Estimating Structural TFP Growth as a Latent Variable

Evidence from the NBER Manufacturing Productivity Database

BART HOBIJN
Department of Economics, New York University

PAUL DE BOER[1]
Econometric Institute, Erasmus University Rotterdam

PIET LESUIS
Department of Economics, Erasmus University Rotterdam

ABSTRACT

In this paper we present a formal operational model, based on the translog cost functional form, that allows for simultaneous determination of factor demands and of technological change. Contrary to translog specifications most commonly used to analyze total factor productivity growth, our specification allows both for smooth adjustment processes and irreversibilities in input demands that might cause input demand rigidities. Estimating the demand equations allowing for these rigidities enables us to separate measured total factor productivity growth into a cyclical and a structural component. We apply this model to 9 sectors of economic activity belonging to the manufacturing sector of the US economy. Our data span a period of 37 years, from 1958 to 1994.

[1] Corresponding author: Econometric Institute, Erasmus University Rotterdam, P.O.Box 1738, 3000 DR Rotterdam, The Netherlands.
E-mail: pmdeboer@ect.few.eur.nl.

1. Introduction

In a paper that appeared in Economic Systems Research (Lesuis et al.,1996), we presented a formal operational model based on the translog cost functional form that consists of a set of long term equilibrium relationships which describe how factor demands and technological change depend on factor prices and time. Because of informational and reactional delays in the adaptation of factor demands and technological change to price changes, we proposed to combine the model with an adjustment process according to the error correction mechanism (ECM).

According to Slade (1989) the commonly used specification of the equation that describes technological change, that we adapted in our previous paper, has two drawbacks. First, a deterministic trend is used that does not allow for the level and the slope to evolve slowly over time. Second, it does not allow for a cyclical component in measured total factor productivity (TFP) that may arise due to input demand rigidities. Instead, Slade introduces a specification in which aggregate output growth is used as a proxy for cyclical effects and TFP growth is estimated as a latent variable. The latter is done by assuming that structural TFP growth follows a stochastic trend, as introduced by Harvey (1981).

In this paper we illustrate that Slade's stochastic specification of structural TFP growth boils down to the conventional deterministic trend approach, when applied to annual data for 9 manufacturing sectors of the US economy over the period 1958-1994. We therefore consider a generalization of Slade's specification that allows for both a stochastic specification of the cyclical and the structural components in TFP growth. When applied to our data, this approach yields much better results than those obtained with Slade's specification.

The structure of the paper is as follows. In Section 2, that consists of four parts, we derive the econometric model. In the first part, we briefly review the standard cost share equations derived from the translog specification of the cost function. In the second part, we consider the equation that describes technological change. We introduce our generalization of the time varying parameter approach used by Slade and show how this specification also includes the standard model with deterministic trend. In the third part, we write the model in the state space representation that enables us to apply the Kalman Filter techniques to estimate the path of the (latent) structural and cyclical components in TFP growth. Finally, in the last part of Section 2, we briefly describe the EM-algorithm and the Kalman Filtering techniques that are used to obtain the maximum likelihood estimates of our parameters. Technical details can be found in Appendix 2. After a brief discussion of our data, we compare the results obtained using Slade's and our approach in Section 3. We show that our generalization yields a more satisfactory explanation of structural TFP growth. Section 4 contains our conclusions and suggestions for further research.

2. Model Specification

In this section we will follow Slade (1989) and consider the cost share equations and the equation for technological change implied by the following generalization of the translog unit cost function.

$$\ln C_t = F(t, y) + \sum_{i=1}^{n} a_i \ln P_{ti} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} \ln P_{ti} \ln P_{tj} + \sum_{i=1}^{n} c_i t \ln P_{ti} \tag{1}$$

where C_t denotes the unit cost at time t ($t=1\ldots T$), $F(t,y)$ represents the Hicks-neutral portion of productivity that is assumed to depend on time and on a vector of exogenous variables (y) and P_{ti} is the price of factor of production i at time t. Twice continuous differentiability of $\ln C_t$ implies that

$$b_{ij} = b_{ji} \quad \text{for} \quad i \neq j = 1\ldots n \tag{2}$$

2.1 Cost share equations

By virtue of Shephard's lemma, we can derive instantaneous conditional factor demand equations at time t in terms of cost shares, s_{ti}, by logarithmic differentiation of (1) with respect to $\ln P_{ti}$:

$$s_{ti} = a_i + c_i t + \sum_{j=1}^{n} b_{ij} \ln P_{tj} \qquad i = 1\ldots n \tag{3}$$

Since the cost shares have to add up to one by definition, the parameters have to satisfy the following additivity restrictions

$$\sum_{i=1}^{n} a_i = 1 \text{ and } \sum_{i=1}^{n} c_i = \sum_{i=1}^{n} b_{ij} = 0 \tag{4}$$

Obviously, equations (2) and (4) imply that

$$\sum_{j=1}^{n} b_{ij} = 0 \text{ for all } i$$

i.e. the cost share equations are homogeneous of degree zero in prices and can thus be written in terms of *relative prices* in the following way

$$s_{ti} = a_i + c_i t + \sum_{j=1}^{n-1} b_{ij} \ln p_{tj} \qquad i = 1\ldots n \tag{5}$$

where p_{ti} is the *relative price* of factor i in terms of factor n, i.e. $p_{ti}=P_{ti}/P_{tn}$.

2.2 Technological change

In this part we consider the equation for technological change, i.e. the percentage change of unit costs over time. Following Lesuis and de Boer (1994) we will denote this by tc_t. Partial differentiation of (1) with respect to t yields

$$tc_t = \frac{\partial \ln C_t}{\partial t} = \frac{\partial F(t,y)}{\partial t} + \sum_{i=1}^{n} c_i \ln P_{ti}$$

Using the additivity restriction (4) and defining $f(t,y) = \frac{\partial F(t,y)}{\partial t}$ this equation can be rewritten to

$$tc_t = f(t,y) + \sum_{i=1}^{n-1} c_i \ln p_{ti} \tag{6}$$

In this equation *f(t,y)* represents the Hicks neutral part of technological change and the last term represents its non-neutral part. In the following we will focus on the specification of *f(t,y)*.

