

## Analyse Multikollinearität am Beispiel der Abhängigkeit der Schweizer Exporte von der Weltkonjunktur

Die Exporte der Schweiz sollen mit der BIP-Entwicklung in den wichtigsten Abnehmerländern und einem Mass für die preisliche Wettbewerbsfähigkeit der Schweizer Exporteure erklärt werden.

Quartalsdaten 1980Q1 – 2008Q3

EXTOT	Gesamtexport real (zu Preisen von 2000)
GDPFRA	reales BIP Frankreich (Index 1995=1)
GDPGER	reales BIP Deutschland (Index 1995=1)
GDPITA	reales BIP Italien (Index 1995=1)
GDPUKI	reales BIP Grossbritannien (Index 1995=1)
GDPUSA	reales BIP USA (Index 1995=1)
GDPJAP	reales BIP Japan (Index 1995=1)
PEXTOT	Exportpreis in CHF (Index 2000=1)
PCOMP	Weltmarktpreis ausgedrückt in CHF (Index 2000=1)

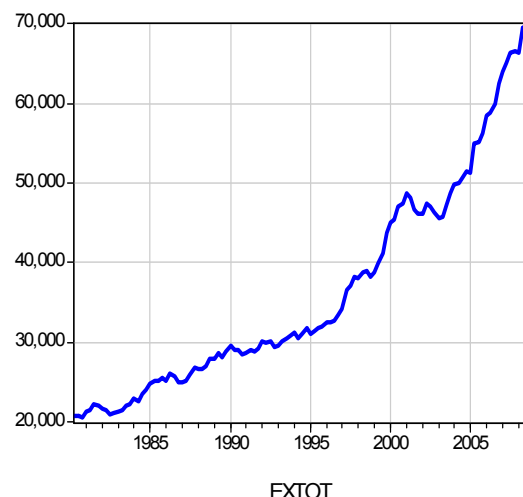
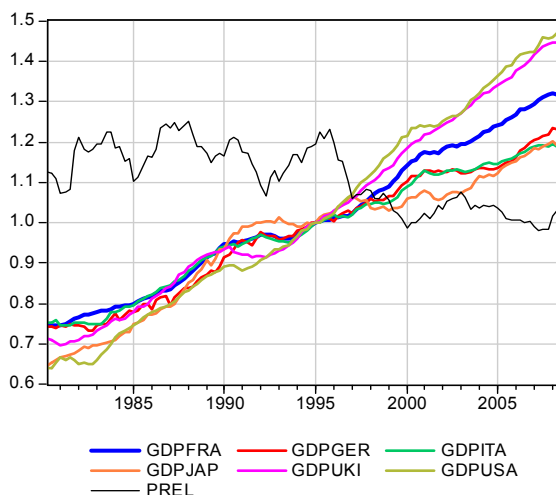
Die Regressionsgleichung wird als "partial adjustment"-Modell formuliert:

$$(1) \quad \log(\text{EXTOT}) = c(1) + c(2) \cdot \log(\text{GDPFRA}) + c(3) \cdot \log(\text{GDPGER}) + c(4) \cdot \log(\text{GDPITA}) \\ + c(5) \cdot \log(\text{GDPUKI}) + c(6) \cdot \log(\text{GDPUSA}) + c(7) \cdot \log(\text{GDPJAP}) \\ + c(8) \cdot \log(\text{PREL}) + c(9) \cdot \log(\text{EXTOT}(-1))$$

Die preisliche Wettbewerbsfähigkeit wird mit relativen Preis

$$\text{PREL} = \text{PEXTOT} / \text{PCOMP}$$

gemessen. Entsprechend wird für  $c(8)$  ein negatives Vorzeichen erwartet.



Ein mögliches Multikollinearitätsproblem besteht in der vermutlich sehr ähnlichen Entwicklung der sechs ausländischen Konjunkturvariablen. Erste Hinweise darauf liefern die paarweisen Korrelationen zwischen diesen Variablen. Die logarithmierten Variablen, wie sie in der Exportgleichung erscheinen, sind im Folgenden mit einem vorangestellten L bezeichnet, d.h. es gilt z.B.  $\text{LGDPFRA} = \log(\text{GDPFRA})$ .

