3

<sup>6)</sup> Wartość krytyczna rozkładu serii dla lewostronnego obszaru krytycznego i poziomu istotności 0,05.

**Wnioski:**

1.  $K > K^*$  co oznacza, że nie ma podstaw do odrzucenia hipotezy o losowości składnika resztowego ut.

**Wniosek końcowy:**

Wyniki testów Durбина - Watsona i Goldfeldta - Quandta świadczą o tym, że klasyczna MNK jest właściwą metodą estymacji tego równania. Oznacza to, że parametry równania ze zmienną V1 można szacować klasyczną MNK. Ponadto składnik resztowy wykazuje losowość i symetrię. Wszystkie zmienne objaśniające istotnie wpływają na P\_CNC, a parametry stochastyczne świadczą o bardzo dobrym dopasowaniu równania do empirycznych danych. Równanie uzupełnione o zmienną V1 można zatem uznać za postać ostateczną.

## 2.6. Prognoza zmiennych egzogenicznych ( objaśniających )

Rok	L15-19	L20-34	LBZ	t
1991	2998,6	8179,6	9001	
1992	3067,7	8057,4	10862	
1993	3139	7956,6	11013	
1994	3194,9	7919,1	11781	1
1995	3236,8	8131,5	12102	2
1996	3244,8	8012,7	12422	3
1997	3292,5	8101,3	12675	4
1998	3343,1	8212,1	12899	5
1999	3365,7	8345,2	13174	6
2000	3356,9	8503,8	13371	7
2001	3322,3	8659	13664	8
2002	3231,1	8639,5	14012	9

Rok	^L15-19 (prognoza)	^L20-34 (prognoza)	^LBZ (prognoza)
1991			
1992			
1993			
1994			
1995			
1996			
1997			
1998			
1999			
2000			
2001			
2002	3231,1	8639,5	14012
2003	3104,8425	8831,98	14233,917
2004	2978,585	9010,56	14500,7
2005	2852,3275	9169,94	14767,4833
2006	2726,07	9284,72	15034,2667

### 2.6.1. Prognoza zmian poziomu ludności w wieku od 20 do 34 roku życia ( zmienna L20-34 )

Uwaga:

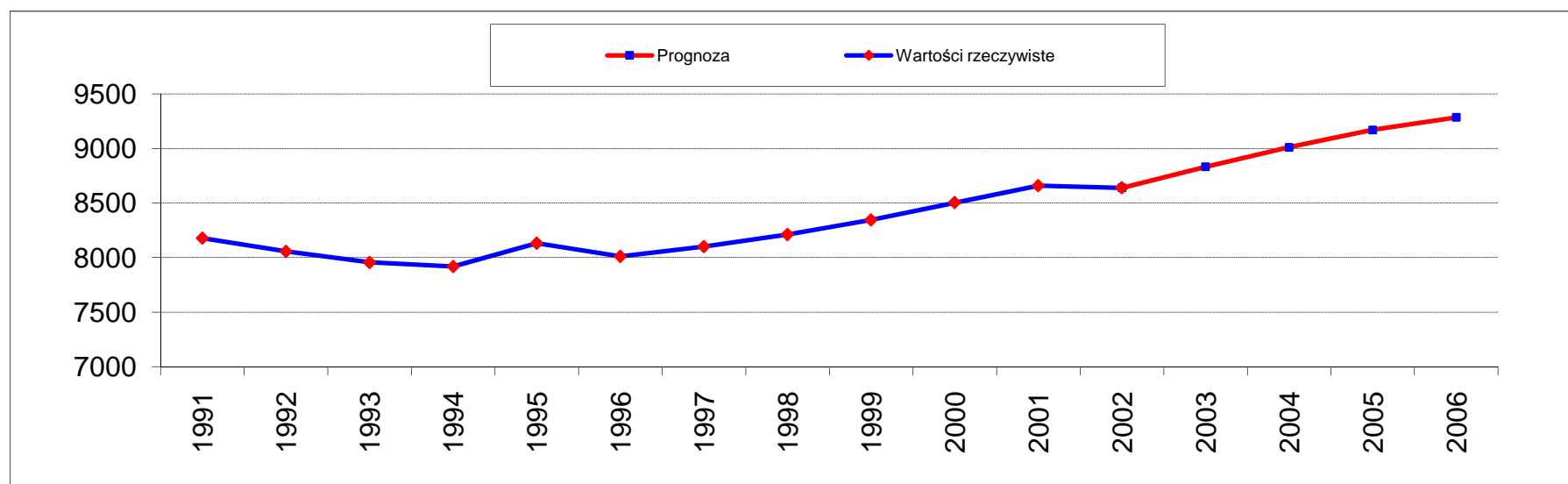
- Prognozując poziom ludności w latach 2003 -2006 posługiwano się:
  - stanem faktycznym ludności w wieku 16,17, 18 i 19 lat.
  - z braku danych statystycznych jako liczbę ludności w wieku 31, 32, 33 i 34 lata przyjęto 1/5 liczby ludności w wieku 30 - 34 lat
- Do liczby ludności w wieku od 20 do 34 lat w roku 2002 dodano osoby mające 19 lat i odjęto mające 34 lata <sup>1)</sup>.
- Analogicznie w następnych latach dodawano liczby odpowiednio 18, 17, i 16 - latków odejmując jednocześnie oszacowane liczby 33, 32, 31- latków.



1) Nie uwzględniono ruchu naturalnego i wędrownego ludności . Jednak podejście takie wydaje się uprawnione zważywszy na to, że liczba 16, 17 , 18 i 19 - latków w 2002 była bardzo zbliżona do liczby odpowiednio 15, 16, 17,18 - latków w 2001 r. ( różnice poniżej 1 % ).

	16	17	18	19	20-34	30-34	31	32	33	34
2002	615,10	659,70	678,90	692,80	<b>8639,50</b>	2501,60	500,32	500,32	500,32	500,32
2003					<b>8831,98</b>					
2004					<b>9010,56</b>					
2005					<b>9169,94</b>					
2006					<b>9284,72</b>					

Rys. nr 1 Wartości empiryczne i prognozowane zmiennej L 20-34 na lata 2003 - 2006



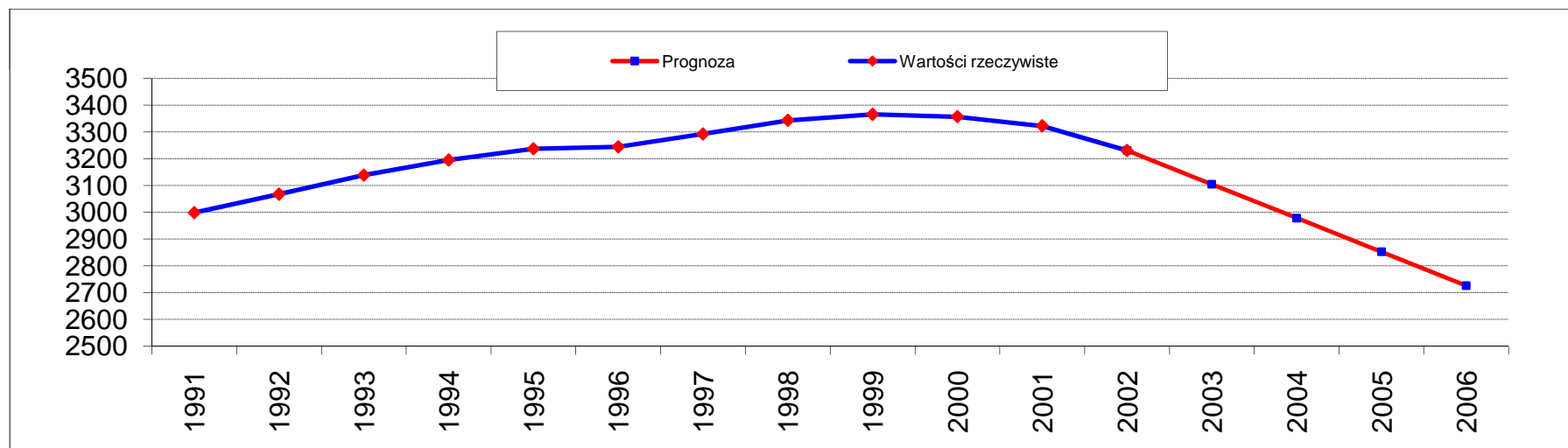
## 2.6.2. Prognoza zmian poziomu ludności w wieku od 15 do 19 roku życia ( zmienna L15-19 )

### Uwaga:

1. Analogicznie prognozowano liczbę ludności w przedziale wieku od 15 do 19 lat. W tym przypadku za liczbę 21, 22, 23 i 24 - latków przyjęto 1/5 liczby ludności w wieku od 20 do 24 lat, a za 11, 12, 13, 14, 1/8 liczby ludności w wieku od 7 do 14 lat.

	11	12	13	14	07-14	15-19	20-24	21	22	23	24
2002	513,16	513,16	513,16	513,16	4105,30	<b>3231,10</b>	3197,10	639,42	639,42	639,42	639,42
2003						<b>3104,84</b>					
2004						<b>2978,59</b>					
2005						<b>2852,33</b>					
2006						<b>2726,07</b>					

Rys. nr 2 Wartości empiryczne i prognozowane zmiennej L 15-19 na lata 2003 - 2006



### 2.6.3. Prognoza zmian poziomu ludności biernej zawodowo ( zmienna LBZ )

Uwaga:

1. W celu zaprognozowania liczby biernych zawodowo ( LBZ ) posłużono się trendem liniowym liczonym dla 9 ostatnich obserwacji. Jest to okres , w którym liczba biernych zawodowo wykazuje stałą tendencję wzrostową a dopasowanie funkcji trendu jest niemal idealne.

#### 1. Klasyczną MNK oszacowano parametry trendu

<i>Miary dopasowania</i>	
R	0,998
R <sup>2</sup>	0,997
dopas. R <sup>2</sup>	0,996
Su	46,030
Vu	0,36%

<i>Współczynniki</i>	
t	266,783
Prze- cięcie	11566,083

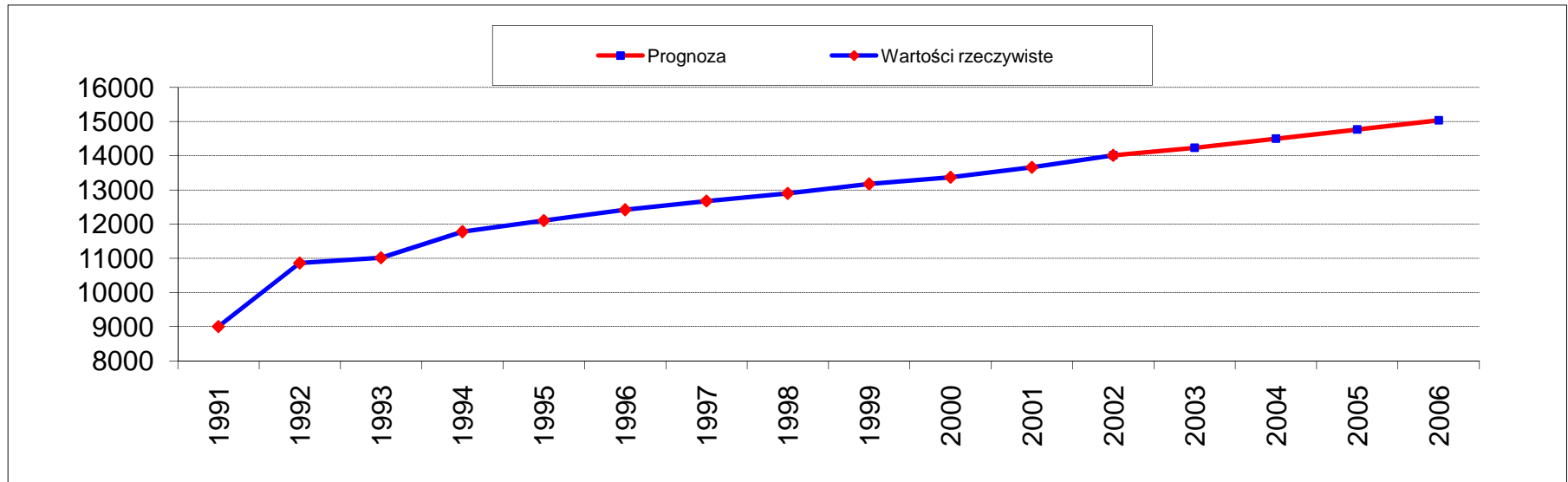
$$\hat{LBZ} = 266,783t + 11566,083$$

gdzie t = 1,2, 3... , począwszy od 1994

#### 2. Następnie na podstawie oszacowanej funkcji trendu postawiono prognozy LBZ na lata 2003 - 2006

rok	t	$\hat{LBZ}$
2003	10	14233,917
2004	11	14500,70
2005	12	14767,483
2006	13	15034,267