When introduced by Christensen et al. (1973) the translog model was specified as an arbitrary second order approximation of the logarithm of the unit cost function. This specification implied the conventional model where

$$F(t,y) = a_0 + a_{n+1}t + \frac{1}{2}c_{n+1}t^2$$

leading to

$$f(t,y) = a_{n+1} + c_{n+1}t \tag{7}$$

Hence, according to this specification the Hicks neutral part of technological change, or *TFP growth* contains a deterministic trend. This implies that, when the length of the period considered goes to infinity, and $c_{n+1} \neq 0$, the TFP growth becomes unbounded.

Slade (1989) argues that when tc_t is measured according to the commonly used Törnqvist index, see for example Lesuis and de Boer (1994, page 362), equation (7) suffers from two problems. First, the Törnqvist index might contain a procyclical bias because in reality firms might face input demand rigidities, whereas the index is derived assuming fully flexible input demands. Secondly, it would be preferable to allow for TFP growth to slowly fluctuate over time, instead of it being completely deterministic.

In order to overcome the former, she proposes to include aggregate output growth as a proxy for input demand rigidities in the specification of *f(t,y)*. She proposes to solve the latter problem by allowing for a time varying trend, as used by for example Harvey (1981). Letting Δy_t represent the growth rate of aggregate output, the specification proposed by Slade (1989) reads

$$\begin{aligned} f(t,y) &= \delta \Delta y_t + \mu_t \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_{t1} \\ \beta_t &= \beta_{t-1} + \eta_{t2} \end{aligned} \tag{8}$$

where $\delta \Delta y_t$ represents the cyclical component and μ_t the structural component of TFP growth. β_t is the time varying trend parameter. η_{t1} and η_{t2} are assumed to be independently normally distributed white noise with variances q_1 and q_2 respectively. It can be easily seen that when there is no cyclical effect, i.e. $\delta=0$, and when the trend is not time varying, i.e. $q_1=q_2=0$, (8) reduces to the deterministic specification (7). Although this specification allows for more flexible TFP growth and for a cyclical component, it still suffers from the fact that the structural component of TFP growth is still estimated by a process with a unit root, implying unbounded growth as t goes to infinity. Furthermore, as we will show in the next section this specification will actually yield results for which $q_1=q_2=0$ and δ is insignificant. It would therefore be worthwhile to consider a specification that allows for a stationary, mean reverting, process.

Since the latent growth rate in our data seems to be stationary, as can be seen from the figures that we will present in the next section, we propose to generalize specification (8) by allowing μ_t to be an arbitrary AR(1)-process with a possible non-unit root and a time varying mean. That is, we propose to use

$$\begin{aligned} f(t,y) &= \delta\Delta y_t + \mu_t \\ \mu_t &= \rho\mu_{t-1} + \beta_{t-1} + \eta_{t1} \\ \beta_t &= \beta_{t-1} + \eta_{t2} \end{aligned} \tag{9}$$

where we consider $|\rho| \leq 1$. Thus, in principle we allow for μ_t to contain a unit root but in practice we will find that it is stationary with $|\rho|<1$.

For the relevant case that $|\rho|<1$ specification (9) implies

$$\lim_{k\to\infty} E\left[\mu_{s+k} | t=s\right] = \frac{\beta_s}{1-\rho} \tag{10}$$

Hence, at time s the expected structural growth rate depends only on β_s and ρ. Like Slade (1989) we allow structural TFP growth to evolve slowly over time and, just like in equation (8), we use a martingale specification for the parameter fluctuation. Therefore in (9) we allow β_t to follow a random walk. Thus, in specification (9), μ_t represents latent TFP growth and $\beta_t/(1-\rho)$ its structural component. However, specification (9) has one drawback. It can be easily seen that it is not identified for $\rho=0$, which actually turns out not to be a relevant case when applying the model in Section 3.

2.3 State space representation and ECM model

In this part we will rewrite the model in terms of a state space representation, see Hamilton (1994), which is used for the application of the Kalman Filter. This representation consists of two parts. The first part, known as the measurement equation, consists of the conventional share equations and the equation for technological change derived in the previous section. The second part, known as the state equation, describes the dynamic transition of the latent variables μ_t and β_t.

Measurement equation

In order to be able to write the measurement equation in matrix notation we introduce the following notation

$$w_t' = \left[s_{t1}, \ldots, s_{tn}, tc_t\right]$$

$$\xi_t' = \left[\mu_t, \beta_t, \mu_{t-1}\right]$$

$$x_t' = \left[1, t, \ln p_{t1}, \ldots, \ln p_{tn-1}, \Delta y_t\right]$$

and the parameter matrix

$$\Gamma' = \begin{bmatrix} a_1 & c_1 & b_{1,1} & \cdots & b_{1,n-1} & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_n & c_n & b_{n,1} & \cdots & b_{n,n-1} & 0 \\ 0 & 0 & c_1 & \cdots & c_{n-1} & \delta \end{bmatrix}$$

Furthermore we define

$$H_1' = \begin{bmatrix} 0_n & 0_n & 0_n \\ 1 & 0 & 0 \end{bmatrix}$$

This enables us to write the share equations (5) and the equation for technological progress, i.e. (6) in which the first equation of (9) has been substituted, as

$$w_t = \Gamma' x_t + H_1' \xi_t \tag{11}$$

Following Lesuis et al. (1996), we will consider the ECM specification of (11), which is of the form

$$w_t - w_{t-1} = \Phi\left\{\Gamma'\left(x_t - x_{t-1}\right) + H_1'\left(\xi_t - \xi_{t-1}\right)\right\} + \Psi\left\{\Gamma' x_{t-1} + H_1' \xi_{t-1} - w_{t-1}\right\} + e_t \tag{12}$$

where Φ and Ψ are $(n+1 \times n+1)$-matrices of adjustment parameters and where

$$e_t = \left(e_{t1}, \ldots, e_{tn+1}\right)' \sim IN\left(0_{n+1}, R\right)$$

It is assumed that e_t is not only independent over time but also independent from η_{t1} and η_{t2}. As explained in Lesuis et al (1996), due to the constant term appearing in (12) and the additivity restrictions, there is a perfect multicollinearity between the explanatory variables and the matrix R is singular. They show that these problems can be solved by defining

