Trends in Soviet labour productivity, 1928–1985: war, postwar recovery, and slowdown

MARK HARRISON

Department of Economics
University of Warwick
Coventry, England CV4 7AL
+44 1203 523030 (direct)
+44 1203 523032 (fax)
+44 1203 523055 (secretary)
Mark.Harrison@warwick.ac.uk

Appendices

A. Basic data and sources

B. Unit root tests

C. Regression equations and symbols

Date of last revision: 20 February, 1998

Appendix A. Basic data and sources

A.1. GNP

The basic series are obtained as below. They are not the only available growth estimates, but they are the only ones suitable for dynamic testing. My view is that they are also the most reliable. More detailed discussion, omitted here for reasons of space, can be found in a earlier draft of this article circulated as Harrison (1996b) and available from the author.

GNP at 1937 factor cost

1928–39, 1946–62	Moorsteen and Powell 1966, pp. 622–4.
1940–5	Harrison 1996a, p. 92.
1962–6	Becker, Moorsteen, and Powell 1968, p. 51.

GNP at 1982 factor cost

1950–85	CIA 1990, pp. 54–7.

A.2. Industry

Value added at 1937 factor costs

1928–39, 1946–62	Moorsteen and Powell 1966, pp. 622–4.
1940–1945	Harrison 1996a, p. 92.
1962–6	Becker, Moorsteen, and Powell 1968, p. 51.

Value added at 1982 factor costs

1950–85	CIA 1990, pp. 54–7.

A.3. Employment and hours

Total employment, full–time equivalents (whole economy)

1928–40, 1945–60	Moorsteen and Powell 1966, p. 643.
1941–4	Powell 1968, p. 35.
1961–6	Becker, Moorsteem, and Powell 1968, p. 57 (there is some break of continuity between this section of the series and the figure for 1961 provided by Moorsteen and Powell 1966, p. 643).
1950–85	Rapawy 1987, pp. 194–5, 202–3 (armed forces personnel, plus total employment in the national economy).

Industrial employment in full–time equivalents

1928–50	For 1928, 1932, 1937, 1940, figures are as from Barber and Davies 1994, p. 282. For other years before 1940 and for 1946–50 employment in largescale industry is given by Zaleski 1971, pp. 342–3, and Zaleski 1980, pp. 562–3, 590–1, 628–9 (before 1930 these figures must be adjusted to a calendar–year basis), and for 1941–5 from Harrison 1996a, p. 256 (table G–2). Artisan employment in

	1928, 1932, 1937, and 1940 is calculated as the difference between employment figures for industry as a whole and for largescale industry by Barber and Davies 1994, p. 282. For 1941–5, employment in largescale and artisan industry is taken from Harrison 1996a, p. 256 (table G–3), and for 1950 from Rapawy 1987, p. 202. Between other benchmark years, artisan employment is estimated by geometric interpolation. Total employment in industry is then employment in large–scale industry plus artisan employment in all years.
1950–85	Employment in industry, plus independent artisans, from Rapawy 1987, p. 202.

Average annual hours worked in industry

1928–50	Average days multiplied by average hours in largescale industry, from Moorsteen and Powell 1966, p. 647, except 1940–5 from Harrison 1996a, p. 259.
1950–85	Total work–hour employment in industry and of independent artisans, divided by total numbers employed, calculated from Rapawy 1987, pp. 202–3, 210–11.

Appendix B. Unit root tests

The procedure I follow with each series is first to check whether its upward path can be described as a trend rather than a drift. Let's write down the path of y as:

1. $$y_t = \alpha_0 + \alpha_1 \cdot y_{t-1} + \alpha_2 \cdot t + \sum_{j=1}^{p} \delta_i \cdot \Delta y_{t-i} + \varepsilon_t$$

where t is time and the lag length p is chosen such that the error term ε is approximately white noise. The value of α_1 tells us about the persistence of shocks. As long as $\alpha_1 < 1$, the influence of random fluctuations will tend to die away, the more rapidly the smaller is α_1. In the long run, the behaviour of y will be dominated by $\alpha_2 \cdot t$, the time trend.

Upward drift describes the path of y when $\alpha_1 = 1$ (when y has a unit root). In this case, any random fluctuation will not die away but will prove to be permanent. When there is a unit root, then as long as $\alpha_2 > 0$, y may well move upwards through time under the influence of $\alpha_2 \cdot t$, but its path will follow a drift, not a trend. In this case we would not be able to identify a 'normal' peacetime trendline from which Soviet growth series were disturbed by war or postwar retardation.