Rys. nr 3 Wartości empiryczne i prognozowane zmiennej LBZ na lata 2003 - 2006



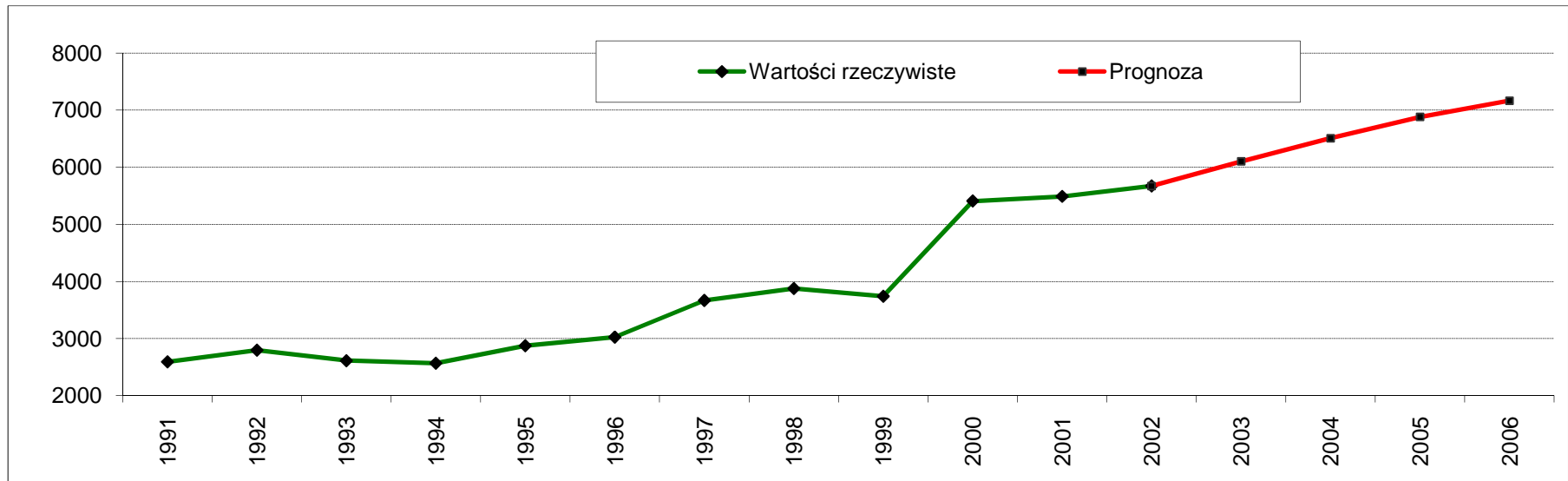
**2.7. Obliczenie prognozy dla przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_CNC ) na lata 2003 - 2006**

<i>Współczynniki</i>	
Przecięcie	-15310,304
L20-34	1,898
LBZ	0,256
V1	1001,582

<i>Miary dopasowania</i>	
R	0,986
R <sup>2</sup>	0,972
dopas. R <sup>2</sup>	0,961
Su	234,666
Vu	6,35%

rok	<sup>^</sup> L20-34 ( prognoza )	<sup>^</sup> LBZ ( prognoza )	V1	P_CNC	<sup>^</sup> P_CNC (prognoza)
1991				2592	
1992				2797	
1993				2613	
1994				2568	
1995				2875	
1996				3024	
1997				3668	
1998				3876	
1999				3742	
2000				5407	
2001				5489	
2002				5672	
2003	8831,98	14233,917	1,0		6102,862
2004	9010,56	14500,700	1,0		6510,193
2005	9169,94	14767,483	1,0		6881,078
2006	9284,72	15034,267	1,0		7167,298

Rys. nr 5 Wartości empiryczne i prognozowane zmiennej P\_CNC



### 2.7.1. Obliczenie średniego błędu predykcji

X =

L20-34	LBZ	V1	
8179,6	9001	0,0	1,0
8057,4	10862	0,0	1,0
7956,6	11013	0,0	1,0
7919,1	11781	0,0	1,0
8131,5	12102	0,0	1,0
8012,7	12422	0,0	1,0
8101,3	12675	0,0	1,0
8212,1	12899	0,0	1,0
8345,2	13174	0,0	1,0
8503,8	13371	1,0	1,0
8659	13664	1,0	1,0
8639,5	14012	1,0	1,0

$X^T =$ 

<b>L20-34</b>	8179,6	8057,4	7956,6	7919,1	8131,5	8012,7	8101,3	8212,1	8345,2
<b>LBZ</b>	9001	10862	11013	11781	12102	12422	12675	12899	13174
<b>V1</b>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0

8503,8	8659	8639,5
13371	13664	14012
1,0	1,0	1,0
1,0	1,0	1,0

 $X^T X =$ 

812817699,7	1211635264	25802,3	98717,8
1211635264	1822262446	41047	146976
25802,3	41047,0	3,0	3,0
98717,8	146976,0	3,0	12,0

 $(X^T X)^{-1} =$ 

0,0000069	0,00000	-0,00305	-0,05324
-0,0000002	0,00000	-0,00005	0,00066
-0,0030494	-0,00005	2,06498	25,20199
-0,0532412	0,00066	25,20199	423,72306

 $X_T =$ 

	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>
<b>^L20-34</b>	8831,98	9010,56	9169,94	9284,72
<b>^LBZ</b>	14233,91667	14500,7	14767,48333	15034,26667
<b>V1</b>	1,0	1,0	1,0	1,0
	1,0	1,0	1,0	1,0

 $X_T^T =$ 

	<b>^L20-34 (proгноza)</b>	<b>^LBZ (proгноza)</b>	<b>V1</b>	
<b>2003</b>	8831,98	14233,917	1,0	1,0
<b>2004</b>	9010,56	14500,7	1,0	1,0
<b>2005</b>	9169,94	14767,483	1,0	1,0
<b>2006</b>	9284,72	15034,267	1,0	1,0

$$X_T^T (X^T X)^{-1} =$$

<b>2003</b>	0,0014776	-0,0000020	-0,4002097	-11,947647
<b>2004</b>	0,0026498	-0,0000159	-0,9585463	-21,280170
<b>2005</b>	0,0036903	-0,0000261	-1,4583346	-29,590460
<b>2006</b>	0,0044252	-0,0000277	-1,8221201	-35,526190

$$X_T^T (X^T X)^{-1} X^T =$$

	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>
<b>2003</b>	0,673868749	0,937206648	1,172173839	1,341238239
<b>2004</b>	0,937206654	1,406151878	1,824221467	2,124111195
<b>2005</b>	1,172173848	1,824221472	2,405414876	2,822019832
<b>2006</b>	1,341238247	2,124111119	2,822019814	3,322563407

## 2.7.2. Prognoza dla zmiennej P\_CNC na lata 2003 - 2006 i średnie błędy predykcji

rok	prognoza	śr. błąd	śr. wzgl. błąd
<b>2003</b>	6102,862	303,607	4,97%
<b>2004</b>	6510,193	364,009	5,59%
<b>2005</b>	6881,078	433,047	6,29%
<b>2006</b>	7167,298	487,889	6,81%



### **3. PRZESTĘPSTWA PRZECIWKO MIENIU SPOŁECZNEMU I PRYWATNEMU**

- 3.1. Dobór zmiennych objaśniających**
- 3.2. Estymacja parametrów równania**
- 3.3. Weryfikacja równania**
- 3.4. Powtórna estymacja równania**
- 3.5. Weryfikacja poprawionego równania**
- 3.6. Obliczenie prognozy dla przestępstw przeciwko mieniu społecznemu i prywatnemu**

**3.1. Dobór zmiennych objaśniających ( egzogenicznych ) dla równania modelu , w którym zmienną objaśnianą stanowią przestępstwa przeciwko mieniu społecznemu i prywatnemu ( P\_PM )**

**3.1.1. Baza danych niezbędna do wytypowania zmiennych równania**

P_PM	L	L15-19	L20-34	L35-50	LB	LBZ	SB	WBZ	B_W	B_SO	B_SZ
620705	38309	2998,6	8179,6	8090	2155,6	9001	12,2	37,81	64,1	155,5	482,5
591250	38418	3067,7	8057,4	8387,4	2509,3	10862	14,3	38,3	56,6	177,9	527,9
561001	38505	3139	7956,6	8675,5	2889,6	11013	16,4	38,8	52,1	193,2	581,5
582539	38581	3194,9	7919,1	8913,7	2838	11781	16	39,1	47,6	194,4	570,2
715355	38609	3236,8	8131,5	9114,2	2628,8	12102	14,9	41,6	38,9	188,2	531,7
613900	38639	3244,8	8012,7	9152	2359,5	12422	13,2	42,4	31,4	151,7	471
676836	38660	3292,5	8101,3	9130,5	1826,4	12675	10,3	42,6	25,1	115,2	364,2
747015	38667	3343,1	8212,1	9060,8	1831,4	12899	10,4	42,9	30,2	111,2	369
803977	38654	3365,7	8345,2	8974,1	2349,8	13174	13,1	43,3	47,7	142	483
925433	38644	3356,9	8503,8	8835,9	2702,6	13371	15,1	43,6	69,4	168,7	561,9
904623	38632	3322,3	8659	8678,9	3115,1	13664	17,5	44,2	100,5	192,1	666
890085	38219	3231,1	8639,5	8371,4	3217	14012	18	44,81	126,7	199	683,6

B_Z	B_P	DRG	WRG
805,6	663,3	618,64	566,03
964,6	782,4	631,16	620,39
1131,2	931,5	573,80	559,58
1118,3	907,5	577,31	546,51
1025	845	587,93	540,51
907,7	797,6	625,55	572,64
700,7	621,3	672,73	607,59
698	622,8	664,14	638,86
898,2	778,9	663,34	650,71
998,8	903,8	656,33	644,48
1142,5	1014	656,73	621,30
1163,6	1044,1	664,21	624,99

### 3.1.2. Macierz korelacji zmiennych

	P_PM	L	L15-19	L20-34	L35-50	LB	LBZ	SB	WBZ	B_W
P_PM	1,000	0,087	0,681	0,950	0,016	0,322	0,773	0,297	0,850	0,602
L	0,087	1,000	0,699	-0,157	0,872	-0,350	0,369	-0,358	0,305	-0,622
L15-19	0,681	0,699	1,000	0,465	0,703	-0,045	<b>0,885</b>	-0,071	<b>0,857</b>	-0,051
L20-34	0,950	-0,157	0,465	1,000	-0,253	0,406	0,638	0,385	0,745	0,778
L35-50	0,016	0,872	0,703	-0,253	1,000	-0,328	0,476	-0,338	0,398	-0,644
LB	0,322	-0,350	-0,045	0,406	-0,328	1,000	0,272	0,999	0,141	0,760
LBZ	0,773	0,369	<b>0,885</b>	0,638	0,476	0,272	1,000	0,247	<b>0,940</b>	0,310
SB	0,297	-0,358	-0,071	0,385	-0,338	0,999	0,247	1,000	0,111	0,751
WBZ	0,850	0,305	<b>0,857</b>	0,745	0,398	0,141	<b>0,940</b>	0,111	1,000	0,333
B_W	0,602	-0,622	-0,051	0,778	-0,644	0,760	0,310	0,751	0,333	1,000
B_SO	0,052	-0,437	-0,326	0,154	-0,409	0,935	-0,034	0,944	-0,170	0,618
B_SZ	0,371	-0,445	-0,107	0,493	-0,461	0,985	0,223	0,983	0,131	0,844
B_Z	0,190	-0,295	-0,089	0,257	-0,262	0,985	0,199	0,988	0,034	0,646
B_P	0,402	-0,251	0,095	0,456	-0,215	0,987	0,401	0,982	0,269	0,742
DRG	0,673	0,104	0,557	0,665	0,045	-0,288	0,591	-0,312	0,713	0,237
WRG	0,681	0,120	0,561	0,654	-0,026	-0,121	0,579	-0,141	0,606	0,281

B_SO	B_SZ	B_Z	B_P	DRG	WRG
0,052	0,371	0,190	0,402	0,673	0,681
-0,437	-0,445	-0,295	-0,251	0,104	0,120
-0,326	-0,107	-0,089	0,095	0,557	0,561
0,154	0,493	0,257	0,456	0,665	0,654
-0,409	-0,461	-0,262	-0,215	0,045	-0,026
0,935	0,985	0,985	0,987	-0,288	-0,121
-0,034	0,223	0,199	0,401	0,591	0,579
0,944	0,983	0,988	0,982	-0,312	-0,141
-0,170	0,131	0,034	0,269	0,713	0,606
0,618	0,844	0,646	0,742	0,237	0,281
1,000	0,921	0,957	0,877	-0,546	-0,383
0,921	1,000	0,948	0,959	-0,215	-0,067
0,957	0,948	1,000	0,966	-0,425	-0,235
0,877	0,959	0,966	1,000	-0,209	-0,052
-0,546	-0,215	-0,425	-0,209	1,000	0,884
-0,383	-0,067	-0,235	-0,052	0,884	1,000

### 3.1.3. Zasady doboru zmiennych do równania

1. Wektor współczynników korelacji zmiennej objaśnianej z potencjalnymi zmiennymi objaśniającymi zaznaczono kolorem
2. Wybór najsilniej skorelowanej potencjalnej zmiennej egzogenicznej ze zmienna endogeniczną (objaśnianą) został zaznaczony kolorem . Zmienna została automatycznie zakwalifikowana do modelu. Ponadto przyjęto maksymalną wartość modułu współczynnika korelacji między zmiennymi egzogenicznymi na poziomie 0,8 wyższe współczynniki pomiędzy potencjalnymi zmiennymi objaśniającymi zaznaczono kolorem zakwalifikować jedną z nich.
3. W następnej kolejności rozważano zasadność doboru drugiej (obok L\_20\_34) zmiennej objaśniającej. W celu wstępnej selekcji pozostałych zmiennych zastosowano następującą procedurę:
  - a) 

	>	
--	---	--

 Dobór zmiennej może posiadać uzasadnienie.
  - b) 

	<	
--	---	--

 Dobór zmiennej nie posiada uzasadnienia.