$$\tilde{\Phi} = \left[\left(\Phi_1 - \Phi_n \right) \quad \ldots \quad \left(\Phi_{n-1} - \Phi_n \right) \quad \Phi_{n+1} \right]$$
$$\tilde{\Psi} = \left[\left(\Psi_1 - \Psi_n \right) \quad \ldots \quad \left(\Psi_{n-1} - \Psi_n \right) \quad \Psi_{n+1} \right]$$

where Φ_i and Ψ_i denote the ith columns of Φ and Ψ respectively[2]. Furthermore, we use the result derived by Barten (1969) and delete an arbitrary share equation, in our case equation n. In the following the superscript *(n)* denotes that the n^{th} row, or in case of a vector the n^{th} element, has been deleted. Defining

$$H_2^{(n)'} = \begin{bmatrix} 0_{n-1} & 0_{n-1} & 0_{n-1} \\ 1 & 0 & -1 \end{bmatrix} \text{ and } H_3^{(n)'} = \begin{bmatrix} 0_{n-1} & 0_{n-1} & 0_{n-1} \\ 0 & 0 & 1 \end{bmatrix}$$

we can use

$$H_1^{(n)'} \left(\xi_t - \xi_{t-1} \right) = H_2^{(n)'} \xi_t$$

$$H_1^{(n)'} \xi_{t-1} = H_3^{(n)'} \xi_t$$

to rewrite (12) as

$$w_t^{(n)} - w_{t-1}^{(n)} = \tilde{\Phi}^{(n)} \left\{ \Gamma^{(n)'} \left(x_t - x_{t-1} \right) + H_2^{(n)'} \xi_t \right\} + \tilde{\Psi}^{(n)} \left\{ \Gamma^{(n)'} x_{t-1} + H_3^{(n)'} \xi_t - w_{t-1}^{(n)} \right\} + e_t^{(n)}$$

In order to introduce the state space representation of the ECM model in matrix notation, we define the following matrices

$$A' = \left[\tilde{\Phi}^{(n)} \Gamma^{(n)'} \quad | \quad \left(\tilde{\Psi}^{(n)} - \tilde{\Phi}^{(n)} \right) \Gamma^{(n)'} \quad | \quad - \tilde{\Psi}^{(n)} \right]$$

$$H' = \left[\tilde{\Phi}^{(n)} H_2^{(n)'} + \tilde{\Psi}^{(n)} H_3^{(n)'} \right] \text{ and } z_t' = \left[x_t' \quad | \quad x_{t-1}' \quad | \quad w_{t-1}^{(n)'} \right]$$

such that we can write

$$\Delta w_t^{(n)} = A' z_t + H' \xi_t + e_t^{(n)} \tag{13}$$

In our empirical application we resort the a simplification: we impose the restriction that the matrices of adjustment parameters are diagonal. As shown by Lesuis et al. (1996) the adjustment matrices then reduce to

$$\tilde{\Phi}^{(n)} = \begin{bmatrix} \phi_1 I_{n-1} & 0_{n-1} \\ 0_{n-1}' & \phi_{n+1} \end{bmatrix} \text{ and } \tilde{\Psi}^{(n)} = \begin{bmatrix} \psi_1 I_{n-1} & 0_{n-1} \\ 0_{n-1}' & \psi_{n+1} \end{bmatrix}$$

Moreover, because we already impose a dynamic structure on the equation describing structural technological change, i.e. (9), we fix ϕ_{n+1} and ψ_{n+1} at 1, implying a static specification.

[2] It can be shown, see Lesuis et al. (1996) that the first n elements of the columns of $\tilde{\Phi}$ and $\tilde{\Psi}$ add up to zero.

State equation

We define

$$\eta_t' = [\eta_{t1}, \eta_{t2}, 0] \text{ and } Q = \begin{bmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and the transition matrix

$$F = \begin{bmatrix} \rho & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

so that we can rewrite the last two equations of (9) as

$$\xi_t = F\xi_{t-1} + \eta_t \tag{14}$$

which is the state equation that combined with the measurement equation derived in (13) yields the state space representation used in Hamilton (1994) to derive the Kalman Filter.

2.4 Estimation

In appendix 2 we describe the Kalman filter and the maximum likelihood procedures that were applied to obtain the results presented in the following section. The appendix consists of three parts. In the first part we review the Kalman filter equations that can be used to obtain forecasts of the state variable, ξ_t, conditional on previous and current information. Furthermore, we also review the Kalman smoother that calculates forecasts of the state variable, ξ_t, conditional on all the data available. In the second part, we introduce the likelihood function and describe how the Kalman filter and smoother can be used in conjunction with the EM algorithm, first introduced by Dempster et al. (1977) and applied to state space models by Shumway and Stoffer (1982), to obtain maximum likelihood estimates of the unknown parameters in the model. The EM algorithm iterates over two steps; the expectation step (*E*-step) and the maximization step (*M*-step). Shumway and Stoffer (1982) show that for the state space model considered here in the *E*-step the missing data, in our case the latent variable ξ_t, are replaced by their sufficient statistics, which are in our case obtained from the Kalman smoother. In the second step new parameter estimates are then obtained by maximizing the conditional expection of the likelihood for the model that is derived assuming that all variables, including ξ_t are observed. This expectation is evaluated conditional on the available data. The *M*-step thus yields updated parameter estimates that can again be used in the Kalman smoother in the *E*-step. Finally, in the third part, we consider the practical implementation of the maximization step (*M*-step) of the EM algorithm.

