Identifying Regression Parameters When Variables are Measured with Error

A. N. Rambaldi[1]*, T. H. Y. Tran[2]† and A. Peyrache[1]

[1]School of Economics, The University of Queensland, St Lucia, QLD 4072. Australia

[2]Department of Economics, Yale University

this version: 14 December, 2015

Abstract

The paper proposes an approach for identifying and estimating the economic parameters of interest when all the variables are measured with errors and these are correlated. Two propositions show how the parameters of interest and the bias are identified. Three Monte Carlo simulations illustrate the results. The empirical application estimates returns to scale and technological progress in US manufacturing sectors. The results can be linked to previous works in the literature to demonstrate the ambiguous bias in least squares estimates of returns to scale parameters and to compare estimates of trends in technological change using two alternative identification approaches.

*Corresponding author. Tel.: +61 7 33656576, Email: a.rambaldi@uq.edu.au

†The work was carried out while a student at The University of Queensland

Keywords: unobserved components, time-varying parameters, least squares bias, returns to scale, technological change

JEL Code: *C18, C32, E23*

1 Introduction

This paper proposes an approach to identifying and estimating the parameters of interest in an economic relationship characterized by variables measured with errors. A main departure from one of the basic assumptions of the classical errors in variables model (EIV) is to allow for correlation across the measurement errors in the variables of the model. The bias of the least squares estimator is no longer downwards, as the sign is ambiguous. The essence of the proposal is to treat the observed variables as jointly distributed in an unobserved components structural time series framework. Some of the advantages are that time-varying relationships can be estimated using standard maximum likelihood and Kalman filtering techniques and additional features of the data being studied (such as trends, cycles, and seasonality) are easily incorporated. The paper provides two propositions to show how the parameters of interest and the bias can be identified when there is non-zero correlation in the measurement errors across the variables in the model.

Measurement errors are well documented in many settings in economics. Devereux (2001) raises the problem in survey data of earnings and work hours. Klette and Griliches (1996), Ornaghi (2006) and Van Beveren (2012) show that the use of industry-level prices to deflate firm-level sales in order to obtain a proxy for firm's output quantity instead of the real quantity, likely results in biased input coefficients and unreasonably low estimates of returns to scale. Diewert and Fox (2008) derive an economic model that has as parameters returns to scale and technological change and show how the EIV is the relevant econometric model to estimate the economic relationship.

In the single regressor static EIV the assumption is that the regressor is measured with error and the measurement error is uncorrelated with the error in the regression equation. In this case, under the assumption of normality, the distribution of the dependent and regressor variables are completely characterized by the first and second moments. Early results on identification were established by Reisersol (1950), Kendall and Stuart (1977) and Kapteyn and Wansbeek (1983). A review of early works can be found in Fuller (1987). Identification is possible when some prior information on the measurement error is available (usually it is either assumed that the variance of the measurement error is known or the ratio of this variance to the variance of the regression is known). Since in the EIV model OLS is downward biased, this prior information is used to correct for the bias of the OLS estimator. In practice, neither of these estimators are readily available, and they are often estimated through replication (Buonaccorsi, 2010).

Alternative strategies include the use of instrumental variables, abandoning the one equation model and specifying a more complete simultaneous equation model which identifies truly exogenous variables (measured without error) from endogenous variables and variables measured with error (e.g. Marschak and Andrews (1944), Anderson (1976), and Klette and Griliches (1996)), and making use of the dynamic structure of the data. Maravall and Aigner (1977) pointed out that a static model that is unidentified in the presence of serially uncorrelated measurement errors could be identifiable when the model has a dynamic structure. Intuitively, if the measurement errors are white noise but the model has dynamics, the model can be identified by the suitable choice of instruments, such as lags. Anderson and Deistler (1984) assume that the regressor and the measurement errors are autoregressive moving average processes (ARMA). Nowak (1993) extends this to a multivariate setting where the underlying series follows a VARMA process. Solo (1986) and Komumjer and Ng (2014) consider an ARMA model with exogenous variables (ARMAX) where measurement errors are serially autocorrelated. Sargent (1989) uses an autoregressive state-space system and shows how the Kalman filter can be used when the variables in a dynamic model are subject to autoregressive measurement errors. Harvey (2013) remarks that the identification problem can be resolved if the dynamic properties of the components of interest are different from those of the errors. In Harvey (2011) the Phillips curve is built from the joint cyclical components of the unemployment rate and GDP, after trend, seasonal, and irregular components are extracted. The extraction is possible because their dynamics differ from those of the cycles. The approach in this paper follows these ideas.