W świetle powyższych kryteriów można rozważyć dobór zmiennych oznaczonych kolorem  
 Należy zatem rozważyć następujące pary cech:



A L20-34 L15-19

B L20-34 LBZ

C L20-34 WBZ

D L20-34 DRG

E L20-34 WRG

Wektor modułów współczynnika korelacji zmiennej objaśnianej ( endogenicznej ) z badanymi zmiennymi

	<i>P_PM</i>
<b>L20-34</b>	0,950
<b>L15-19</b>	0,681
<b>LBZ</b>	0,773
<b>WBZ</b>	0,850
<b>DRG</b>	0,673
<b>WRG</b>	0,681

Macierz modułów współczynników korelacji między badanymi zmiennymi

	<b>L20-34</b>	<b>L15-19</b>	<b>LBZ</b>	<b>WBZ</b>	<b>DRG</b>	<b>WRG</b>
<b>L20-34</b>	1,000	0,465	0,638	0,745	0,665	0,654
<b>L15-19</b>	0,465	1,000	0,885	0,857	0,557	0,561
<b>LBZ</b>	0,638	0,885	1,000	0,940	0,591	0,579
<b>WBZ</b>	0,745	0,857	0,940	1,000	0,713	0,606
<b>DRG</b>	0,665	0,557	0,591	0,713	1,000	<b>0,884</b>
<b>WRG</b>	0,654	0,561	0,579	0,606	<b>0,884</b>	1,000

**A**

h(A,L20-34)	0,615738571
h(A,L15-19)	0,316185583
<b>H(A)</b>	<b>0,931924154</b>

**B**

h(B,L20-34)	0,550618161
h(B,L20-34)	0,365200109
<b>H(B)</b>	<b>0,91581827</b>

**C**

h(C,L20-34)	0,51674753
h(C,WBZ)	0,41384389
<b>H(C)</b>	<b>0,93059143</b>

**D**

h(D,L20-34)	0,541
<b>h(D, DRG)</b>	<b>0,272</b>
<b>H(D)</b>	<b>0,813</b>

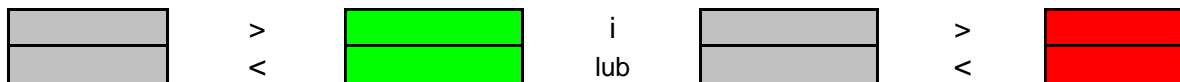
**E**

h(E,L20-34)	0,545
h(E, WRG)	0,280
<b>H(E)</b>	<b>0,826</b>

**Wniosek:**

Z kombinacje A, B, C, D i E najkorzystniejszą wydaje się kombinacja A, ponieważ ma największy wskaźnik pojemności indywidualnej nośnika informacji H(A).

4. Nie ma jednak pewności, czy nie jest możliwa poprawa jakości równania poprzez włączenie trzeciej zmiennej. Jako potencjalną trzecią zmienną rozważono każdą z cech oznaczonych wcześniej kolorem [ ]



Dobór zmiennej może mieć uzasadnienie. Dobór zmiennej nie ma uzasadnienia, ponieważ w większym stopniu "powiela" ona jedną ze zmiennych egzogenicznych niż opisuje zmienną P\_PM.

Odpowiednie wiersze i kolumny macierzy korelacji kształtują się następująco:

	P_PM	L20-34	L15-19
LBZ	0,773	0,638	0,885
WBZ	0,850	0,745	0,857
DRG	0,673	0,665	0,557
WRG	0,681	0,654	0,561

Można zatem sprawdzić zasadność włączenia do równania jednej ze zmiennych oznaczonych na   
Pozostałe cechy są bardzo silnie skorelowane z L15-19.

Obliczono zatem wskaźnik pojemności indywidualnej nośnika informacji dla dwóch kombinacji:

F

G

h(F,L20-34)	0,423372676
h(F,L15-19)	0,229088561
h(F,DRG)	0,203788977
<b>H(F)</b>	<b>0,856250213</b>

h(G,L20-34)	0,42565458
h(G,L15-19)	0,22863594
h(G,WRG)	0,20931512
<b>H(G)</b>	<b>0,86360564</b>

**Wniosek:**

1. Dodanie jednej ze zmiennych **DRG** lub **WRG** nie poprawi jakości modelu, gdyż wskaźnik pojemności indywidualnej nośnika informacji H(A) dla kombinacji A jest wyższy od wskaźników H(F), H(G).

### 3.2. Estymacja parametrów równania opisującego zmienną P\_PM

#### 3.2.1. Baza danych niezbędna do oszacowania równania

rok	P_PM	L15-19	L20-34
1991	620705	2998,6	8179,6
1992	591250	3067,7	8057,4
1993	561001	3139	7956,6
1994	582539	3194,9	7919,1
1995	715355	3236,8	8131,5
1996	613900	3244,8	8012,7
1997	676836	3292,5	8101,3
1998	747015	3343,1	8212,1
1999	803977	3365,7	8345,2
2000	925433	3356,9	8503,8
2001	904623	3322,3	8659
2002	890085	3231,1	8639,5

#### 3.2.2. Klasyczną MNK za pomocą MC Excel oszacowano parametry strukturalne oraz parametry stochastyczne modelu

$$P\_PM = 350,308[L15-19_t] + 423,079[L20-34_t] - 389326,444 + u_t$$

**(68,538)                      (31,269)                      (249691,264)**

Interpretacja :

1. Jeżeli liczba ludności w wieku od 15 do 19 lat ( zmienna L15-19 ) wzrośnie o 1000 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko mieniu społecznemu i prywatnemu ( zmienna P\_PM ) wzrośnie przeciętnie o 350,308 przypadków.
2. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej L15-19 na podstawie 12-elementowych prób, pomyliliibyśmy się przeciętnie rzecz biorąc o 68,538.



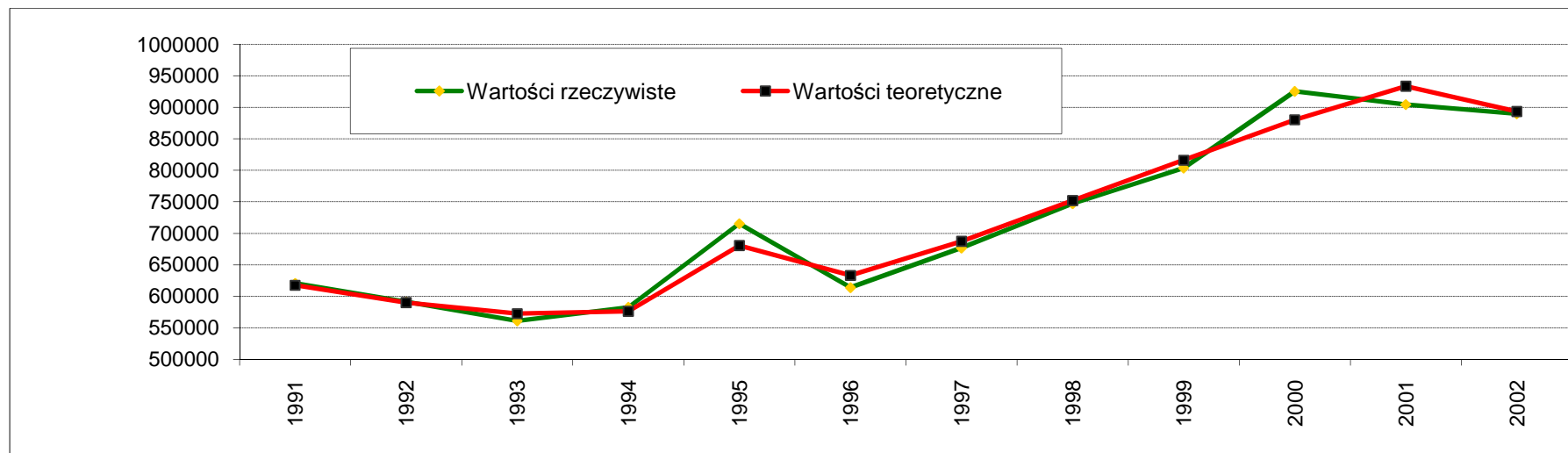
3. Jeżeli liczba ludności w wieku od 20 do 34 lat ( zmienna L20-34 ) wzrośnie o 100 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko mieniu społecznemu i prywatnemu ( zmienna P\_PM ) wzrośnie przeciętnie o 423,079 przypadków.
4. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej L20-34 na podstawie 12-elementowych prób, pomylibyśmy się przeciętnie rzecz biorąc o 31,269.

<i>Miary dopasowania</i>	
R	0,987
R <sup>2</sup>	0,975
dopas. R <sup>2</sup>	0,969
Su	23451,610
Vu	3,26%

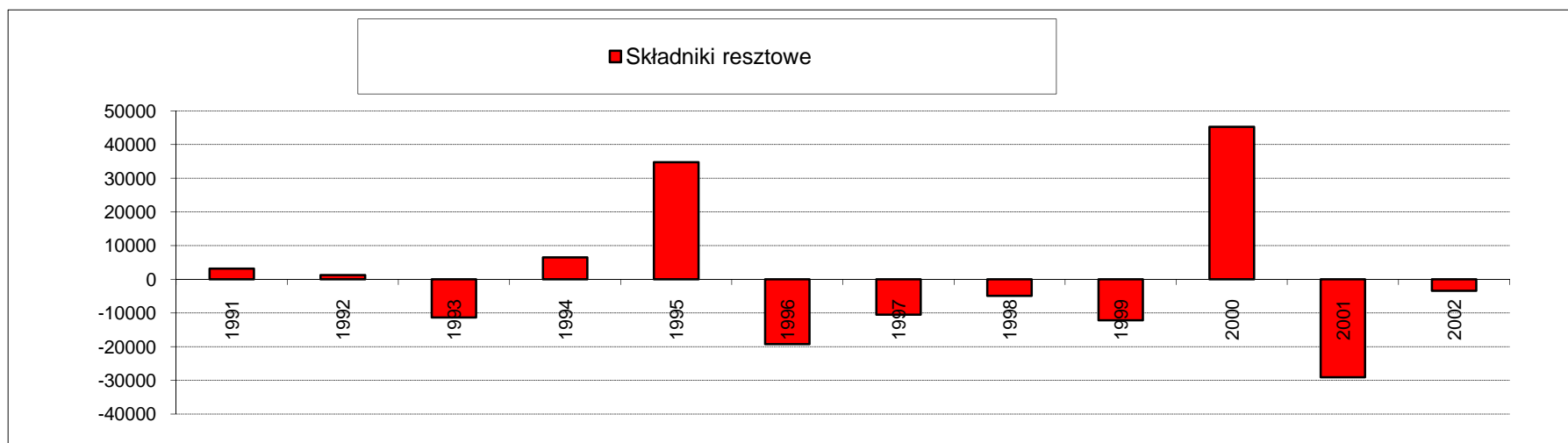
**Wnioski:**

1. Ponad 97 % różnicowania liczby przestępstw przeciwko mieniu społecznemu i prywatnemu ( zmienna P\_PM ) jest wyjaśniane przez model.
2. Empiryczna liczba przestępstw przeciwko czci i nietykalności cielesnej ( zmienna P\_PM ) odchyła się przeciętnie o 23451,610 przypadków od wartości teoretycznej wyznaczonej na podstawie modelu.
3. Odchylenia te stanowią 3,26% przeciętnej liczby przestępstw przeciwko mieniu społecznemu i prywatnemu.