**Korrelationsmatrix der erklärenden Variablen**

	LGDPFRA	LGDPGER	LGDPITA	LGDPJAP	LGDPUKI	LGDPUSA	LPREL
LGDPFRA	1.000						
LGDPGER	0.989	1.000					
LGDPITA	0.991	0.994	1.000				
LGDPJAP	0.952	0.974	0.977	1.000			
LGDPUKI	0.995	0.978	0.985	0.938	1.000		
LGDPUSA	0.995	0.986	0.990	0.949	0.997	1.000	
LPREL	-0.771	-0.746	-0.726	-0.633	-0.761	-0.769	1.000

Alle ausländischen BIPs sind hoch miteinander korreliert. Am stärksten ist die Korrelation zwischen Grossbritannien und den USA (0.997). Der relative Preis ist eine relativ unabhängige Variable.

**Multiple Korrelation unter den erklärenden Variablen**

Sample: 1980Q2 2008Q3

Included observations: 114

LOG(GDPGER)= C(1)+C(2)\*LOG(GDPFRA)+C(3)\*LOG(GDPITA)+C(4)  
\*LOG(GDPUKI)+C(5)\*LOG(GDPUSA)+C(6)\*LOG(GDPJAP)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.005155	0.002050	-2.515074	0.0134
C(2)	0.692265	0.084279	8.213993	0.0000
C(3)	0.522915	0.103377	5.058348	0.0000
C(4)	-0.708452	0.079651	-8.894471	0.0000
C(5)	0.422802	0.066257	6.381212	0.0000
C(6)	0.056529	0.038718	1.460019	0.1472
R-squared	0.994654	Mean dependent var		-0.037550
Adjusted R-squared	0.994407	S.D. dependent var		0.162503
S.E. of regression	0.012153	Akaike info criterion		-5.931222

Das deutsche BIP lässt sich mit einem  $R^2$  von 0.994654 mit den BIPs der anderen Länder erklären.

Aufgrund der hohen Korrelationen unter den erklärenden Variablen unterliegt die Exportgleichung (1) einem ausgeprägten Multikollinearitätsproblem. Schätzt man die Gleichung trotzdem ohne irgendwelche Parameterrestriktionen, so ergibt sich:

**Exportgleichung (ohne Restriktionen)**

Sample (adjusted): 1980Q2 2008Q3

Included observations: 114 after adjustments

LOG(EXTOT)=C(1)+C(2)\*LOG(GDPFRA)+C(3)\*LOG(GDPGER)

C(4)\*LOG(GDPITA)+C(5)\*LOG(GDPUKI)+C(6)\*LOG(GDPUSA)

C(7)\*LOG(GDPJAP)+C(8)\*LOG(PEXTOT/PCOMP)+C(9)\*LOG(EXTOT(-1))

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	1.822675	0.448787	4.061337	0.0001
C(2)	-0.095377	0.182314	-0.523147	0.6020
C(3)	0.016120	0.137941	0.116862	0.9072
C(4)	-0.530059	0.180698	-2.933395	0.0041
C(5)	0.430043	0.178429	2.410163	0.0177
C(6)	0.080139	0.123045	0.651296	0.5163
C(7)	0.157744	0.063665	2.477730	0.0148
C(8)	-0.226986	0.054174	-4.189981	0.0001
C(9)	0.827582	0.043067	19.21615	0.0000
R-squared	0.997643	Mean dependent var	10.43857	
Adjusted R-squared	0.997463	S.D. dependent var	0.343904	
S.E. of regression	0.017322	Akaike info criterion	-5.198029	

Die in Form der Elastizitäten C(2) bis C(7) gemessenen Einflüsse der ausländischen Konjunkturvariablen sind teilweise unplausibel, mit hohen Standardfehlern unpräzise geschätzt und insignifikant. Der mit C(2) gemessene Einfluss von Frankreich auf die Schweizer Exporte ist z.B. insignifikant negativ, der mit der Elastizität C(3) gemessene Einfluss von Deutschland auf die Schweizer Exporte ist insignifikant positiv. Aufgrund der hohen Multikollinearität lassen sich die Einzeleinflüsse der ausländischen Konjunkturvariablen nicht zuverlässig bestimmen. Der "Variance Inflation Factor" für z.B. C(3) beträgt:

$$VIF = \frac{1}{1 - R_{GDPGER}^2} = \frac{1}{1 - 0.994654} = 187.1$$

(Das  $R_{GDPGER}^2$  stammt aus der Regression von GDPGER auf die anderen erklärenden Variablen.)