3. Data and empirical results

In this section we present our empirical results for the US economy using data from the NBER Manufacturing Productivity Database for the period 1958-1994. A more detailed description of the data can be found in Bartelsman and Gray (1996). In this paper we use three categories of production factors, i.e. capital, labor and materials. Our TFP growth measure , based on the three input case, perfectly matches the

Table 1. Estimation results for turbine sector (3511)

	Restricted		*Unrestricted*	
a_1	.289	(.020)	.288	(.020)
a_2	.245	(.013)	.244	(.014)
c_1	.005	(.002)	.005	(.002)
c_2	-.005	(.002)	-.005	(.002)
b_{11}	.131	(.070)	.133	(.074)
b_{12}	-.069	(.038)	-.063	(.035)
b_{22}	.078	(.048)	.082	(.050)
δ	.289	(.175)	.357	(.321)
ϕ_1	1.203	(.604)	1.215	(.626)
ψ_1	.144	(.071)	.143	(.072)
ρ	1.000		-.098	
q_1^*	.000		.327	
q_2^*	.000		.370	

* q_1 and q_2 have to multiplied by 10^{-2}.
standard errors between parentheses

NBER measure that is calculated for the five input case. Capital income is treated as a value added residual. A rental price of capital was obtained implicitly by dividing capital income by the real capital stock figures available in the database. Likewise a wage deflator was obtained implicitly from the total payroll and the number of employees.

The discussion of our detailed results in this section will be limited to the Turbine and Turbine Generator Sets industry (SIC 3511). The complete estimation results for all sectors are presented in Appendix 1, Table 2 and Table 3.

Table 1 lists the parameter estimates for both Slade’s (1989) specification, denoted by “Restricted”, and our generalization, denoted by “Unrestricted”. Two things are immediately apparent. First, the parameter estimates for the state equations do not differ much for both specifications. Furthermore, the coefficient for the proxy for irreversibilities, i.e. δ, turns out to be insignificant, as judged by the "two-sigma" rule of thumb . As listed in Appendix 1, the results concerning the parameter estimates for the state equations seem to be robust across the sectors considered. Significant cyclical effects however are present in the sectors Furniture (SIC 2511), Newspapers (SIC 2711), Steel mills (SIC 3312) and Car bodies (SIC 3711) for both type of specifications.

The peculiarity of the results obtained with Slade’s specification are most clear when considering estimated structural and measured TFP growth. From Table 1 it can be seen that, in case of the restricted specification, q_1 and q_2 are both estimated equal to zero. This implies that Slade’s (1989) specification, for our data, yields the same result as the conventional Translog model with a deterministic trend, i.e. as specification (7). That is, it fits structural TFP growth as a deterministic trend. That this specification leads to a poor fit is immediately clear when one considers figure 1, which depicts measured and estimated structural TFP growth. Where measured TPF growth, i.e. the Törnqvist index, seems to be a mean reverting process, estimated structural TFP growth is a slightly positive deterministic trend.

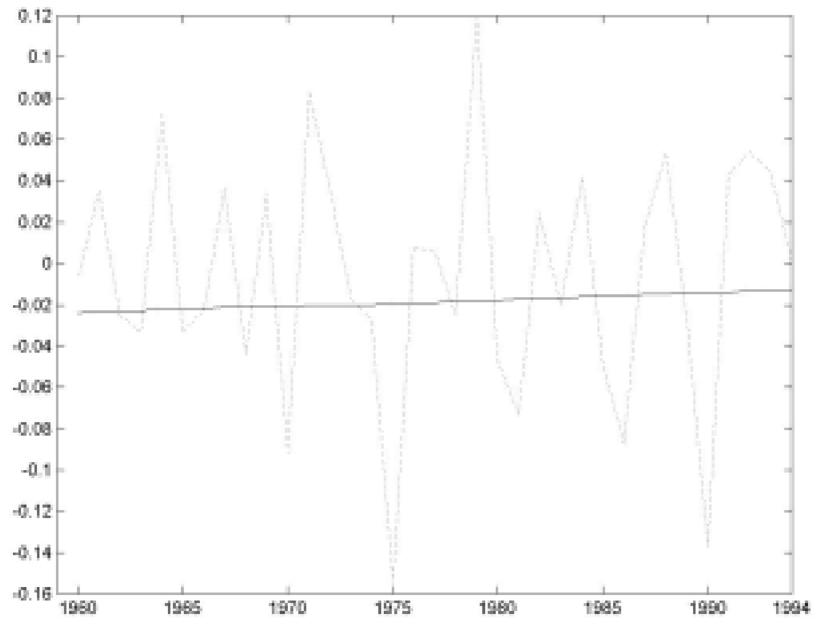


figure 1. Estimated and measured TFP growth, restricted specification
solid: estimated structural TFP growth (μ_t), *dashed*: Törnqvist index.

This fit contrasts sharply with the fit obtained with our specification. From figure 2 it can be easily seen that our model yields a much better fit for TFP-growth. More importantly, instead of being a positive deterministic trend as obtained with Slade's (1989) specification estimated structural TFP growth, as depicted in figure 3, seems to be fluctuating much more. As expected, the outliers for the oilcrises in 1974 and subsequently 1979, are mainly attributed to cyclical effects. More importantly, where the results obtained with the restricted specification suggest that the growth rate of TFP will eventually be infinite, because of the deterministic trend, the results of the unrestricted specification suggest that structural TFP growth peaked especially after the oilcrisis in 1975 and in 1979.

.