Rys. nr 1 Empiryczne wartości zmiennej P\_PM a wartości teoretyczne



Rys. nr 2 Składniki resztowe równania modelującego zmienną P\_CNC



### 3.3. Weryfikacja równania

#### 3.3.1. Badanie istotności parametrów strukturalnych równania

	<i>Współczynniki</i>	<i>Błąd stand.</i>	<i>t Stat</i>
Przecięcie	-3893526,444	249691,264	-15,593
L15-19	350,308	68,538	5,111
L20-34	423,079	31,269	13,530

obserw.	12
zm. objaśnij.	2
poziom ist.	0,05
wartość t* <sup>1)</sup>	2,262

1) Wartość krytyczna rozkładu t - Studenta dla poziomu istotności 0,05 i dla 9 stopni swobody.

#### Wnioski:

1. Dla zmiennych ( L15-19 ); ( L20-34 )  $|t| > t^*$  hipotezę o statycznej nieistotności parametrów strukturalnych należy odrzucić. Oznacza to, że z prawdopodobieństwem 95 % można twierdzić, że obie zmienne uwzględnione w modelu wywierają istotny wpływ na zmienną P\_PM.

#### 3.3.2. Badanie autokorelacji składnika losowego

W celu zbadania autokorelacji rzędu pierwszego posłużono się dwoma metodami:

1. Test Durbina - Watsona.
2. Wyznaczono wartość współczynnika autokorelacji rzędu pierwszego i przeprowadzono test t - Studenta na istotność współczynnika korelacji, gdzie statystyka testowa ma postać:

$$t = \frac{r}{\sqrt{1 - r^2}} * \sqrt{n - 3}$$

i rozkład t - Studenta o n - 3 stopniach swobody

Obserwacja	$\hat{P}_M$	reszty ( $u_t$ )	$u_t - u_{t-1}$
1991	617521,668	3183,332	-
1992	590027,706	1222,294	-1961,038
1993	572358,307	-11357,307	-12579,601
1994	576075,056	6463,944	17821,251
1995	680614,886	34740,114	28276,170
1996	633155,588	-19255,588	-53995,702
1997	687350,048	-10514,048	8741,539
1998	751952,751	-4937,751	5576,298
1999	816181,493	-12204,493	-7266,742
2000	880199,081	45233,919	57438,412
2001	933740,261	-29117,261	-74351,180
2002	893542,157	-3457,157	25660,104
<b>suma kwadratów</b>		4949802278	13840791829

d	2,79623125
d'	1,20376875
obserwacji	12
poziom ist.	0,05
zm.objaśnaj.	2
$d_L$	0,812
$d_U$	1,579
$r_1$	-0,4011207
t	-1,3136779
$t^*$	2,26215716

### Wnioski:

1. Wartość statystyki  $d'$  Durбина -Watsona zawiera się w przedziale  $[d_L - d_U]$ .

Na podstawie testu Durбина - Watsona nie można wyciągnąć wniosków dotyczących autokorelacji.

Moduł statystyki  $t$  - Studenta nie przekracza wartości krytycznej, oznacza to, że można przyjąć hipotezę o braku autokorelacji rzędu pierwszego.

### 3.3.3. Badanie jednorodności wariancji składnika losowego test Goldfeldta - Quandta

Celem przeprowadzenia testu Goldfeldta - Quanta próbę podzielono na 2 połowy i oszacowano dla każdej z nich z osobna parametry strukturalne i stochastyczne modelu:

<i>Miary dopasowania</i>	
R	0,960
R <sup>2</sup>	0,922
dopas. R <sup>2</sup>	0,870
Su	19488,054

<i>Miary dopasowania</i>	
R	0,970
R <sup>2</sup>	0,940
dopas. R <sup>2</sup>	0,900
Su	31275,938

F	2,575631521
obserw. (1)	6
obserw. (2)	6
zm.objaśnaj.	2
poziom ist.	0,05
F* <sup>2)</sup>	9,277

2) Wartość krytyczna rozkładu F - Snedecora dla poziomu istotności 0,05 oraz dla 3 i 3 stopni swobody.

#### Wnioski:

1. Wyniki powyższych testów nie wskazują ani na autokorelację, ani na niejednorodność wariancji. Warto zatem zwrócić uwagę na stosunkowo wysoki (co do modułu) współczynnik autokorelacji i "rozregulowanie się modelu od 2000 r.". W celu wyeliminowania tej nieprawidłowości wprowadzono zmienną zerojedynkową V2 przyjmującą jednorazowo w 2000 roku wartość równą 1.

### 3.4. Powtórna estymacja parametrów równania

#### 3.4.1. Baza danych niezbędna do oszacowania równania

rok	P_PM	L15-19	L20-34	V2
1991	620705	2998,6	8179,6	0
1992	591250	3067,7	8057,4	0
1993	561001	3139	7956,6	0
1994	582539	3194,9	7919,1	0
1995	715355	3236,8	8131,5	0
1996	613900	3244,8	8012,7	0
1997	676836	3292,5	8101,3	0
1998	747015	3343,1	8212,1	0
1999	803977	3365,7	8345,2	0
2000	925433	3356,9	8503,8	1
2001	904623	3322,3	8659	0
2002	890085	3231,1	8639,5	0

#### 3.4.2. Oszacowanie parametrów strukturalnych oraz parametrów stochastycznych równania

$$P\_PM = 317,652[L15-19_t] + 407,379[L20-34_t] + 58517,199V2_t - 3663682,317 + u_t$$

**(50,741)                      (23,206)                      (19297,300)                      (195898,940)**

Interpretacja :

1. Jeżeli liczba ludności w wieku od 15 do 19 lat ( zmienna L15-19 ) wzrośnie o 1000 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko mieniu społecznemu i prywatnemu ( zmienna P\_PM ) wzrośnie przeciętnie o 317,652 przypadków.

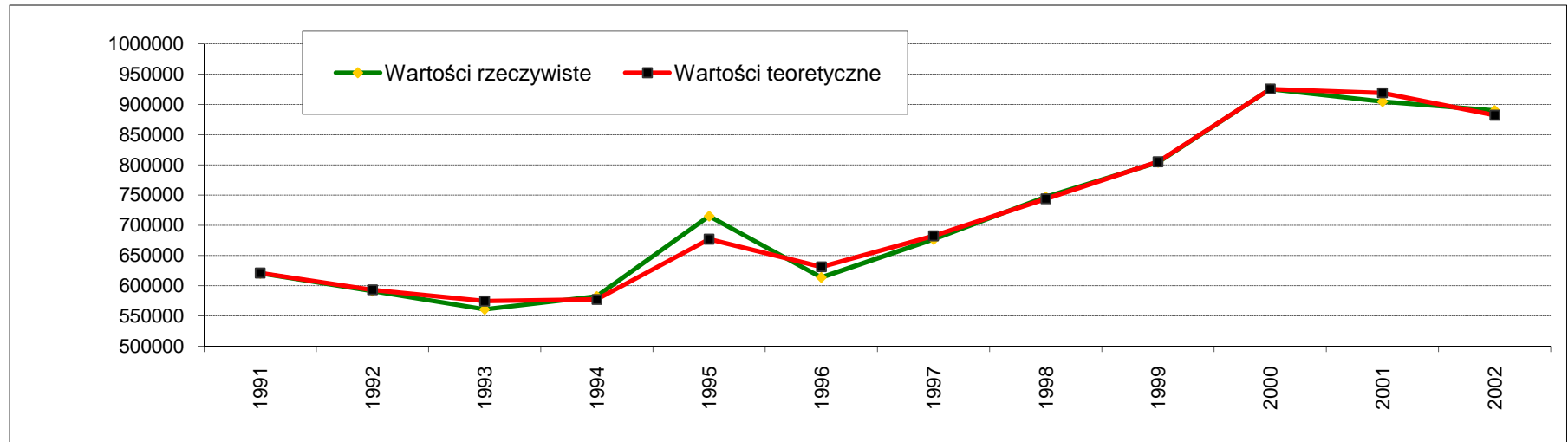
2. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej L15-19 na podstawie 12-elementowych prób, pomyliliśmy się przeciętnie rzecz biorąc o 50,741.
3. Jeżeli liczba biernych zawodowo ( zmienna L20-34 ) wzrośnie o 1000 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko mieniu społecznemu i prywatnemu ( zmienna P\_PM ) wzrośnie przeciętnie o 407,379 przypadków .
4. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej L20-34 na podstawie 12-elementowych prób, pomyliliśmy się przeciętnie rzecz biorąc o 23,206.
5. Znowelizowana klasyfikacja przestępstw w kodeksie karnym przy niezmienności pozostałych czynników równania spowodowała jednorazowy wzrost przestępstw przeciwko mieniu społecznemu i prywatnemu ( zmienna P\_PM ) przeciętnie o 58517,192 przypadków .
6. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej V2 na podstawie 12-elementowych prób, pomyliliśmy się przeciętnie rzecz biorąc o 195898,940.

<i>Miary dopasowania</i>	
R	0,994
R <sup>2</sup>	0,988
dopas. R <sup>2</sup>	0,984
Su	16966,290
Vu	2,36%

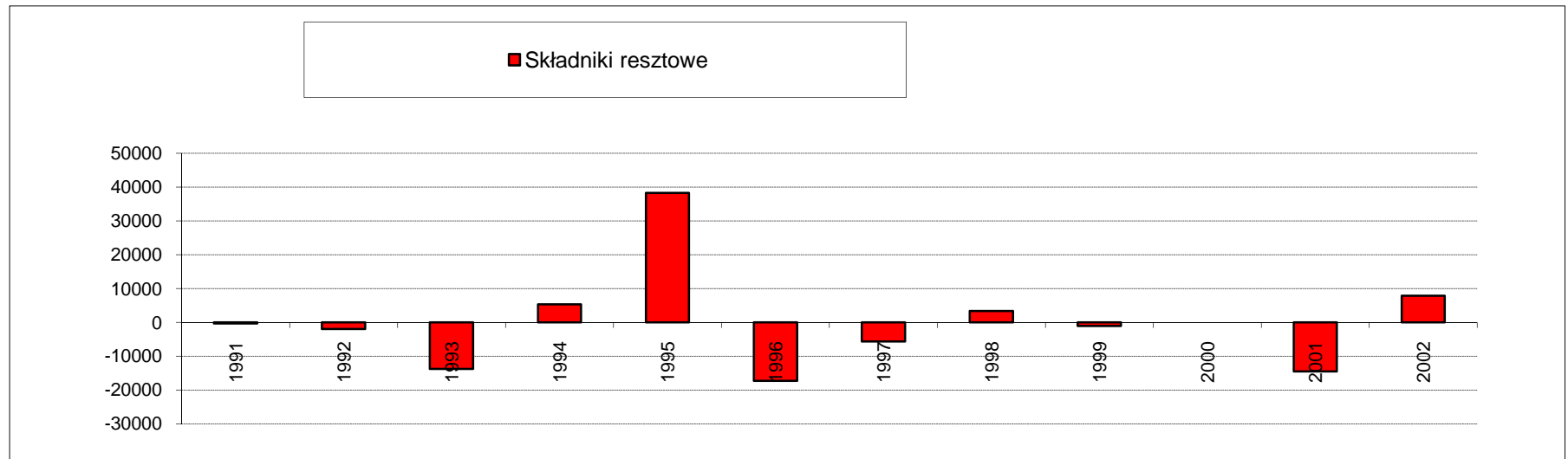
#### **Wnioski:**

1. Prawie 99 % zróżnicowania liczby przestępstw przeciwko mieniu społecznemu i prywatnemu ( zmienna P\_PM ) jest wyjaśniane przez równania.
2. Empiryczna liczba przestępstw przeciwko czci i netykalności cielesnej ( zmienna P\_PM ) odchylna się przeciętnie o 16966,290 przypadków od wartości teoretycznej wyznaczonej na podstawie równania.
3. Odchylenia te stanowią 2,36% przeciętnej liczby przestępstw przeciwko mieniu społecznemu i prywatnemu.