Der Standardfehler von C(3) ist aufgrund der hohen Korrelation von GDPGER mit den anderen erklärenden Variablen um den Faktor  $\sqrt{187.1} = 13.7$  aufgebläht. Den kleinsten Standardfehler weist C(7) auf (Elastizität der Schweizer Exporte in bezug auf die Konjunktur in Japan). Dies ist damit zu erklären, dass das japanische BIP sich "relativ unabhängig" entwickelte.

Weiter sind z.B. die Parameter C(5) und C(6) wegen der hohen Korrelation zwischen LGDPUKI und LGDPUSA von 0.997 gut "gegeneinander austauschbar": Je tiefer man C(5) festlegt, desto grösser und signifikanter wird C(6). Setzt man z.B. - wie in der folgenden Gleichung gezeigt - C(5) für Grossbritannien auf 0.2 herunter (was angesichts des hohen Standardfehlers eine statistisch zulässige Restriktion ist), so wird dafür C(6) für die USA grösser und signifikant. Auch andere Parameter ändern sich etwas. Das  $R^2$  der Gleichung bleibt dabei praktisch unverändert.

Sample (adjusted): 1980Q2 2008Q3

Included observations: 114 after adjustments

LOG(EXTOT)=C(1)+C(2)\*LOG(GDPFRA)+C(3)\*LOG(GDPGER)+C(4)

\*LOG(GDPITA)+0.2\*LOG(GDPUKI)+C(6)\*LOG(GDPUSA)+C(7)

\*LOG(GDPJAP)+C(8)\*LOG(PEXTOT/PCOMP)+C(9)\*LOG(EXTOT(-1))

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	1.627873	0.423905	3.840180	0.0002
C(2)	0.037484	0.150865	0.248461	0.8043
C(3)	-0.085277	0.113677	-0.750169	0.4548
C(4)	-0.450605	0.170394	-2.644487	0.0094
C(6)	0.205684	0.075454	2.725941	0.0075
C(7)	0.132386	0.060739	2.179578	0.0315
C(8)	-0.191021	0.046584	-4.100564	0.0001
C(9)	0.845983	0.040760	20.75519	0.0000
R-squared	0.997605	Mean dependent var		10.43857
Adjusted R-squared	0.997447	S.D. dependent var		0.343904
S.E. of regression	0.017376	Akaike info criterion		-5.199867