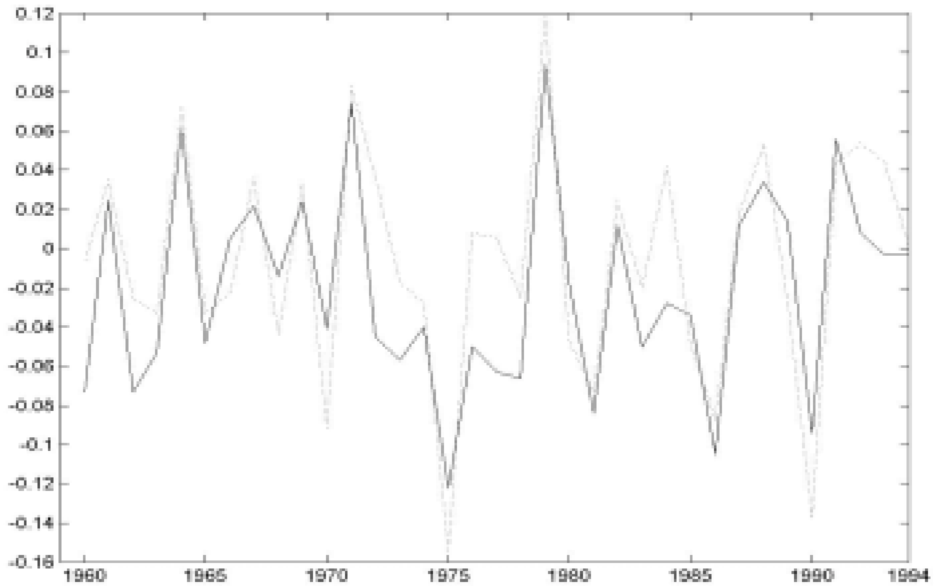


figure 2. Estimated and measured TFP growth, unrestricted specification
solid: estimated latent Hicks-neutral TFP growth (μ_t), *dashed*: Törnqvist index.

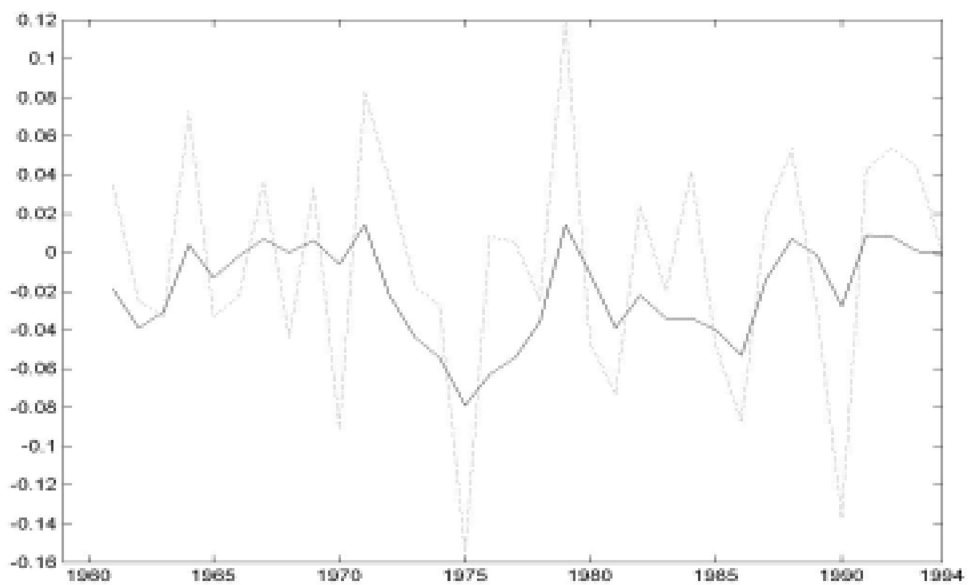


figure 3. Estimated and measured TFP growth, unrestricted specification
solid: Estimated structural TFP growth ($\beta_{t-1}/(1-\rho)$), *dashed*: Törnqvist index.

5. Concluding remarks

In this paper we have extended Slade's (1989) idea to allow for the estimation of structural TFP-growth as a latent variable. We have shown that, using data for 9 sectors for the US economy, Slade's specification boils down to the conventional Translog model in which structural TFP growth is modeled as a deterministic trend. Our specification allows for both structural as well as cyclical TFP growth being latent variables.

Our results suggest that structural TFP growth does not follow any deterministic pattern at all but, instead, fluctuates significantly over time. Furthermore, contrary to Slade's specification, which predicts that the structural component of the productivity growth rate will grow beyond bounds, the results obtained with our generalized specification do not seem to suggest any trend in the growth rate of TFP.

Many extensions of the analysis that we presented in this paper are possible. Two of them are especially worthwhile mentioning. First, one could consider a specification of the state equation that does not suffer from an identification problem for certain parameter values, as is the case in our specification when $\rho=0$. Second, one could substitute the deterministic trend in the share equations by the latent TFP growth variable. One thing should be clear from the results presented in this paper: Any specification of TFP growth as a latent variable should allow for it to follow a general stochastic process that can preferably be divided in a structural and cyclical component.

References

Bartelsman E.J. & Gray W., (1996) The NBER Manufacturing Database, *NBER Technical Working Paper Series 205*.

Barten, A.P. (1969) Maximum likelihood estimation of a complete system of demand equations, *European Economic Review*, 1, pp. 7-73.

Christensen, L.R., Jorgenson, D.W. & Lau, L.J. (1973) Transcendental logarithmic production frontiers, *Review of Economics and Statistics*, 55, pp. 28-45.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society*, *Series B*, 39, pp. 1-38.

Hamilton, J.D. (1994) *Time Series Analysis*, Princeton University Press, Princeton.

Harvey, A.C. (1981) *Time Series Models*, Philip Allen Publishers, Oxford.

Lesuis, P.J.J. & de Boer, P.M.C. (1994) Interrelated factor demands and technological change: an empirical assessment of adjustment for sectors of the Dutch economy, *Economic Systems Research*, 6, pp. 347-362.

Lesuis, P.J.J., de Boer, P.M.C., Harkema, R. & Hobijn, B. (1996) Structural models of factor demands and technological change: an empirical assessment of dynamic adjustment specifications for sectors of the Dutch economy, *Economic Systems Research*, 8, pp. 341-360.

Schumway, R.H. & Stoffer, D.S. (1982) An approach to time series smoothing and forecasting using the EM algorithm, *Journal of Time Series Analysis*, 3, pp. 253-264.

Slade, M.E. (1989) Modeling stochastic and cyclical components of technical change: an application of the Kalman filter, *Journal of Econometrics*, 41, pp.363-383.

Watson, M.W. & Engle, R.F. (1983) Alternative algorithms for the estimation of dynamic factor, MIMIC and varying coefficient regression models, *Journal of Econometrics*, 23, pp. 385-400.