Rys. nr 3 Empiryczne wartości zmiennej P\_PM a wartości teoretyczne



Rys. nr 4 Składniki resztowe równania modelującego zmienną P\_PM





### 3.5. Weryfikacja poprawionego równania

#### 3.5.1. Badanie istotności parametrów strukturalnych równania

	Współczynniki	Błąd stand.	t Stat
Przecięcie	-3663682,317	195898,940	-18,702
L15-19	317,652	50,741	6,260
L20-34	407,379	23,206	17,555
V2	58517,199	19297,300	3,032

obserw.	12
zm. objaśnij.	3
poziom ist.	0,05
wartość t*	2,306

#### Wnioski:

1. Dla zmiennych ( L15-20 ); ( L20-34 ); ( V2 )  $|t| > t^*$  zatem hipotezę o statycznej nieistotności parametrów strukturalnych należy odrzucić. Oznacza to, że z prawdopodobieństwem 95 % można twierdzić, że zmienne uwzględnione w modelu wywierają istotny wpływ na zmienną P\_PM

#### 3.5.2. Badanie autokorelacji składnika losowego

Obserwacja	$\hat{P}_M$	reszty ( $u_t$ )	$u_t - u_{t-1}$
1991	621028,617	-323,617	-
1992	593196,671	-1946,671	-1623,055
1993	574781,475	-13780,475	-11833,803
1994	577261,529	5277,471	19057,946
1995	677098,493	38256,507	32979,036
1996	631243,073	-17343,073	-55599,579
1997	682488,885	-5652,885	11690,188
1998	743699,706	3315,294	8968,179
1999	805100,813	-1123,813	-4439,107
2000	925433,000	0,000	1123,813
2001	919150,267	-14527,267	-14527,267
2002	882236,471	7848,529	22375,796
suma kwadratów		2302840043	5634583609

d	2,44679765
d'	1,55320235
obserwacji	12
poziom ist.	0,05
zm.objaśnij.	3
$d_L$	0,812
$d_U$	1,579
$r_1$	-0,2402329
t	-0,7424409
t*	2,26215716

**Uwaga:**

1. W tablicach nie ma wartości krytycznych testu Durбина - Watsona dla 3 zmiennych objaśniających. W tabeli podano wartości krytyczne dla dwóch zmiennych i zaznaczono je kolorem . Wartość statystyki  $d'$  zawiera się w przedziale  $[d_L - d_U]$ . Wraz ze wzrostem ilości zmiennych objaśniających przedział ten się " rozszerza " w związku z tym również wartość statystyki  $d'$  zawiera się w przedziale  $[dL - dU ]$  dla trzech zmiennych. Wyniki testu t - Studenta pozwalają przyjąć hipotezę o braku autokorelacji.

### 3.5.3. Badanie jednorodności wariancji składnika losowego test Goldfeldta - Quandta

Celem przeprowadzenia testu Goldfeldta - Quandta próbę podzielono na 2 połowy i oszacowano dla każdej z nich z osobna parametry strukturalne i stochastyczne modelu:

<i>Miary dopasowania</i>	
R	0,960
R <sup>2</sup>	0,922
dopas. R <sup>2</sup>	0,870
Su	19488,054

<i>Miary dopasowania</i>	
R	0,998
R <sup>2</sup>	0,995
dopas. R <sup>2</sup>	0,988
Su	10947,388

F	3,168954072
n (1)	6
n (2)	6
zm. objaśn. (1) <sup>4)</sup>	2
zm. objaśn. (2) <sup>4)</sup>	3
poziom ist.	0,05
F* <sup>3)</sup>	19,164

<sup>3)</sup> Wartość krytyczna rozkładu F - Snedecora dla poziomu istotności 0,05 dla 2 i 3 stopni swobody.

<sup>4)</sup> Różna liczba zmiennych objaśniających wynika z tego, że w pierwszych 6-ciu latach ( 1991 -1996 ) nie uwzględniono zmiennej V1, która przyjmuje wartość 0

**Wniosek:**

1. Wartość statystyki F nie przekracza wartości krytycznej  $F^*$  w związku z tym nie ma podstaw do odrzucenia hipotezy o jednorodności wariancji składnika losowego.

**3.5.4. Badanie symetrii reszt**

m/n	0,333333333
t	-1,17260394
n	12
poziom ist.	0,05
$t^*$	2,200985159

**Wnioski:**

1.  $|t| < t^*$  co pozwala przyjąć hipotezę o symetrii reszt.

**3.5.5. Badanie losowości reszt**

K	6
liczba +	4
liczba -	7
poziom ist.	0,05
$K^{* 4)}$	3

- 4) Wartość krytyczna rozkładu serii dla lewostronnego obszaru krytycznego i poziomu istotności 0,05.

**Wnioski:**

1.  $K > K^*$  co oznacza, że nie ma podstaw do odrzucenia hipotezy o losowości składnika resztowego ut.

**Wniosek końcowy:**

Wyniki testów Durбина - Watsona i Goldfeldta - Quandta świadczą o tym, że klasyczna MNK jest właściwą metodą estymacji tego równania. Oznacza to, że parametry równania ze zmienną  $V_2$  można szacować klasyczną MNK. Ponadto składnik resztowy wykazuje losowość i symetrię.

Wszystkie zmienne objaśniające istotnie wpływają na  $P\_CNC$ , a parametry stochastyczne świadczą o bardzo dobrym dopasowaniu równania do empirycznych danych. Równanie uzupełnione o zmienną  $V_2$  można zatem uznać za postać ostateczną.

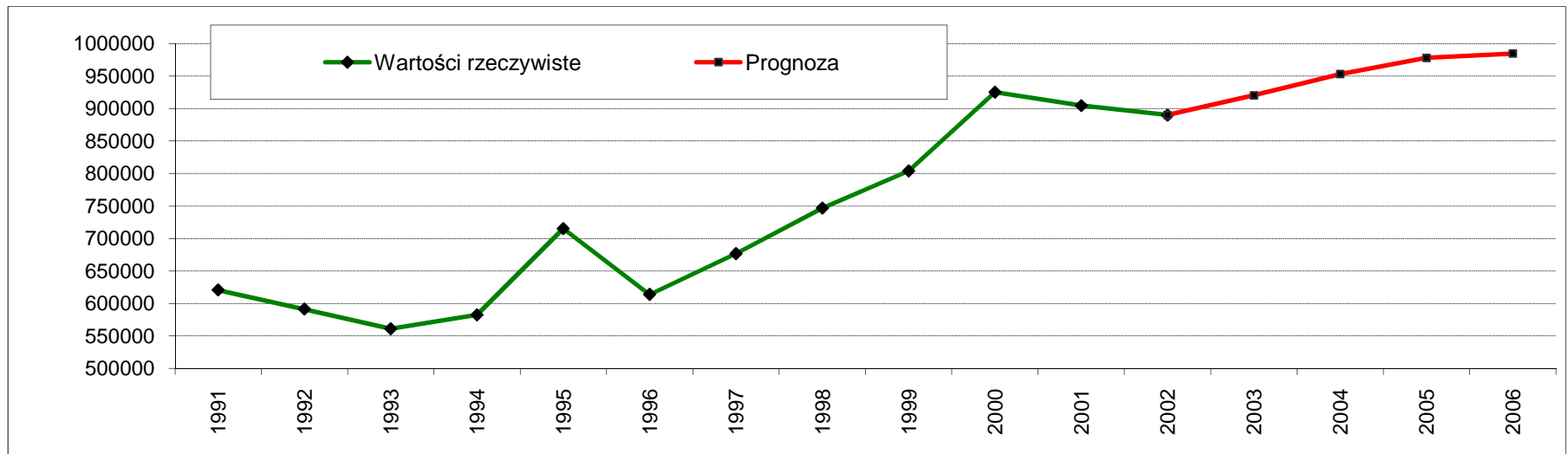
**3.6. Obliczenie prognozy dla przestępcstw przeciwko mieniu społecznemu i prywatnemu ( zmienna P\_M ) na lata 2003 - 2006**

<i>Współczynniki</i>	
Przecięcie	-3663682,317
L15-19	317,652
L20-34	407,379
V2	58517,199

<i>Miary dopasowania</i>	
R	0,994
R <sup>2</sup>	0,988
dopas. R <sup>2</sup>	0,984
Su	16966,290
Vu	2,36%

rok	$\wedge$ L15-19 ( prognoza )	$\wedge$ L20-34 ( prognoza )	V2	P_M	$\wedge$ P_M ( prognoza )
1991				620705	
1992				591250	
1993				561001	
1994				582539	
1995				715355	
1996				613900	
1997				676836	
1998				747015	
1999				803977	
2000				925433	
2001				904623	
2002				890085	
2003	3104,8425	8831,98	0,0		920542,802
2004	2978,585	9010,56	0,0		953186,562
2005	2852,3275	9169,94	0,0		978008,644
2006	2726,07	9284,72	0,0		984661,616

Rys. nr 5 Wartości empiryczne i prognozowane zmiennej P\_M



### 3.6.1. Obliczenie średniego błędu predykcji

X =

L15-19	L20-34	V2	
2998,6	8179,6	0,0	1,0
3067,7	8057,4	0,0	1,0
3139	7956,6	0,0	1,0
3194,9	7919,1	0,0	1,0
3236,8	8131,5	0,0	1,0
3244,8	8012,7	0,0	1,0
3292,5	8101,3	0,0	1,0
3343,1	8212,1	0,0	1,0
3365,7	8345,2	0,0	1,0
3356,9	8503,8	1,0	1,0
3322,3	8659	0,0	1,0
3231,1	8639,5	0,0	1,0

$$X^T =$$

<b>L15-19</b>	2998,6	3067,7	3139	3194,9	3236,8	3244,8	3292,5	3343,1	3365,7
<b>L20-34</b>	8179,6	8057,4	7956,6	7919,1	8131,5	8012,7	8101,3	8212,1	8345,2
<b>V2</b>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0

3356,9	3322,3	3231,1
8503,8	8659	8639,5
1,0	0,0	0,0
1,0	1,0	1,0

$$X^T X =$$

125559966	319285314,5	3356,9	38793,4
319285314,5	812817699,7	8503,8	98717,8
3356,9	8503,8	1,0	1,0
38793,4	98717,8	1,0	12,0

$$(X^T X)^{-1} =$$

0,000009	-0,000002	-0,000722	-0,015554
-0,000002	0,000002	-0,000347	-0,010135
-0,000722	-0,000347	1,293658	5,081234
-0,015554	-0,010135	5,081234	133,318491

$$X_T =$$

	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>
<b>^L15-19</b>	3104,8425	2978,585	2852,3275	2726,07
<b>^L20-34</b>	8831,98	9010,56	9169,94	9284,72
<b>V2</b>	0,0	0,0	0,0	0,0
	1,0	1,0	1,0	1,0

$$X_T^T =$$

	<b>^L15-19 (proгноza)</b>	<b>^L20-34 (proгноza)</b>	<b>V2</b>	
<b>2003</b>	3104,8425	8831,98	0,0	1,0
<b>2004</b>	2978,585	9010,56	0,0	1,0
<b>2005</b>	2852,3275	9169,94	0,0	1,0
<b>2006</b>	2726,07	9284,72	0,0	1,0

$$X_T^T (X^T X)^{-1} =$$

<b>2003</b>	-0,0020631	0,00136858	-0,2255957	-4,4868508
<b>2004</b>	-0,0034811	0,0019068	-0,1964289	-4,3329581
<b>2005</b>	-0,004868	0,00240911	-0,1605981	-3,9844724
<b>2006</b>	-0,0061829	0,00282798	-0,1092877	-3,1839634

$$X_T^T (X^T X)^{-1} X^T =$$

	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>
<b>2003</b>	1,19475675	1,69963986	2,17824631	2,5958143
<b>2004</b>	1,69963986	2,47967124	3,223092	3,8814693
<b>2005</b>	2,17824631	3,223092	4,22168276	5,1128272
<b>2006</b>	2,59581429	3,88146932	5,11282717	6,2180572

### 3.6.2. Prognoza dla zmiennej P\_M na lata 2003 - 2006 i średnie błędy predykcji

rok	prognoza	śr. błąd	śr. wzgl. błąd
<b>2003</b>	920542,802	25135,069	2,73%
<b>2004</b>	953186,562	31648,709	3,32%
<b>2005</b>	978008,644	38769,673	3,96%
<b>2006</b>	984661,616	45582,386	4,63%