## Hauptkomponentenanalyse

Principal Components Analysis

Date: 02/05/09 Time: 16:22

Sample: 1980Q2 2008Q3

Included observations: 114

Computed using: Ordinary correlations

Extracting 7 of 7 possible components

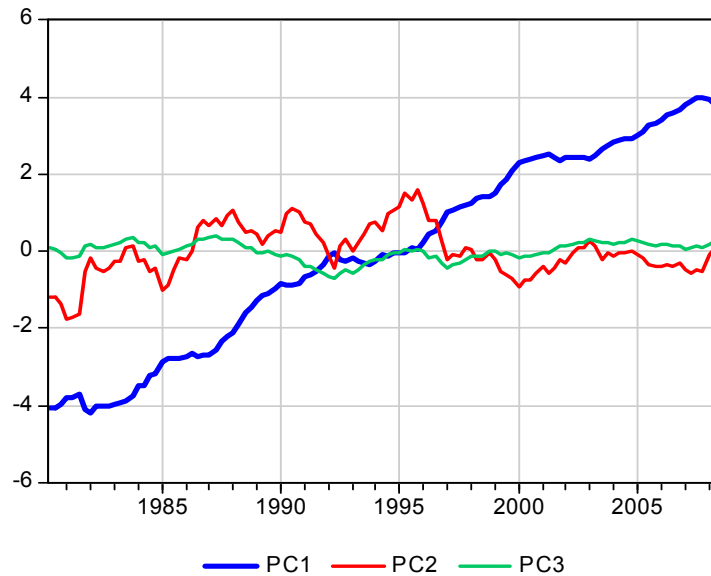
Eigenvalues: (Sum = 7, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	6.487402	6.050252	0.9268	6.487402	0.9268
2	0.437150	0.379833	0.0625	6.924552	0.9892
3	0.057317	0.048708	0.0082	6.981869	0.9974
4	0.008609	0.003614	0.0012	6.990478	0.9986
5	0.004995	0.001393	0.0007	6.995473	0.9994
6	0.003602	0.002676	0.0005	6.999075	0.9999
7	0.000925	---	0.0001	7.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7
LGDPFRA	0.391180	0.041129	0.244584	0.104127	0.749653	0.092255	0.451810
LGDPGER	0.389935	0.111567	-0.209963	0.769959	-0.118699	0.324903	-0.280945
LGDPITA	0.389846	0.157459	-0.048163	0.144470	-0.166554	-0.873584	0.095639
LGDPJAP	0.376526	0.345048	-0.672729	-0.499541	0.087670	0.161787	-0.056602
LGDPUKI	0.388938	0.046755	0.535545	-0.301531	0.070222	0.014668	-0.680930
LGDPUSA	0.390809	0.041694	0.315119	-0.165139	-0.619332	0.308941	0.489839
LPREL	-0.311726	0.915470	0.234737	0.087842	0.001010	0.030716	0.031292

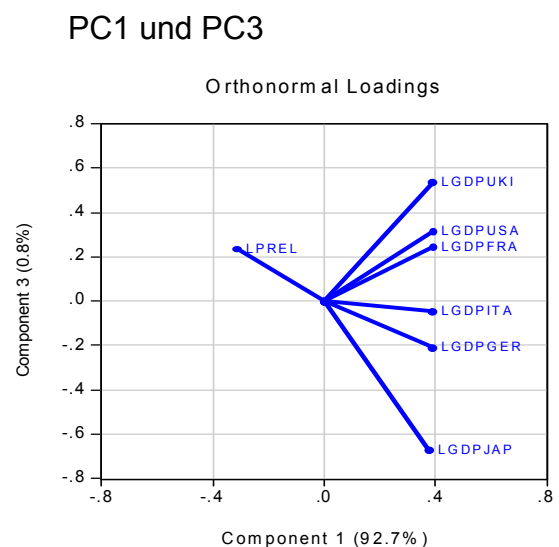
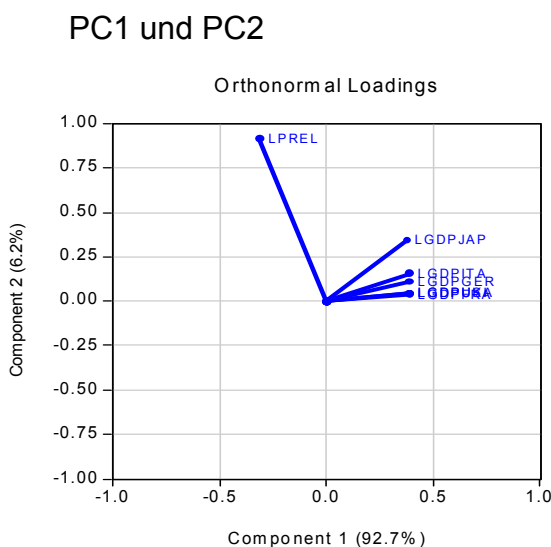
Die ersten drei Hauptkomponenten fangen 99.74% der Varianz der 7 Variablen auf. Deren Informationsgehalt lässt sich also im Wesentlichen auf 3 Faktoren reduzieren.



Die erste Hauptkomponente (PC1) kann als durchschnittliche Konjunkturentwicklung in den sechs Abnehmerländern interpretiert werden. Die zweite Hauptkomponente (PC2) widerspiegelt den relativen Preis, die dritte Hauptkomponente (PC3) fängt Unterschiede in der Konjunktur zwischen den Ländern auf.