In testing for a unit root we can subtract y_{t-1} from both sides to give equation 1 in an alternative form:

2. $$\Delta y_t = \alpha_0 + \beta_1 \cdot y_{t-1} + \alpha_2 \cdot t + \sum_{j=1}^{p} \delta_i \cdot \Delta y_{t-i} + \varepsilon_t, \quad \beta_1 = \alpha_1 - 1$$

This is the form of the model tested initially. A unit–root to y is not rejected if β_1 is not significantly different from zero, which simplifies somewhat the testing procedure. Critical values used to test significance are MacKinnon's for β_1 and Dickey–Fuller for α_2 .

Table B–1 (col. 1) shows a unit–root test using equation 2 for GNP per worker with $p = 0$; these results have also been tested for various values of p, and appear to be robust. In the present case β_1 is estimated at not far from zero, with a t–ratio which is well short of the critical values for 5 or 10 per cent significance in a unit–root test. MacKinnon's critical value for 10 per cent significance of β_1 in a constant–plus–trend model with a sample size of 57 is -3.17. The Dickey–Fuller equivalent for the α_2 coefficient of the time trend t is 2.38, so α_2 is also not significant. At this stage a unit–root is not rejected.

The long–run GNP series of many countries display a unit root; however, even when this is the case at first sight, the unit root usually disappears once a segmented trend is allowed for, with periodic, but infrequent, large exogenous shocks giving rise to permanent breaks of trend and/or level (e.g. Crafts, Mills (1996)). Critical values were tabulated by Perron (1979) for three models. specifying a break of level (Model A), a break of trend (Model B), and a simultaneous break of both level and trend (Model C). However, as will become clear below, if the unit–root hypothesis is to be excluded we must allow for more than one break in the Soviet series between 1928 and 1985. For present purposes, since critical values are not available for such a model, I divide the sample into two overlapping subperiods.

In table B–1 (col. 2) I test Perron's Model C (a simultaneous break of both level and trend) on the period 1929–65. The record shows that, on this hypothesis, there was a negative exogenous shock to the productivity level in 1945, and at the same time an acceleration of the productivity trend from its new, lower level. The point of hypothetical trend break, and the final year of the sample, are found endogenously by searching for maximum values of the t–ratio for β_1 and the F–statistic respectively. The estimated value of β_1 for this subperiod falls to -0.79 which is significantly less than zero at 1 per cent, so over this period a unit root is rejected. This result does not depend on the autoregressive order (p) of the estimating model, and proved robust for varying lag lengths of the dependent variable. Thus in the long run growth was significantly accounted for by an upward trend.

Table B–1 (col. 3) deals with the second subperiod from 1945 (the year of the crash and trend–break identified in col. 2) to the end of the sample. Here I test Perron's Model B (a break of trend). The break point is identified endogenously and is located in 1970. The estimated value of β_1 for the second subperiod is -0.72 which is significantly less than zero at 1 per cent, so again we can reject a unit root. Again this result proved robust for varying lag lengths when $p > 0$.

Table B–2 shows results for industry value added per worker and per hour worked. In each case, as for GNP per worker, on an initial test (cols 1 and 3) the unit–root hypothesis is not rejected.

In the case of industry value added per worker, when a break of level in 1945 is permitted (Perron's Model A, shown in col. 2), the β_1 coefficient becomes less than zero at the 2.5 per cent level, which allows us to reject the unit–root hypothesis straight away. This result also held for varying lag lengths.

The case of industry value added per hour is more doubtful. As in the case of GNP per worker, a unit root cannot be rejected for the full sample on the basis of a Perron–type segmented–trend model with a single break. I proceed by dividing the full sample into subperiods over which a Perron–type model may be tested. However, there is also no subperiod beginning in 1929 over which it is possible to reject a unit root. To put the argument at its strongest: in the 1930s, productivity shocks which are potentially candidates to satisfy the conditions of being large, infrequent, and exogenous arrive too frequently to distinguish them statistically from random disturbances.

The first run of years over which it is possible to reject a unit root begins in 1932 and runs to 1944. Within this period a break of both level and trend (Perron's Model C, shown in col. 4) may be identified in 1939, i.e. in that year the productivity level collapsed and growth became negative. On this hypothesis it is possible just to reject a unit root at the 10 per cent level. However, neither the crash nor the trend–break coefficient is significant even at this level. A unit root is not rejected over this period if sufficient lag lengths ($p \geq 2$) are allowed, but this is due to the reduced power of the test when the number of insignificant variables is increased.

Over the postwar period 1945–85, the unit–root hypothesis can be rejected very confidently on the basis of Perron's Model A (a break of trend growth), shown in col. 5. The most likely point of trend break marking postwar deceleration is found relatively early, in 1962.