Appendix 1: Complete estimation results

Table 2. Estimation results: restricted model

	Measurement equation										*state equation*		
Sector	a_1	a_2	c_1	c_2	b_{11}	b_{12}	b_{22}	δ	ϕ_1	ψ_1	ρ#	q_1^*	q_2^*
2011 Meatpacking	.071	.059	.0005	-.0011	.031	-.002	.023	.215	2.014	.158	1.	.000	.000
	.003	.009	.0003	.0003	.019	.002	.014	.081	1.190	.061			
2511 Furniture	.270	.259	.0024	-.0026	.035	-.028	.055	.596	2.007	.324	1.	.000	.000
	.013	.012	.0004	.0003	.025	.021	.040	.091	1.350	.104			
2711 Newspapers	.478	.300	.0079	-.0041	.154	-.077	.105	.312	1.012	.098	1.	.000	.000
	.054	.029	.0018	.0009	.077	.039	.054	.064	.491	.044			
2911 Petroleum re	.106	.025	.0018	-.0005	.078	-.003	.019	.313	1.214	.091	1.	.000	.000
	.018	.005	.0016	.0005	.031	.002	.008	.138	.461	.046			
3111 Leather	.181	.176	.0004	-.0039	.036	-.017	.035	-.240	2.256	.151	1.	.000	.000
	.016	.097	..0018	.0011	.057	.028	.006	.215	3.511	.081			
3312 Steel mills	.206	.184	.0018	-.0034	.048	-.029	.050	.567	1.770	.202	1.	.000	.000
	.001	.015	.0008	.0005	.022	.013	.024	.120	.768	.065			
3511 Turbines	.289	.245	.0053	-.0047	.131	-.069	.078	.289	1.203	.144	1.	.000	.000
	.020	.013	.0023	.0018	.070	.038	.048	.175	.604	.071			
3711 Car bodies	.190	.080	-.0011	-.0021	.057	-.007	-.018	.679	.996	.337	1.	.000	.000
	.008	.005	.0009	.0007	.013	.006	.025	.100	.187	.090			
3911 Precious met.	.255	.187	.0024	-.0034	.074	-.054	.043	.343	1.623	.164	1.	.000	.000
	.020	.022	.0011	.0008	.040	.028	.022	.316	.800	.067			

* q_1 and q_2 have to be multiplied by 10^{-2}

#preset at value 1

Table 3. Estimation results: unrestricted model

		measurement equation										state equation		
Sector		a_1	a_2	c_1	c_2	b_{11}	b_{12}	b_{22}	δ	ϕ_1	ψ_1	ρ	q_1^*	q_2^*
2011	Meatpacking	.070	.057	.005	-.0011	.036	-.002	.027	.239	1.745	.147	-.060	.088	.081
		.004	.008	.0004	.0003	.021	.002	.016	.162	.989	.063			
2511	Furniture	.268	.262	.0023	-.0025	.033	-.023	.049	.604	2.265	.312	-.017	.140	.224
		.016	.015	.0004	.0003	.027	.019	.041	.233	1.730	.103			
2711	Newspapers	.478	.300	.0076	-.0040	.151	-.078	.102	.332	1.050	.108	-.239	.050	.103
		.046	.025	.0016	.0008	.070	.036	.048	.145	.468	.045			
2911	Petroleum ref	.113	.030	.0011	-.0007	.051	-.002	.014	.255	1.862	.121	-.205	.220	.312
		.016	.006	.0012	.0004	.032	.001	.009	.277	1.129	.057			
3111	Leather	.187	.149	.0015	-.0036	.054	-.023	.067	-.210	1.487	.298	-.100	.603	.578
		.011	.012	.0009	.0007	.028	.014	.038	.416	.711	.092			
3312	Steel mills	.206	.183	.0018	-.0035	.052	-.029	.051	.620	1.733	.207	-.114	.191	.239
		.009	.014	.0009	.0005	.022	.012	.023	.253	.698	.068			
3511	Turbines	.288	.244	.0051	-.0049	.133	-.063	.082	.357	1.215	.143	-.098	.327	.370
		.020	.014	.0024	.0017	.731	.035	.050	.321	.626	.072			
3711	Car bodies	.187	.082	-.0010	-.0022	.048	-.008	.008	.732	1.322	.332	-.064	.093	.097
		.009	.006	.0009	.0006	.015	.005	.020	.166	.366	.095			
3911	Precious met	.261	.181	.0018	-.0032	.075	-.053	.052	.444	1.571	.225	-.163	.983	1.236
		.011	.012	.0008	.0007	.031	.021	.020	.612	.562	.076			

* q_1 and q_2 have to be multiplied by 10^{-2}

Appendix 2: Maximum likelihood estimation, Kalman filter and the EM algorithm

In this appendix we briefly describe the Kalman filter and the maximum likelihood procedures that were applied to obtain the results reported in the paper and Appendix 1.

The Kalman filter and smoother

Let $\xi_{t|s}$ denote the minimum mean squared error forecast of the state variable at time t on the basis of the information available at time s and let $P_{t|s}$ be the corresponding covariance matrix of the forecast errors, then the Kalman filter generates series $\xi_{t|t-1}$, $\xi_{t|t}$, $P_{t|t}$ and $P_{t|t-1}$, conditional on the values of $\xi_{0|0}$ and $P_{0|0}$. These forecasts can be shown to be the result of a series of recursive linear projections. That is, given $\xi_{0|0}$ and $P_{0|0}$, the Kalman filter first calculates $\xi_{t|t-1}$ and $P_{t|t-1}$ on the basis of $\xi_{t-1|t-1}$ and $P_{t-1|t-1}$ and then $\xi_{t|t}$ and $P_{t|t}$ on the basis of $\xi_{t|t-1}$ and $P_{t|t-1}$ for $t=1\ldots T$. This recursion is captured by the following recursive formulas, which are derived under the assumption that both the parameter matrices A, H and F as well as the covariance matrices R and Q are known.

$$\begin{aligned}
\xi_{t+1|t} &= F\xi_{t|t} \\
\xi_{t|t} &= \xi_{t|t-1} + P_{t|t-1}H\left(H'P_{t|t-1}H + R\right)^{-1}\left(\Delta w_t^{(n)} - A'z_t - H'\xi_{t|t-1}\right) \\
P_{t+1|t} &= FP_{t|t}F' + Q \\
P_{t|t} &= P_{t|t-1} - P_{t|t-1}H\left(H'P_{t|t-1}H + R\right)^{-1}H'P_{t|t-1}
\end{aligned}$$