## **4. PRZESTĘPSTWA PRZECIWKO WOLNOŚCI I OBYCZAJOWOŚCI**

**4.1. Dobór zmiennych objaśniających**

**4.2. Estymacja parametrów równania**

**4.3. Weryfikacja równania**

**3.4. Obliczenie prognozy dla przestępstw przeciwko wolności i obyczajowości**

**4.1. Dobór zmiennych objaśniających ( egzogenicznych ) dla równania modelu, w których zmienną objaśnianą stanowią przestępstwa przeciwko wolności i obyczajowości ( P\_WO )**

**4.1.1. Baza danych niezbędna do wytypowania zmiennych równania**

P_WO	L	L15-19	L20-34	L35-50	LB	LBZ	SB	WBZ	B_W	B_SO
13069	38309	2998,6	8179,6	8090	2155,6	9001	12,2	37,81	64,1	155,5
16317	38418	3067,7	8057,4	8387,4	2509,3	10862	14,3	38,3	56,6	177,9
19268	38505	3139	7956,6	8675,5	2889,6	11013	16,4	38,8	52,1	193,2
24753	38581	3194,9	7919,1	8913,7	2838	11781	16	39,1	47,6	194,4
28516	38609	3236,8	8131,5	9114,2	2628,8	12102	14,9	41,6	38,9	188,2
29463	38639	3244,8	8012,7	9152	2359,5	12422	13,2	42,4	31,4	151,7
35611	38660	3292,5	8101,3	9130,5	1826,4	12675	10,3	42,6	25,1	115,2
36032	38667	3343,1	8212,1	9060,8	1831,4	12899	10,4	42,9	30,2	111,2
36085	38654	3365,7	8345,2	8974,1	2349,8	13174	13,1	43,3	47,7	142
44494	38644	3356,9	8503,8	8835,9	2702,6	13371	15,1	43,6	69,4	168,7
47556	38632	3322,3	8659	8678,9	3115,1	13664	17,5	44,2	100,5	192,1
47816	38219	3231,1	8639,5	8371,4	3217	14012	18	44,81	126,7	199

B_SZ	B_Z	B_P	DRG	WRG
482,5	805,6	663,3	618,64	566,03
527,9	964,6	782,4	631,16	620,39
581,5	1131,2	931,5	573,8	559,58
570,2	1118,3	907,5	577,31	546,51
531,7	1025	845	587,93	540,51
471	907,7	797,6	625,55	572,64
364,2	700,7	621,3	672,73	607,59
369	698	622,8	664,14	638,86
483	898,2	778,9	663,34	650,71
561,9	998,8	903,8	656,33	644,48
666	1142,5	1014	656,73	621,3
683,6	1163,6	1044,1	664,21	624,99

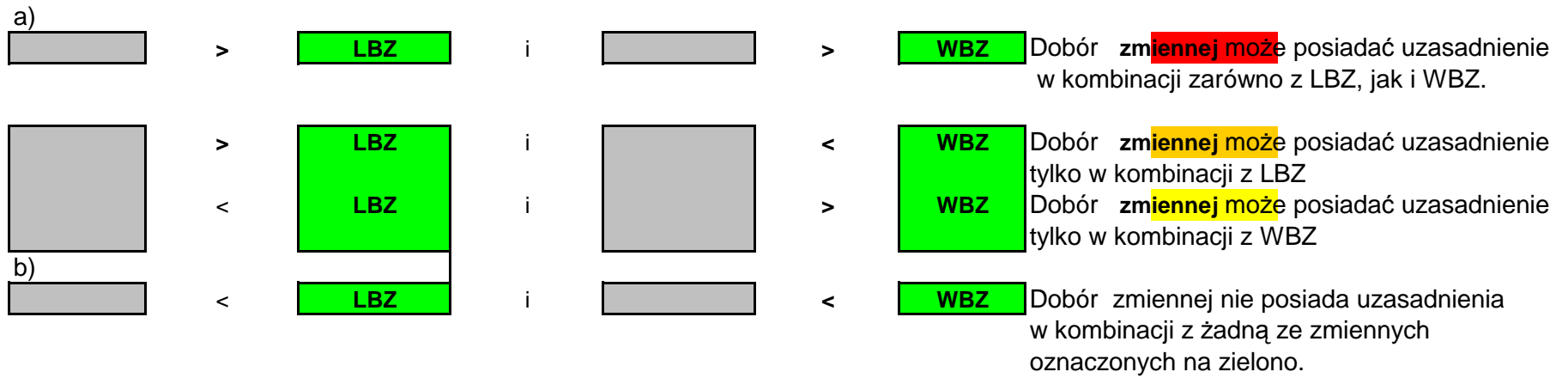
#### 4.1.2. Macierz korelacji zmiennych

	P_WO	L	L15-19	L20-34	L35-50	LB	LBZ	SB	WBZ	B_W
P_WO	1,000	0,257	0,832	0,808	0,296	0,285	0,948	0,257	0,963	0,465
L	0,257	1,000	0,699	-0,157	0,872	-0,350	0,369	-0,358	0,305	-0,622
L15-19	0,832	0,699	1,000	0,465	0,703	-0,045	0,885	-0,071	0,857	-0,051
L20-34	0,808	-0,157	0,465	1,000	-0,253	0,406	0,638	0,385	0,745	0,778
L35-50	0,296	0,872	0,703	-0,253	1,000	-0,328	0,476	-0,338	0,398	-0,644
LB	0,285	-0,350	-0,045	0,406	-0,328	1,000	0,272	0,999	0,141	0,760
LBZ	0,948	0,369	0,885	0,638	0,476	0,272	1,000	0,247	0,940	0,310
SB	0,257	-0,358	-0,071	0,385	-0,338	0,999	0,247	1,000	0,111	0,751
WBZ	0,963	0,305	0,857	0,745	0,398	0,141	0,940	0,111	1,000	0,333
B_W	0,465	-0,622	-0,051	0,778	-0,644	0,760	0,310	0,751	0,333	1,000
B_SO	-0,031	-0,437	-0,326	0,154	-0,409	0,935	-0,034	0,944	-0,170	0,618
B_SZ	0,280	-0,445	-0,107	0,493	-0,461	0,985	0,223	0,983	0,131	0,844
B_Z	0,174	-0,295	-0,089	0,257	-0,262	0,985	0,199	0,988	0,034	0,646
B_P	0,404	-0,251	0,095	0,456	-0,215	0,987	0,401	0,982	0,269	0,742
DRG	0,676	0,104	0,557	0,665	0,045	-0,288	0,591	-0,312	0,713	0,237
WRG	0,613	0,120	0,561	0,654	-0,026	-0,122	0,579	-0,141	0,606	0,281

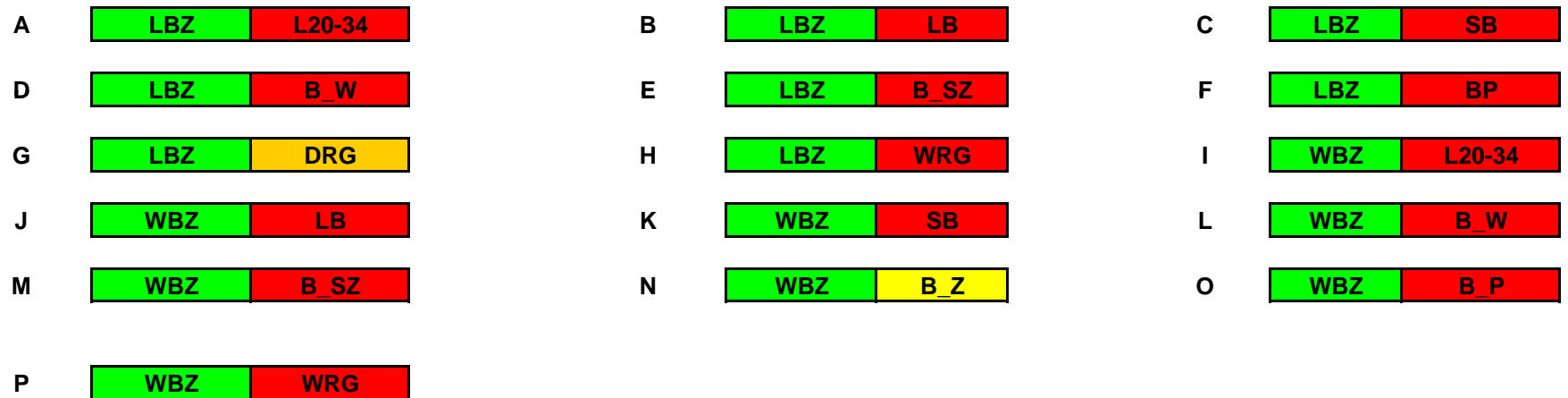
B_SO	B_SZ	B_Z	B_P	DRG	WRG
-0,031	0,280	0,174	0,404	0,676	0,613
-0,437	-0,445	-0,295	-0,251	0,104	0,120
-0,326	-0,107	-0,089	0,095	0,557	0,561
0,154	0,493	0,257	0,456	0,665	0,654
-0,409	-0,461	-0,262	-0,215	0,045	-0,026
0,935	<b>0,985</b>	<b>0,985</b>	<b>0,987</b>	-0,288	-0,122
-0,034	0,223	0,199	0,401	0,591	0,579
0,944	<b>0,983</b>	<b>0,988</b>	<b>0,982</b>	-0,312	-0,141
-0,170	0,131	0,034	0,269	0,713	0,606
0,618	<b>0,844</b>	0,646	0,742	0,237	0,281
1,000	<b>0,921</b>	0,957	0,877	-0,546	-0,383
<b>0,921</b>	1,000	<b>0,948</b>	<b>0,959</b>	-0,215	-0,067
0,957	<b>0,948</b>	1,000	<b>0,966</b>	-0,425	-0,235
0,877	<b>0,959</b>	<b>0,966</b>	1,000	-0,209	-0,052
-0,546	-0,215	-0,425	-0,209	1,000	<b>0,884</b>
-0,383	-0,067	-0,235	-0,052	<b>0,884</b>	1,000

### 4.1.3. Zasady doboru zmiennych do równania

1. Wektor współczynników korelacji zmiennej objaśnianej z potencjalnymi zmiennymi objaśniającymi zaznaczono kolorem [szary].
2. Wybór najsilniej skorelowanej potencjalnej zmiennej egzogenicznej ze zmienna endogeniczną (objaśnianą) został zaznaczony kolorem [niebieski]. W tym wypadku zaznaczono dwie najsilniej skorelowane zmienne, ponieważ ich współczynniki korelacji różniły się istotnie od pozostałych. Jako że przyjęto maksymalną wartość współczynnika korelacji między zmiennymi egzogenicznymi na poziomie 0,8 (wyższe wartości oznaczono kolorem) [niebieski]. Do modelu należy zakwalifikować jedną z nich.
3. W następnej kolejności rozważano zasadność doboru drugiej zmiennej objaśniającej. Rozważano tu równolegle kombinacje zarówno z LBZ jak i WBZ. W celu wstępnej selekcji pozostałych zmiennych zastosowano następującą procedurę:



Kierując się powyższymi regułami, zastosowano metodę Hellwiga dla następujących par zmiennych.



