COWLES FOUNDATION DISCUSSION PAPER NO. 12

Note: Cowles Foundation Discussion Papers are preliminary materials circulated privately to stimulate private discussion and critical comment. References in publications to Discussion Papers (other than mere acknowledgment by a writer that he has access to such unpublished material) should be cleared with the author to protect the tenative character of these papers.

The Klein-Goldberger forecasts for 1951, 1952 and 1954, Compared With Naive-Model Forecasts.

Tjalling C. Koopmans

May 10, 1956

The Klein-Goldberger Forecasts for 1951, 1952 and 1954, Compared With Naive-Model Forecasts

In their recent book, "An Econometric Model of the United States, 1929 - 1952" (Amsterdam, 1955), Klein and Goldberger fitted a system of 20 equations (behavioral, techological, and definitional) containing 20 endogenous and 18 exogenous variables, to data for 1929 -'41 and 1946 -'50 and made " ex-post " forecasts for 1951 and 1952 of the current endogenous variables, using observed values for the exogenous and lagged endogenous variables. They also made actual forecasts of the endogenous variables for 1953 and 1954, using their own best projections for the values of the exogenous variables their years, and using for the 1954 forecast a re-estimated model based on data for 1929-'41 and 1946 -'52.

Klein and Goldberger do not report any comparisons between the forecasting performance of this model and naive-model forecasts such as were proposed by Friedman and applied by Andrew Marshall and by Carl Christ to an earlier model of Klein*. For the 1951 and 1952 ex-post forecasts, the

---

* See "Conference on Business Cycles, "Nat. Bur. of Ec. Res., New York 1951, Paper by Christ and discussion by Friedman and by Klein.

---

comparison can be made readily from the data supplied by Klein and Goldberger. I am circulating these comparisons to save my colleagues the work of making them for themselves.

The following two "naive models" are considered

N_I: $\quad y^*_{t+1} = y_t$

N_{II}: $\quad y^*_{t+1} = y_t + (y_t - y_{t-1})$

Here y^*_{t+1} is the forecast value of the endogenous variable x for the year t + 1, and y_t and y_{t-1} are observed values for the years t and t - 1. These are compared with two alternative Klein-Goldberger forecasts

K_I: $y^*_{t+1} = f\,(y_t^{(1)},\ldots, y_t^{(20)}; y_{t-1}^{(1)},\ldots; z_{t+1}^{(1)},\ldots,z_{t+1}^{(18)}; z_t^{(1)},\ldots)$

$= f(----)_t$ (abbreviated)

K_{II}: $y^*_{t+1} = f\,(----)_t + v^*_t$, (where $v^*_t = y_t - f\,(---)_{t-1}$ if the model were linear).

In the first (K_I) the forecast value of any one of the twenty endogenous variables, denoted y^*_{t+1}, is solved from the equation system on the "assumption" that all disturbances (random terms) in all equations are going to be zero in the year t + 1. In the second (K_{II}), each of the endogenous variables for the year t + 1 is solved on the "assumption" that all disturbances in the year t + 1 are going to have the same value as they were found to have in the year t (i.e., the disturbances in the structural equations)

In applying N_{II} to 1951, the difficulty arises that data for 1949 are not given by Klein and Goldberger on the same basis as those for 1950 and 1951 used in the ex-post forecast, but only on the revised basis used for the forecast for 1954. Denoting data on the revised basis by $\bar{x}$, the application of N_{II} to 1951 has therefore been made from the formula

$\bar{N}_{II}$ $\quad x^*_{t+1} = x_t + (\bar{x}_t - \bar{x}_{t-1})$.

Klein and Goldberger give forecasts for 14 of the 20 endogenous variables, and note that since between these 14 variables there are 4 definitional

identities, only 10 random disturbances are operative in the comparisons. Table 1 gives the comparison in numerical terms. The first of the last five lines counts for each year the number of cases in which N_I is at least as close (in absolute terms) to the observed value as K_I and conversely. The next three lines make similar comparisons for N_I versus K_{II}, N_{II} versus K_{II}, and N_{II} versus K_I. The last line counts number of cases where any given method is at least as close as any of the four. Page references give sources of data in the K.-G.-book.

Similar comparisons for the 1953 and 1954 forecasts would require data on the observed values for these years, which are not given in the book. However, the Preface gives a table that permits a comparison with N_I for 12 variables for the year 1954 only. This table is reproduced as table 2 with score columns added. The Klein-Goldberger forecast is labeled K because the choice between K_I and K_{II} has been made by the authors for each variable separately. The Klein-Goldberger forecasts for 1954 stand up better in the comparison than the 1951 and 1952 ex-post forecasts, in spite of the fact that they are based on projected rather than on observed values of exogenous variables. The possibility that this is connected with the fact that 1954 was a year of mild downturn in activity constitutes the first ray of hope for structural estimation in a long time. For the same reason, N_{II} is likely to do worse than N_I for 1954.