However, we are more interested in the forecast of the structural TFP-growth on the basis of our whole sample of data, i.e. in $\xi_{t|T}$ and $P_{t|T}$. Using the results of the Kalman filter, these forecasts can be obtained using the following recursion, known as the Kalman smoothing procedure.

$$\begin{aligned}
\xi_{t|T} &= \xi_{t|t} + J_t\left(\xi_{t+1|T} - \xi_{t+1|t}\right) \\
P_{t|T} &= P_{t|t} + J_t\left(P_{t+1|T} - P_{t+1|t}\right)J_t'
\end{aligned}$$

where $J_t = P_{t|t}F'P_{t+1|t}^{-1}$ and $t=T\text{-}1\ldots 0$. This backward recursion, also yields a new forecast for $\xi_{0|0}$ and $P_{0|0}$. It turns out that for the implementation of the EM algorithm, later described in this section, we also need to calculate $P_{t,t-1|T}=E[(\xi_{t|T}-\xi_t)\ (\xi_{t-1|T}-\xi_{t-1})'|\Omega_T]$, where Ω_t denotes the information available at time t. In order to obtain $P_{t,t-1|T}$ for $t=T\text{-}1,\ldots,1$ we follow Schumway and Stoffer (1982) and use the following backward recursion

$$\begin{aligned}
P_{T,T-1|T} &= \left(I - P_{T|T-1}H\left(H'P_{T|T-1}H + R\right)^{-1}H'\right)FP_{T-1|T-1} \\
P_{t,t-1|T} &= P_{t|t}J_{t-1}' + J_t\left[P_{t+1,t|T} - FP_{t|t}\right]J_{t-1}'
\end{aligned}$$

Maximum likelihood estimation and the EM algorithm

For given priors on the state variable, i.e. $\xi_{0|0}$ and $P_{0|0}$, we can now calculate the maximum likelihood estimates of the unknown parameters. In order to derive the likelihood function it is important to realize that the distributional assumptions about the residual vectors e_t and η_t imply that

$$\Delta w_t^{(n)} | z_t, \Omega_{t-1} \sim N\left(\left(A' z_t + H' \xi_{t|t-1}\right), \left(H' P_{t|t-1} H + R\right)\right)$$

That is, the likelihood value associated with the t^{th} observation equals

$$f\left(\Delta w_t^{(n)} | z_t, \Omega_{t-1}\right) = (2\pi)^{-n/2} \left| H' P_{t|t-1} H + R \right|^{-1/2} \times \exp\left\{ -\frac{1}{2}\left(\Delta w_t^{(n)} - A' z_t - H' \xi_{t|t-1}\right)' \left(H' P_{t|t-1} H + R\right)^{-1} \left(\Delta w_t^{(n)} - A' z_t - H' \xi_{t|t-1}\right)\right\}$$

Thus, the log-likelihood function, conditional on the priors $\xi_{0|0}$ and $P_{0|0}$, equals

$$LnL = \sum_{t=2}^{T} \ln f\left(\Delta w_t^{(n)} | z_t, \Omega_{t-1}\right) \tag{15}$$

In principle the likelihood function in (15) can be maximized using standard numerical methods, like the score algorithm and Newton Raphson. However, as argued by Shumway and Stoffer (1982) and Engle and Watson (1983), these methods require a lot of calculational effort and do not assure us of an increase in the log-likelihood value in every iteration. We therefore follow Shumway and Stoffer (1982) and Engle and Watson (1983) and use the EM algorithm, introduced by Dempster et. Al (1977) to maximize the log-likelihood function (15). The EM algorithm iterates over two steps; the expectation step (*E*-step) and the maximization step (*M*-step). Shumway and Stoffer (1982) show that for the state space model considered here in the *E*-step the missing data, in our case the latent variable ξ_t, are replaced by their sufficient statistics, which are in our case obtained from the Kalman smoother. In the second step new parameter estimates are then obtained by maximizing conditional expection of the likelihood for the model that is derived assuming that all variables, including ξ_t are observed. This expectation is evaluated conditional on the available data. The *M*-step thus yields updated parameter estimates that can again be used in the Kalman filter and in the *E*-step to obtain new values for $\xi_{t|T}$. In the following section we will describe the implementation of the *M*-step in more detail.

The M-step

In this section we will first consider the likelihood function for our model under the assumption that ξ_t was actually observed. We will then derive an expression for the expectation of the likelihood function conditional on the available data and briefly describe an iterative method that can be used to maximize this expectation. For notational purposes we introduce

$$\begin{aligned} W &= \left[w_2^{(n)}, \ldots, w_T^{(n)}\right] \\ W_{-1} &= \left[w_1^{(n)}, \ldots, w_{T-1}^{(n)}\right] \\ Z &= \left[\xi_2, \ldots, \xi_T\right] \\ Z_{-1} &= \left[\xi_1, \ldots, \xi_{T-1}\right] \end{aligned} \qquad \begin{aligned} X &= \left[x_2, \ldots, x_T\right] \\ X_{-1} &= \left[x_1, \ldots, x_{T-1}\right] \\ E &= \left[e_2^{(n)}, \ldots, e_T^{(n)}\right] \end{aligned} \quad \text{and} \quad H_4' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Furthermore, let $\Delta W = W - W_{-1}$, $\Delta X = X - X_{-1}$ and $\Delta Z = Z - Z_{-1}$. Under the assumption that we would have observed all relevant variables, including ξ_t, and that $H'_4 \xi_1$ is normally distributed, such that $H_4' \xi_1 \sim N\left(H_4' \mu_\xi, H_4' P_1 H_4\right)$, we can write the log likelihood function as

$$
\begin{aligned}
LnL \sim & -\frac{1}{2}\ln\left|P_1\right| - \frac{1}{2}\operatorname{tr}\left(P_1^{-1}H_4'\left(\xi_1-\mu_\xi\right)\left(\xi_1-\mu_\xi\right)' H_4 \right) \\
& -\frac{T-1}{2}\ln\left|Q\right| - \frac{1}{2}\operatorname{tr}\left(Q^{-1}H_4'\left(Z-FZ_{-1}\right)\left(Z-FZ_{-1}\right)' H_4 \right) \\
& -\frac{T-1}{2}\ln\left|R_n\right| - \frac{1}{2}\operatorname{tr}\left(\left(R^{(n)}\right)^{-1} EE' \right)
\end{aligned} \tag{16}
$$