In invertierter Betrachtung zeigt die Analyse, wie die Variablen als gewichtete Kombination aus den Hauptkomponenten aufgebaut sind. Die Hauptkomponenten werden auch "Faktoren" und die Gewichte "Faktorladungen" genannt. PC1 geht z.B. mit fast identischen Ladungen von rund 0.38 in die sechs Ländervariablen ein. Unterschiede zwischen den Ländervariablen zeigen sich vor allem bei der Gewichtung von PC3. Wenn PC3 negativ ist (z.B. von 1991-1993), entwickelt sich Japan besser und Grossbritannien schlechter als der Durchschnitt. Der von PC3 erklärte Varianzanteil beträgt aber nur 0.82%.

Die Faktorladungen der sieben Variablen in bezug auf die ersten drei Hauptkomponenten lassen sich mit EViews wie folgt veranschaulichen:



Was folgt aus der Hauptkomponentenanalyse hinsichtlich der Exportgleichung? Der Informationsgehalt in den sieben erklärenden Variablen dieser Gleichung lässt sich fast vollständig auf drei Faktoren reduzieren (PC1: Weltkonjunktur, PC2: relativer Exportpreis, PC3: Abweichungen der Länder von der Weltkonjunktur). Der Versuch, die Einflüsse der sechs Abnehmerländer auf die Schweizer Exporte separat zu bestimmen, scheitert an der Ähnlichkeit der Länderkonjunkturen. Die Analyse legt eine Zusammenfassung zu einer einzigen Variablen "Weltkonjunktur" nahe. Daneben kann man problemlos den relativen Preis in die Gleichung aufnehmen, denn diese Variable erweist sich als ziemlich eigenständig. Ob Unterschiede in den Länderkonjunkturen, wie sie von PC3 aufgefangen werden, für die Schweizer Exporte eine Rolle spielen, ist im Rahmen der Gleichungsschätzung abzuklären. PC3 fängt nur 0.82% der Gesamtvarianz der sieben Ausgangsvariablen auf, kann also möglicherweise ebenfalls vernachlässigt werden.

Stimmt man dieser Interpretation zu, so kann man die Exportgleichung unter Verwendung der Hauptkomponenten wie folgt schätzen:

Sample (adjusted): 1980Q2 2008Q3

Included observations: 114 after adjustments

$\text{LOG}(\text{EXTOT}) = C(1) + C(2) * \text{PC1} + C(3) * \text{PC2} + C(4) * \text{PC3} + C(5) * \text{LOG}(\text{EXTOT}(-1))$

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	1.274680	0.397139	3.209658	0.0017
C(2)	0.017211	0.004990	3.449393	0.0008
C(3)	-0.010572	0.003043	-3.473939	0.0007
C(4)	0.022185	0.010979	2.020713	0.0458
C(5)	0.878771	0.038083	23.07498	0.0000
R-squared	0.997198	Mean dependent var		10.43857
Adjusted R-squared	0.997095	S.D. dependent var		0.343904
S.E. of regression	0.018536	Akaike info criterion		-5.095336

Alternativ kann man auch mit selbstgewählten Gewichten eine Variable GDPWORLD bilden und die Gleichung wie folgt schätzen:

Sample (adjusted): 1980Q2 2008Q3

Included observations: 114 after adjustments

$\text{LOG}(\text{EXTOT}) = C(1) + C(2) * \text{LOG}(\text{GDPWORLD}) + C(8) * \text{LOG}(\text{PEXTOT}/\text{PCOMP}) + C(9) * \text{LOG}(\text{EXTOT}(-1))$

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.919842	0.309233	2.974591	0.0036
C(2)	0.112004	0.042195	2.654455	0.0091
C(8)	-0.150876	0.043879	-3.438437	0.0008
C(9)	0.914436	0.029385	31.11944	0.0000
R-squared	0.997156	Mean dependent var		10.43857
Adjusted R-squared	0.997078	S.D. dependent var		0.343904
S.E. of regression	0.018589	Akaike info criterion		-5.098038