The crudeness of the scoring system hardly needs emphasis. However, direct inspection of the numerical estimates roughly confirms the scores in the cases presented, except that the 1954 appears better than the score indicates, considering also the nature of the variables for which forecasting success is highest.

Table 2 Comparison of Klein-Goldberger 1954 Forecasts With Naive Model Forecasts

(1) Variable and symbol	(2) Observed change 1953 to 1954	(3) Forecast change as of Dec. 53	(4) Forecast change as of June 54	(5) Scores for Dec. 53 forecast N_I	(5) K	(6) Scores for June 54 forecast N_I	(6) K
GNP Y+T+D	-4.6	-5.3	-5.5		*		*
Consumption C	0.8	1.2	0.7		*		*
Investment I	-2.7	-3.3	-2.6		*		*
Imports F_I	-0.3	-0.2	0.0		*	*	*
Private wage bill W_1	-1.2	-2.4	-2.3	*	*		*
Nonwage nonfarm income P	-1.8	-0.1	0.2		*	*	
Farm income A	-0.2	0.5	0.8	*		*	
Depreciation D	1.1	-2.6	-2.2	*		*	
Employment N_W	-0.9	-0.3	-0.9		*		*
Price Level P	0.3	16.1	14.2	*		*	
Wage rate w	10.9	18.1	18.4		*		*
			N_I vs K	4	8	5	7

Units: Billions of $, except for employment (millions of persons) price level and wage rate (index points)

Table 1. Comparison of Klein-Goldberger and naive-model forecasts, 1951, 1952.

(1) Variable and Symbol	(2) 1949 revised p. 131	(3) 1950 observed p. 131	(4) 1950 observed p. 81	(5) 1951 observed p. 81	(6) 1951 forecasts N_I =(4)	(7) 1951 forecasts N_{II} =(4)+(3)-(2)	(8) K_I p.81	(9) K_{II} p.81	(10) 1952 observed p.81	(11) 1952 forecasts N_I =(5)	(12) 1952 forecasts N_{II} =(5)+(5)-(4)	(13) K_I p.81	(14) K_{II} p.81
Consumption C	103.2.	108.9	108.7	108.4	108.7	114.4	108.7	112.7	110.2	108.4	108.1	113.0	112.7
nonwage nonform P income	31.39	35.61	34.9	37.0	34.9	39.1	33.6	37.2	35.5	37.0	39.1	35.7	39.1
depreciation D	14.75	16.25	15.9	17.7	15.9	17.4	15.3	17.0	20.6	17.7	19.5	16.6	19.0
corporate savings S_P	3.92	2.27	2.28	1.43	2.28	0.63	-1.04	1.55	2.30	1.43	0.58	0.92	3.39
investment I	18.0	26.8	24.8	28.0	24.8	33.6	20.2	24.2	23.8	28.0	31.2	19.5	27.3
corporate profits Pc	14.09	17.09	17.4	18.2	17.4	20.4	15.6	19.1	17.8	18.2	19.0	17.1	19.7
private wage bill W_1	65.84	70.85	70.9	75.6	70.9	75.9	72.7	76.1	77.2	75.6	80.3	77.3	80.2
labor force N_W	48.9	50.7	50.7	54.3	50.7	51.5	53.6	55.1	55.2	54.3	57.9	57.7	58.4
wage rate W	276.9	286.9	287.1	309.9	287.1	297.1	304.1	305.4	326.2	309.9	332.7	334.9	340.7
imports F_I	4.6	5.4	4.8	4.2	4.8	5.6	4.0	5.0	4.6	4.2	3.6	4.0	4.2
GNP Y+T+D	144.0	156.2	154.3	167.3	154.3	166.5	159.9	166.9	170.8	167.3	180.3	169.9	177.3
national income Y	116.6	126.6	126.1	135.9	126.1	136.1	130.9	136.1	137.4	135.9	145.7	140.3	143.3
price level P	179.4	183.6	183.2	196.8	183.2	187.4	196.8	195.6	202.0	196.8	210.4	216.6	216.6
farm income A	7.31	7.52	7.52	7.98	7.52	7.54	9.29	7.54	7.70	7.98	8.44	10.3	9.00
N_I vs K_I					7		8			10		4	
N_I vs K_{II}					4			10		11			4
N_{II} vs K_{II}						4		12			8		7
N_{II} vs K_I						9	5				5	9	
best of 4 methods					3	4	4	6		8	2	3	2