Taking an overview of the results for industry value added per hour, a unit root is rejected for the whole period only as long as the years 1929–31 are excluded. However it should be borne in mind that Perron's critical values were not designed to test series with more than one point of segmentation. It seems likely that more purpose–built critical values would permit rejection of a unit root for the whole period. Nonetheless it has to be admitted that in this aspect the testing of my null hypothesis remains indecisive. An interpretation of the outcome is offered in the text.

Table B–1. GNP per worker, 1928–85: unit–root tests

	Col. 1	Col. 2	Col. 3
Dep. var.	*Dgnp/wkr*	*Dgnp/wkr*	*Dgnp/wkr*
R–squared	0.0550	0.5707	0.5367
Obs.	57	37	41
Deg. freedom	54	32	37
F–statistic	1.5720	10.6360	14.2881
Constant	–3.7351	–11.4312	–45.8144
	gnp/wkr–1	*gnp/wkr*–1	*gnp/wkr*–1
X–coeff.	–0.0899	–0.7899	–0.7223
SE	0.0521	0.1322	0.1170
t–statistic	–1.7257	–5.9756	–6.1742
	t	*t*	*t*
X–coeff.	0.0019	0.0059	0.0235
SE	0.0011	0.0019	0.0038
t–statistic	1.7725	3.0205	6.2563
		crash(45)	*trend*(70)
X–coeff.	..	–0.1132	–0.0172
SE	..	0.0283	0.0027
t–statistic	..	–4.0040	–6.2888
		trend(45)	
X–coeff.	..	0.0211	..
SE	..	0.0035	..
t–statistic	..	5.9521	..

Notes
Col. 1 Full sample.
Col. 2 1929–65.
Col. 3 1945–85.
Regression variables and equations are listed and defined in appendix C.

Table B–2. Industry value added per worker and per hour worked, 1929–85: unit–root tests

	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5
Dep. var.	*Dind/wkr*	*Dind/wkr*	*Dind/hr*	*Dind/hr*	*Dind/hr*
R–squared	0.1342	0.2869	0.1156	0.7463	0.6764
Obs.	57	57	57	13	41
Deg. freedom	54	53	54	8	37
F–statistic	4.1836	7.1090	3.5298	5.8834	25.7753
Constant	–11.2047	–23.7901	–8.5401	–110.4386	–67.7708
	ind/wkr–1	*ind/wkr*–1	*ind/hr*–1	*ind/hr*–1	*ind/hr*–1
X–coeff.	–0.2155	–0.3570	–0.1318	–0.9003	–0.5535
SE	0.0770	0.0821	0.0547	0.2142	0.0701
t–statistic	–2.7972	–4.3474	–2.4107	–4.2028	–7.8989
	t	*t*	*t*	*t*	*t*
X–coeff.	0.0058	0.0123	0.0044	0.0570	0.0348
SE	0.0020	0.0027	0.0017	0.0151	0.0040
t–statistic	2.8806	4.6090	2.6518	3.7679	8.6446
		crash(45)		*crash*(39)	*trend*(62)
X–coeff.	..	–0.1434	..	–0.1868	–0.0217
SE	..	0.0426	..	0.0809	0.0026
t–statistic	..	–3.3698	..	–2.3097	–8.2257
				trend(39)	
X–coeff.	..	..	..	–0.0373	..
SE	..	..	..	0.0229	..
t–statistic	..	..	..	–1.6302	..

Notes:
Cols 1–3 Full sample.
Col. 4 1932–44.
Col. 5 1945–85.
Regression variables and equations are listed and defined in appendix C.

Appendix C. Regression equations and symbols

C.1. Real variables (all in natural logarithms)

gnp(37*fc*)	gross national product at 1937 factor cost
gnp(82*fc*)	gross national product at 1982 factor cost
gnp/wkr	adjusted gross national product per worker
hours	average hours worked in industry
ind(37*fc*)	industry value added at 1937 factor
ind(82*fc*)	industry value added at 1982 factor cost
ind/hr	adjusted industry value added per hour worked
ind/wkr	adjusted industry value added per worker