Wektor modułów współczynnika korelacji zmiennej objaśnianej ( endogenicznej ) z badanymi zmiennymi

	P_WO
L20-34	0,808329402
LB	0,2846435
LBZ	0,947918352
SB	0,257435232
WBZ	0,962709106
B_W	0,46475745
B_SZ	0,279957788
B_Z	0,174427002
B_P	0,403538524
DRG	0,676
WRG	0,612516551

Macierz modułów współczynników korelacji między badanymi zmiennymi

	L20-34	LB	LBZ	SB	WBZ	B_W	B_SZ	B_Z	B_P	DRG
L20-34	1,000	0,406	0,638	0,385	0,745	0,778	0,493	0,257	0,456	0,665
LB	0,406	1,000	0,272	<b>0,999</b>	0,141	0,760	<b>0,985</b>	<b>0,985</b>	<b>0,987</b>	0,288
LBZ	0,638	0,272	1,000	0,247	<b>0,940</b>	0,310	0,223	0,199	0,401	0,591
SB	0,385	<b>0,999</b>	0,247	1,000	0,111	0,751	<b>0,983</b>	<b>0,988</b>	<b>0,982</b>	0,312
WBZ	0,745	0,141	<b>0,940</b>	0,111	1,000	0,333	0,131	0,034	0,269	0,713
B_W	0,778	0,760	0,310	0,751	0,333	1,000	<b>0,844</b>	0,646	0,742	0,237
B_SZ	0,493	<b>0,985</b>	0,223	<b>0,983</b>	0,131	<b>0,844</b>	1,000	<b>0,948</b>	<b>0,959</b>	0,215
B_Z	0,257	<b>0,985</b>	0,199	<b>0,988</b>	0,034	0,646	<b>0,948</b>	1,000	<b>0,966</b>	0,425
B_P	0,456	<b>0,987</b>	0,401	<b>0,982</b>	0,269	0,742	<b>0,959</b>	<b>0,966</b>	1,000	0,209
DRG	0,665	0,288	0,591	0,312	0,713	0,237	0,215	0,425	0,052	1,000
WRG	0,654	0,122	0,579	0,141	0,606	0,281	0,067	0,235	0,052	<b>0,884</b>

<b>WRG</b>
0,654
0,122
0,579
0,141
0,606
0,281
0,067
0,235
0,052
<b>0,884</b>
1,000

**A**

h(A, LBZ)	0,548622171
h(A, L20-34)	0,398940607
H(A)	<b>0,947562778</b>

**B**

h(B, LBZ)	0,706337576
h(B, LB)	0,063690255
H(B)	0,770027832

**C**

h(C, LBZ)	0,72055228
h(C, SB)	0,05314466
H(C)	0,77369693

**D**

h(D, LBZ)	0,685752437
h(D, BW)	0,164845926
H(D)	0,850598363

**E**

h(E, LBZ)	0,734950258
h(E, B_SZ)	0,064106371
H(E)	0,799056629

**F**

h(F, LBZ)	0,641351647
h(F, BP)	0,116231637
H(F)	0,757583284

**G**

h(G, LBZ)	0,5649123
h(G, DRG)	0,2875024
H(G)	0,8524147

**H**

h(H, LBZ)	0,569053514
h(H, WRG)	0,237600256
H(H)	0,80665377

**I**

h(I, WBZ)	0,531067265
h(I, L20_34)	0,374400246
H(I)	0,90546751

**J**

h(J, WBZ)	0,812558437
h(J, LB)	0,071034117
H(J)	0,883592555

**K**

h(K, WBZ)	0,83385548
h(K, SB)	0,05962613
H(K)	0,89348161

**L**

h(L, WBZ)	0,695215192
h(L, BW)	0,16202492
H(L)	0,857240112

**M**

h(M, WBZ)	0,819522574
h(M, BSZ)	0,069303612
H(M)	0,888826186

**N**

h(N, WBZ)	0,896170239
h(N, B_Z)	0,029418992
H(N)	0,925589231

**O**

h(O, WBZ)	0,73028516
h(O, BP)	0,12831349
H(O)	0,85859865

**P**

h(P, WBZ)	0,576998199
h(P, WRG)	0,233571557
H(P)	0,810569755

**Wniosek:**

Z powyższego wynika, że spośród wszystkich rozważanych par najlepiej zmienną P\_WO będzie opisywać kombinacja A ( L20-34, LBZ ).

4. Nie ma jednak pewności, czy nie jest możliwa poprawa jakości równania poprzez włączenie trzeciej zmiennej. Jako potencjalną trzecią zmienną rozważono każdą z cech oznaczonych wcześniej na czerwono.


 > 


 i 


 > 


 lub 


 < 


 dobór zmiennej może mieć uzasadnienie  
dobór zmiennej nie ma uzasadnienia,  
ponieważ w większym stopniu "powiela"  
ona jedną ze zmiennych egzogenicznych  
niż opisuje zmienną P\_WO.

	P_WO	L20-34	LBZ
L20-34	0,808	1,000	0,638
LB	0,285	0,406	0,272
LBZ	0,948	0,638	1,000
SB	0,257	0,385	0,247
WBZ	0,963	0,745	0,940
B_W	0,465	0,778	0,310
B_SZ	0,280	0,493	0,223
B_P	0,404	0,456	0,401
DRG	0,676	0,665	0,591
WRG	0,613	0,654	0,579

**R**

h(R, LBZ)	0,403221055
h(R, L20-34)	0,283677224
h(R, DRG)	0,202697219
H(R)	0,889595498



**Wniosek końcowy:**

**Ze względu, że kombinacje od B do R ( .. )  $<H(A)$  nie ma uzasadnienia dodania trzeciej zmiennej objaśniającej DRG, gdyż ta zmienna nie poprawi jakości modelu.**

## 4.2. Estymacja parametrów równania opisującego zmienną P\_WO

### 4.2.1. Baza danych niezbędna do oszacowania równania

Rok	P_WO	L20-34	LBZ
1991	13069	8179,6	9001
1992	16317	8057,4	10862
1993	19268	7956,6	11013
1994	24753	7919,1	11781
1995	28516	8131,5	12102
1996	29463	8012,7	12422
1997	35611	8101,3	12675
1998	36032	8212,1	12899
1999	36085	8345,2	13174
2000	44494	8503,8	13371
2001	47556	8659	13664
2002	47816	8639,5	14012

### 4.2.2. Klasyczną MNK za pomocą MC Excel oszacowano parametry strukturalne oraz parametry stochastyczne modelu

$$P\_WO = 15,859[L20-34_t] + 6,064[LBZ_t] - 173152,776 + u_t$$

(3,546)                      (0,639)                      (24930,574)

### Interpretacja :

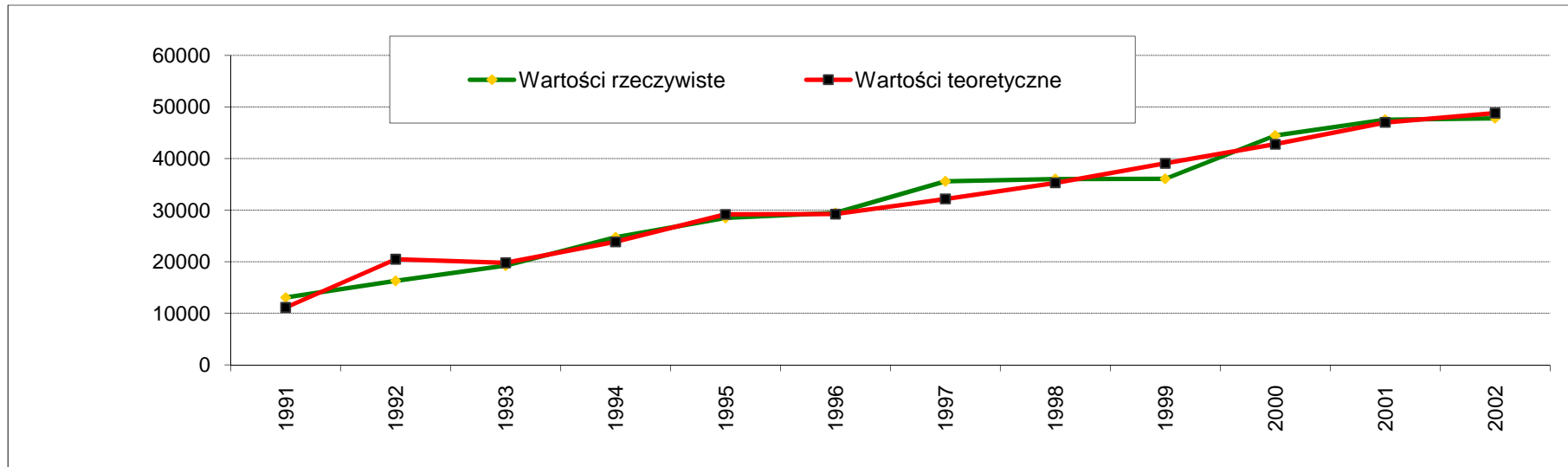
1. Jeżeli liczba ludności w wieku od 20 do 34 lat ( zmienna L20-34 ) wzrośnie o 1000 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko wolności i obyczajowości ( zmienna P\_WO ) wzrośnie przeciętnie o 15,859 przypadków.
2. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej L20-34 na podstawie 12-elementowych prób, pomylibyśmy się przeciętnie rzecz biorąc o 3,546.
3. Jeżeli liczba biernych zawodowo ( zmienna LBZ ) wzrośnie o 1000 osób, a pozostałe czynniki modelu nie ulegną zmianie to ilość przestępstw przeciwko wolności i obyczajowości ( zmienna P\_WO ) wzrośnie przeciętnie o 6,064 przypadków .
4. Gdybyśmy wielokrotnie szacowali parametr przy zmiennej LBZ na podstawie 12-elementowych prób, pomylibyśmy się przeciętnie rzecz biorąc o 0,639.

<i>Miary dopasowania</i>	
R	0,984
R <sup>2</sup>	0,969
dopas. R <sup>2</sup>	0,962
Su	2313,287
Vu	7,32%

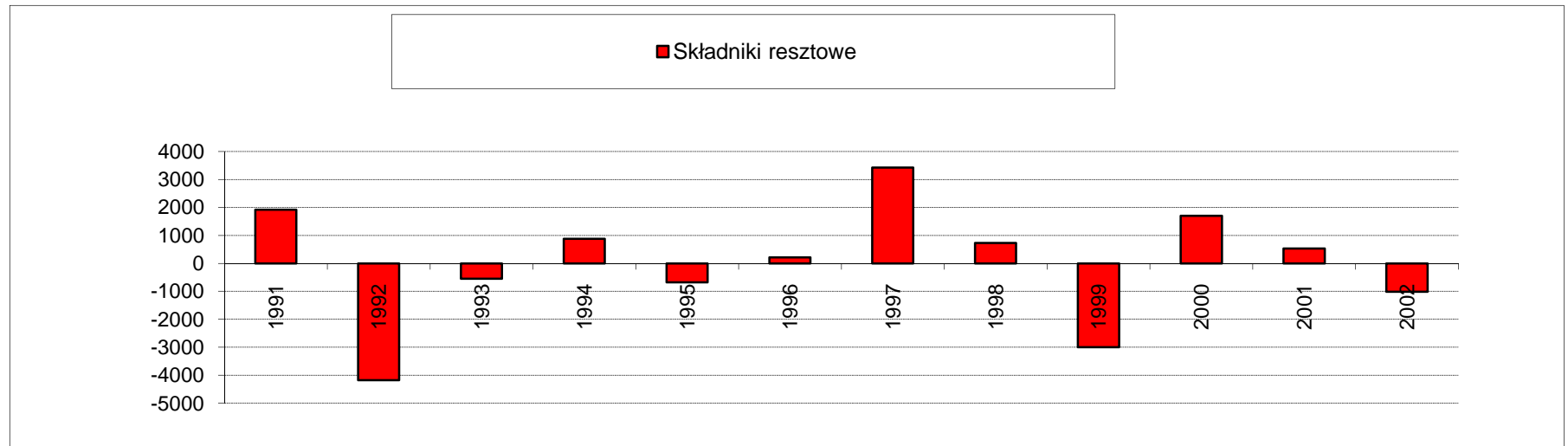
### Wnioski:

1. Prawie 97 % zróżnicowania liczby przestępstw przeciwko wolności i obyczajowości ( zmienna P\_WO ) jest wyjaśniane przez model.
2. Empiryczna liczba przestępstw przeciwko wolności i obyczajowości ( zmienna P\_WO ) odchyła się przeciętnie o 2313,287 przypadków od wartości teoretycznej wyznaczonej na podstawie modelu.
3. Odchylenia te stanowią 7,32% przeciętnej liczby przestępstw przeciwko wolności i obyczajowości.

Rys. nr 1 Empiryczne wartości zmiennej P\_WO a wartości teoretyczne



Rys. nr 2 Składniki resztowe równania modelującego zmienną P\_WO



### 4.3. Weryfikacja modelu

#### 4.3.1. Badanie istotności parametrów strukturalnych równania

	<i>Współczynniki</i>	<i>Błąd stand.</i>	<i>t</i>
Przecięcie	-173151,775	24930,574	-6,945
L20-34	15,859	3,546	4,472
LBZ	6,064	0,639	9,491

obserw.	12
zm. objaśnij.	2
poziom ist.	0,05
wartość $t^*$ <sup>1)</sup>	2,262

1) Wartość krytyczna rozkładu t - Studenta dla poziomu istotności 0,05 i dla 9 stopni swobody.

#### Wnioski:

1. Dla zmiennych ( L20-34 ); ( LBZ )  $|t| > t^*$  hipotezę o statycznej nieistotności parametrów strukturalnych należy odrzucić. Oznacza to, że z prawdopodobieństwem 95 % można twierdzić, że obie zmienne uwzględnione w modelu wywierają istotny wpływ na zmienną P\_WO.

### 4.3.2. Badanie autokorelacji składnika losowego

W celu zbadania autokorelacji rzędu pierwszego posłużono się dwoma metodami:

1. Test Durбина - Watsona.
2. Wyznaczono wartość współczynnika autokorelacji rzędu pierwszego i przeprowadzono test t - Studenta na istotność współczynnika korelacji.