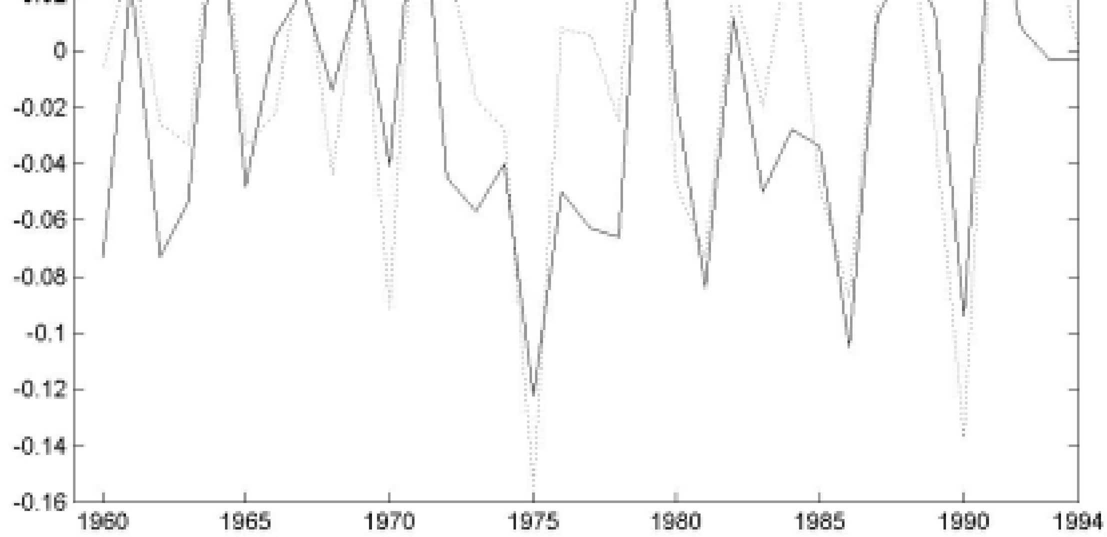
where the first line is the part of the likelihood determined by the initial value ξ_1, the second line is the part associated with the state equation and the third line with the measurement equation. In principle, one would be tempted to maximize (16). However, (16) does not take into account that the latent variable, i.e. *Z*, is unobserved. The objective function that is maximized in the M-step of the EM-algorithm will instead be the expectation of (16) conditional on the available data. Although tedious, this objective function can be derived as a simple extension of the result in Shumway and Stoffer (1982). This derivation yields

$$
\begin{aligned}
LnL = & \, E\left[LnL|data\right] \sim \\
& -\frac{1}{2}\ln\left|\hat{P}_1\right| - \frac{1}{2}\operatorname{tr}\left(\hat{P}_1^{-1}H_4'\left\{P_{1|T} + \left(\hat{\hat{\xi}}_1-\mu_\xi\right)\left(\hat{\hat{\xi}}_1-\mu_\xi\right)'\right\} H_4 \right) \\
& -\frac{T-1}{2}\ln\left|\hat{Q}\right| - \frac{1}{2}\operatorname{tr}\left(\hat{Q}^{-1}H_4'\left\{C - B\hat{F}' - \hat{F}B' + \hat{F}A\hat{F}'\right\} H_4 \right) \\
& -\frac{T-1}{2}\ln\left|\hat{R}^{(n)}\right| - \frac{1}{2}\operatorname{tr}\left(\left(\hat{R}^{(n)}\right)^{-1} \left\{ \hat{E}\hat{E}' + \left(\hat{\hat{\Phi}}^{(n)}H_2' + \hat{\hat{\Psi}}^{(n)}H_2'\right) D \left(\hat{\hat{\Phi}}^{(n)}H_2' + \hat{\hat{\Psi}}^{(n)}H_2'\right)' \right\} \right)
\end{aligned} \tag{17}
$$

where ^ is used to emphasize that we maximize (17) with respect to the estimated parameter values. Furthermore, the matrices *A*, *B*, *C* and *D* are defined as

$$
\begin{aligned}
A &= \left[\left\{\sum_{t=2}^{T} P_{t-1|T}\right\} + \hat{Z}_{-1}\hat{Z}_{-1}'\right] \\
B &= \left[\left\{\sum_{t=2}^{T} P_{t,t-1|T}\right\} + \hat{Z}\hat{Z}_{-1}'\right] \\
C &= \left[\left\{\sum_{t=2}^{T} P_{t|T}\right\} + \hat{Z}\hat{Z}'\right] \\
D &= \left\{\sum_{t=2}^{T} P_{t|T}\right\}
\end{aligned}
$$

where $\hat{Z} = \left[\xi_{2|T}, \ldots, \xi_{T|T}\right]$ and $\hat{Z}_{-1}$ is defined correspondingly. The maximization procedure that we used to maximize (17) is a sequential procedure, that, in every iteration, uses sequentially maximizes (17) with respect to one of the matrices to be estimated, i.e. $\hat{P}_1, \hat{\hat{\xi}}_1, \hat{Q}, \hat{F}, \hat{R}^{(n)}, \hat{\Gamma}^{(n)}, \hat{\hat{\Phi}}^{(n)}$ and $\hat{\hat{\Psi}}^{(n)}$, conditional on all other parameter estimates. The derivation of the exact steps is lengthy and tedious and can be obtained, upon request, from the authors. The resulting maximization procedure is very similar to the one used in Lesuis et. al. (1996).


0
-0.02
-0.04
-0.06
-0.08
-0.1
-0.12
-0.14
-0.16
1960
1965
1970
1975
1980
1985
1990
1994