C.2. Time variables

t time (calendar year)

crash(*T*) break of level in year *T*

$$crash(T_i)_t = 0\ (t < T_i), 1\ (t \geq T_i)$$

crash+(*T*) break of level in year *T*

$$crash(T_i)_t = 0\ (t < T_i), 0.5\ (t = T_i), 1\ (t > T_i)$$

crash#(*T*) break of level in year *T* (persistent in degree *z*), followed by exponential return to preceding trend

$$crash\#(T_i)_t = 0\ (t < T_i), 0.5\ (t = T_i), z^{t-T_i-1}\ (t > T_i); \quad 0 < z < 1$$

trend(*T*) break of trend in year *T*

$$trend(T_i)_t = 0\ (t < T_i), t - T_i\ (t \geq T_i)$$

trend#(*T*) break of trend in year *T* followed after an interval (beyond which the departure from the preceding trend persists in degree *z*) by exponential return to preceding trend

$$trend\#(T_i)_t = 0\ (t < T_i), t - T_i\ (T_i \geq t \geq T_j), z \cdot trend\#(T_i)_{t-1}\ (t > T_j);$$
$$0 < z < 1$$

year(*T*) impulse variable, year T only

$$year(T_i)_t = 0\,(t < T_i, t > T_i), 1\,(t = T_i)$$

C.3. In general

1. $y_t = \alpha_0 + \alpha_1 \cdot y_{t-1} + \alpha_2 \cdot t + \sum_{j=1}^{p} \delta_i \cdot \Delta y_{t-i} + \varepsilon_t$

2. $\Delta y_t = \alpha_0 + \beta_1 \cdot y_{t-1} + \alpha_2 \cdot t + \sum_{j=1}^{p} \delta_i \cdot \Delta y_{t-i} + \varepsilon_t, \quad \beta_1 = \alpha_1 - 1$

C.4. Regression equations

Table 2

Col. 1 $gnp/wkr_t = \alpha_0 + \alpha_1 \cdot gnp/wkr_{t-1} + \alpha_2 \cdot t + \alpha_3 \cdot trend\#(T_3)_t + \alpha_4 \cdot trend(T_4)_t + \varepsilon_t$

T_3 is 1937 ($z = 0.93$), T_4 is 1973

Col. 2 $ind/wkr_t = \alpha_0 + \alpha_2 \cdot t + \Sigma\alpha_i \cdot crash\#(T_i)_t + \alpha_6 \cdot trend(T_6)_t + \varepsilon_t$

T_i ($i = 3, 4, 5$) correspond to 1931, 1942, 1945 (z = 0.93); T_6 is 1977

Col. 3 $ind/hr_t = \alpha_0 + \alpha_1 \cdot ind/hr_{t-1} + \alpha_2 \cdot t + \Sigma\alpha_i \cdot crash\#(T_i)_t + \alpha_6 \cdot trend(T_6)_t + \varepsilon_t$

T_i ($i = 3, 4, 5$) correspond to 1931, 1939, 1945 ($z = 0.92$); T_6 is 1976

Table B–1

Col. 1 $\Delta gnp/wkr_t = \alpha_0 + \beta_1 \cdot gnp/wkr_{t-1} + \alpha_2 \cdot t + \varepsilon_t$

Col. 2 $\Delta gnp/wkr_t = \alpha_0 + \beta_1 \cdot gnp/wkr_{t-1} + \alpha_2 \cdot t + \alpha_3 \cdot crash(T_3)_t + \alpha_4 \cdot trend(T_4)_t + \varepsilon_t$

T_3 is 1945, $T_4 = T_3$ (1929–65 only)

Col. 3 $\Delta gnp/wkr_t = \alpha_0 + \beta_1 \cdot gnp/wkr_{t-1} + \alpha_2 \cdot t + \alpha_3 \cdot crash(T_3)_t + \varepsilon_t$

T_3 is1970 (1945–85 only)

Table B–2

Col. 1 $\Delta ind/wkr_t = \alpha_0 + \beta_1 \cdot ind/wkr_{t-1} + \alpha_2 \cdot t + \varepsilon_t$

Col. 2 $\Delta ind / wkr_t = \alpha_0 + \beta_1 \cdot ind / wkr_{t-1} + \alpha_2 \cdot t + \alpha_3 \cdot crash(T_3)_t + \varepsilon_t$

T_3 is 1945

Col. 3 $\Delta ind / hr_t = \alpha_0 + \beta_1 \cdot ind / hr_{t-1} + \alpha_2 \cdot t + \varepsilon_t$

Col. 4 $\Delta ind / hr_t = \alpha_0 + \beta_1 \cdot ind / hr_{t-1} + \alpha_2 \cdot t + \alpha_3 \cdot crash(T_3)_t + \alpha_4 \cdot trend(T_4)_t + \varepsilon_t$

T_3 is 1939, $T_4 = T_3$ (1932–44 only)

Col. 5 $\Delta ind / hr_t = \alpha_0 + \beta_1 \cdot ind / hr_{t-1} + \alpha_2 \cdot t + \alpha_3 \cdot trend(T_3)_t + \varepsilon_t$

T_3 is1962 (1945–85 only)