Obserwacja	$\hat{P}_{WO}$	reszty ( $u_t$ )	$u_t - u_{t-1}$
1991	11149,097	1919,903	-
1992	20495,783	-4178,783	-6098,686
1993	19812,816	-544,816	3633,967
1994	23875,072	877,928	1422,744
1995	29190,009	-674,009	-1551,937
1996	29246,355	216,645	890,654
1997	32185,602	3425,398	3208,753
1998	35301,071	730,929	-2694,469
1999	39079,450	-2994,450	-3725,379
2000	42789,261	1704,739	4699,188
2001	47027,274	528,726	-1176,012
2002	48828,212	-1012,212	-1540,938
suma kwadratów		48161659,02	112900237,9

d	2,34419329
d'	1,65580671
obserwacji	12
poziom ist.	0,05
zm.objaśnaj.	2
$d_L$	0,812
$d_U$	1,579
$r_1$	-0,229697
t	-0,7080221
$t^*$	2,26215716

#### Wnioski:

1. Zarówno wartości statystyki  $d' > d_U$  ( Durбина - Watsona ) jak i statystyki t - Studenta  $|t| < t$  świadczą o tym, że nie ma podstaw do odrzucenia hipotezy o braku atokorelacji składnika losowego.

### 4.3.3. Badanie jednorodności wariancji składnika losowego - test Goldfeldta - Quandta

Celem przeprowadzenia testu Goldfeldta - Quandta próbę podzielono na 2 połowy i oszacowano dla każdej z nich z osobna parametry strukturalne i stochastyczne modelu:

Statystyki regresji	
R	0,952
R <sup>2</sup>	0,907
dopas. R <sup>2</sup>	0,845
Su	2638,863

Statystyki regresji	
R	0,947
R <sup>2</sup>	0,897
dopas. R <sup>2</sup>	0,829
Su	2476,012

Wyznaczenie wartości statystyki F - Snedecora

F	1,136
obserw. (1)	6
obserw. (2)	6
zm.objaśnaj.	2
poziom ist.	0,05
F* <sup>2)</sup>	9,277

<sup>2)</sup> Wartość krytyczna rozkładu F - Snedecora dla poziomu istotności 0,05 oraz dla trzech i trzech stopni swobody.

#### Wnioski:

1. Wartość statystyki F nie przekracza wartości krytycznej F\* w związku z tym nie ma podstaw do odrzucenia hipotezy o jednorodności wariancji składnika losowego.
2. Wyniki testów DW i G Q świadczą o tym, że klasyczna MNK jest właściwą metodą estymacji tego równania. Uzyskane oceny parametrów strukturalnych są bowiem estymatorami nieobciążonymi i najefektywniejszymi.

#### 4.3.4. Badanie symetrii reszt

m/n	0,583
t	0,561
n	12
poziom ist.	0,05
t*	2,201

**Wnioski:**

1.  $|t| < t^*$  co pozwala przyjąć hipotezę o symetrii reszt.

#### 4.3.5. Badanie losowości reszt

K	8
poziom ist.	0,05
liczba +	4
liczba -	4
$K^*$ <sup>3)</sup>	2

- <sup>3)</sup> Wartość krytyczna serii dla lewostronnego obszaru krytycznego i poziomu istotności 0,05

**Wnioski:**

1.  $K > K^*$  co oznacza, że nie ma podstaw do odrzucenia hipotezy o losowości składnika resztowego ut.



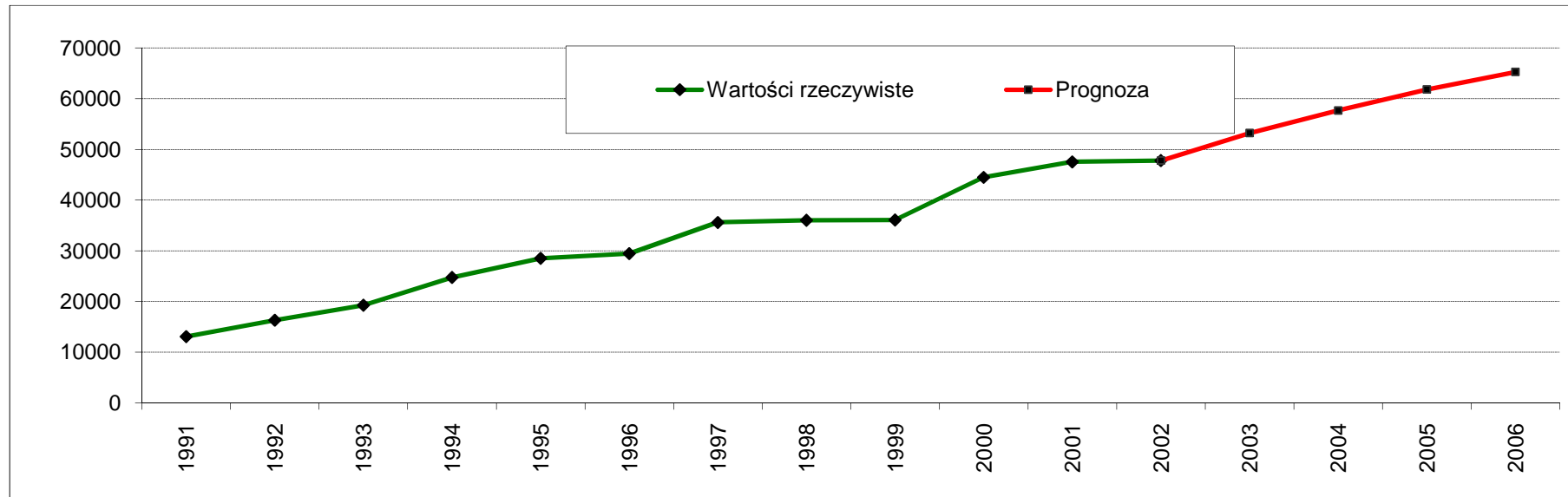
**4.4. Prognoza liczby przestępstw przeciwko wolności i obyczajowości  
( zmienna P\_WO ) na lata 2003 - 2006**

<i>Współczynniki</i>	
Przecięcie	-173151,775
L20-34	15,859
LBZ	6,064

<i>Miary dopasowania</i>	
R	0,984
R <sup>2</sup>	0,969
dopas. R <sup>2</sup>	0,962
Su	2313,287
Vu	7,32%

rok	$\wedge$ L20-34 (prognoza)	$\wedge$ LBZ (prognoza)	P_WO	$\wedge$ P_WO (prognoza)
1991			13069	
1992			16317	
1993			19268	
1994			24753	
1995			28516	
1996			29463	
1997			35611	
1998			36032	
1999			36085	
2000			44494	
2001			47556	
2002			47816	
2003	8831,98	14233,91667		53226,41762
2004	9010,56	14500,7		57676,24337
2005	9169,94	14767,48333		61821,5748
2006	9284,72	15034,26667		65259,59128

Rys. nr 3 Wartości empiryczne i prognozowane zmiennej P\_WO



#### 4.4.1. Obliczenie średniego błędu predykcji

X =

L20-34	LBZ	
8179,6	9001,0	1,0
8057,4	10862,0	1,0
7956,6	11013,0	1,0
7919,1	11781,0	1,0
8131,5	12102,0	1,0
8012,7	12422,0	1,0
8101,3	12675,0	1,0
8212,1	12899,0	1,0
8345,2	13174,0	1,0
8503,8	13371,0	1,0
8659	13664,0	1,0
8639,5	14012,0	1,0

$$X^T =$$

<b>L20-34</b>	8179,6	8057,4	7956,6	7919,1	8131,5	8012,7	8101,3	8212,1	8345,2
<b>LBZ</b>	9001,0	10862,0	11013,0	11781,0	12102,0	12422,0	12675,0	12899,0	13174,0
	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0

8503,8	8659,0	8639,5
13371,0	13664,0	14012,0
1,0	1,0	1,0

$$X^T X =$$

812817699,7	1211635264	98717,8
1211635264	1822262446	146976
98717,8	146976	12

$$(X^T X)^{-1} =$$

0,0000024	-0,0000003	-0,0160251
-0,0000003	0,0000001	0,0012873
-0,016025075	0,00128731	116,1463706

$$X_T =$$

	2003	2004	2005	2006
$\wedge$ L20-34	8831,98	9010,56	9169,94	9284,72
$\wedge$ LBZ	14233,91667	14500,7	14767,48333	15034,26667
	1,0	1,0	1,0	1,0

$$X_T^T =$$

	$\wedge$ L20-34	$\wedge$ LBZ	
<b>2003</b>	8831,98	14233,91667	1,0
<b>2004</b>	9010,56	14500,7	1,0
<b>2005</b>	9169,94	14767,48333	1,0
<b>2006</b>	9284,72	15034,26667	1,0

$$(X_T)^T (X_T X)^{-1} =$$

	$\wedge$ L20-34	$\wedge$ LBZ	
<b>2003</b>	8831,98	14233,91667	1,0
<b>2004</b>	9010,56	14500,70000	1,0
<b>2005</b>	9169,94	14767,48333	1,0
2006	9284,72000	15034,26667	1,0

$$X_T^T (X^T X)^{-1} X^T =$$

	2003	2004	2005	2006
2003	0,596305007	0,751432922	0,889537247	0,98809719
2004	0,751432922	0,961203269	1,147275586	1,278299352
2005	0,889537247	1,147275586	1,375507779	1,535199656
2006	0,98809719	1,278299352	1,535199656	1,714742515

#### 4.4.2. Prognozy dla zmiennej P\_WO na lata 2003 -2006 i średnie błędy predykcji

rok	prognoza	śr. błąd	śr. wzgl. błąd
2003	53226,418	2922,721	5,49%
2004	57676,243	3239,595	5,62%
2005	61821,575	3565,395	5,77%
2006	65259,591	3811,481	5,84%

## **5. PODSUMOWANIE**

### **5.1. Ostateczna postać modelu ekonometrycznego**

## 5.1. Ostateczna postać modelu ekonometrycznego

$$P\_CNC = 1,898[L20-34_t] + 0,256[LBZ_t] + 1001,582V1_t - 15310,304 + u_t$$

(0,614)                      (0,065)                      (337,216)                      (4830,493)

$$P\_PM = 317,652[L15-19_t] + 407,379[L20-34_t] + 58517,199V2_t - 3663682,317 + u_t$$

(50,741)                      (23,206)                      (19297,300)                      (195898,940)

$$P\_WO = 15,859[L20-34_t] + 6,064[LBZ_t] - 173152,776 + u_t$$

(3,546)                      (0,639)                      (24930,574)

### Wnioski:

1. Występowanie zmiennej L20-34 we wszystkich równaniach modelu jest potwierdzeniem, wyników przeprowadzonych badań naukowych dotyczących międzynarodowego terroryzmu i zorganizowanej działalności przestępczej, że najbardziej podatną grupą na czynniki kryminogenne oraz największy udział w działalności przestępczej mają ludzie od 20 do 34 roku życia.
2. Udział drugiej grupy wiekowej tj. od 15 do 19 roku życia ( zmienna L15 - 19 ) w drugim równaniu modelu potwierdza ogólne przekonanie panujące w społeczeństwie, że duża część włamań, kradzieży i przywłaszczeń realizowana jest przez młodocianych.
3. Struktura demograficznych zmiennych ( duży udział zmiennych opisujących wiek, mały dotyczących statusu finansowego ) wchodzących w skład równań, upoważnia do twierdzenia że duża część z opisanych przestępstw dokonuje się w ramach własnej rodziny.
4. Wyniki przeprowadzonego badania w ramach niniejszej pracy z dużym uproszczeniem upowazniają mnie do twierdzenia, że wbrew panującym stereotypom to nie "bieda" popycha ludzi do dokonywania przestępstw. Niski stopień zamożności może wpływać na obniżenie poziomu osobistego kodeksu etyczno-moralnego jednostki.  
Żadna ze zmiennych opisujących w części zamożność społeczeństwa:  
a) DRW                      dochód realny w rodzinie na osobę,  
b) WRG                      wydatki w rodzinie na osobę  
nie weszła w skład równania oraz znikomy udział w równaniach zmiennych opisujących bezrobocie:

- a) LB                    bezrobotni ogółem,
  - b) LBZ                 liczba biernych zawodowo,
  - c) SB                   stopa bezrobocia,
  - d) WBZ                wskaźnik bierności zawodowej,
  - e) B\_W                 bezrobotni z wykształceniem wyższym
  - f) B\_SO                bezrobotni z wykształceniem średnim ogólnym
  - g) B\_SZ                bezrobotni z wykształceniem średnim zawodowym
  - h) B\_Z                 bezrobotni z wykształceniem zawodowym
- jest uzasadnieniem przestawionej opinii.