The paper is organized as follows. Section 2 provides two propositions to show how the parameters of interest and the bias can be identified when there is non-zero correlation in the measurement errors across the variables in the model. These provide the necessary results to a flexible modeling and estimation approach easy to implement using existing commercially available packages. Estimation is discussed in Section 3. Section 4 presents a simulation exercise designed to study the propositions. The empirical example in Section 5 considers the EIV model proposed by Diewert and Fox (2008) to estimate returns to scale and technological change in US industrial sectors. The example is chosen as the estimates of returns to scale can be compared to those obtained by Diewert and Fox (2008) using OLS, and thus assess the extent of the bias induced by measurement errors. Extending their model, technical change is allowed to be time-varying as in the model by Jin and Jorgenson (2010). Jin and Jorgenson (2010) use a very different economic modeling strategy; however, the objective is to estimate technical change of US industrial sectors by treating it as a latent variable. They use an instrumental variable and Kalman filtering approach to account for the correlation between the regressors and the random error. Section 6 concludes.

2 Model and Identification

The relationship of interest is between variables y_t^* and x_t^*, where x_t^* could be a vector of k variables; however, without loss of generality the presentation will be made with a bivariate system. The econometrician observes y_t^* and x_t^* contaminated with measurement errors given by $v_t = \begin{pmatrix} v_{yt} & v_{xt} \end{pmatrix}'$,

$$
\begin{aligned}
y_t &= y_t^* + v_{yt} && (1)\\
x_t &= x_t^* + v_{xt} && (2)
\end{aligned}
$$

The measurement errors are assumed to follow a multivariate normal distribution with zero mean and a constant covariance matrix

$$\Sigma_v = \begin{bmatrix} \sigma_{vy}^2 & \rho_v \sigma_{vy} \sigma_{vx} \\ \rho_v \sigma_{vy} \sigma_{vx} & \sigma_{vx}^2 \end{bmatrix}$$

and the relationship of interest is given by

$$y_t^* = \alpha_t + \beta x_t^* \tag{3}$$

This type of model can describe classical relationships such as an aggregate Cobb-Douglas production function under constant returns to scale where y_t^* is log of output per unit of labour, x_t^* is log of capital per unit of labour and α_t is log of total factor productivity.

Assume that α_t follows a random walk process (other processes are possible, for example trend with drift as in the stochastic growth model)

$$\alpha_t = \alpha_{t-1} + \epsilon_{\alpha t} \tag{4}$$

where $\epsilon_{\alpha t}$ is normally distributed with mean zero and variance σ_α^2. We also accommodate trends in the unobservable x_t^* assuming that it follows a random walk:

$$x_t^* = x_{t-1}^* + \eta_{xt} \tag{5}$$

where η_{xt} is normally distributed with mean zero and variance $\sigma_{\eta x}^2$. Using the previous assumptions, we can write:

$$y_t^* = \alpha_{t-1} + \beta x_{t-1}^* + \epsilon_{\alpha t} + \beta \eta_{xt} = y_{t-1}^* + \eta_{yt} \tag{6}$$

where $\eta_{yt} = \epsilon_{\alpha t} + \beta \eta_{xt}$. Thus, $\sigma_{\eta y}^2 = \sigma_{\alpha}^2 + \beta^2 \sigma_{\eta x}^2$ and $cov(\eta_{yt}, \eta_{xt}) = \beta \sigma_{\eta x}^2$. It is thus possible to write the model in a compact form as an unobserved components model.Write the structural model (1)-(2) in vector form,

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} y_t^* \\ x_t^* \end{pmatrix} + \begin{pmatrix} \upsilon_{yt} \\ \upsilon_{xt} \end{pmatrix} \tag{7}$$

The equations in (7) decompose the vector of observed variables into the true series and the measurement errors. The true series are assumed, for the purpose of the initial discussion, to follow linear trends; however, the results can be generalized as shown in Section 2.2,

$$\begin{pmatrix} y_t^* \\ x_t^* \end{pmatrix} = \begin{pmatrix} y_{t-1}^* \\ x_{t-1}^* \end{pmatrix} + \begin{pmatrix} \eta_{yt} \\ \eta_{xt} \end{pmatrix} \tag{8}$$

The signal vector $\eta_t = \begin{pmatrix} \eta_{yt} & \eta_{xt} \end{pmatrix}'$ is assumed to follow a multivariate normal distribution with zero mean and covariance matrix.

$$\Sigma_\eta = \begin{bmatrix} \sigma_{\eta y}^2 & \rho_\eta \sigma_{\eta y} \sigma_{\eta x} \\ \rho_\eta \sigma_{\eta y} \sigma_{\eta x} & \sigma_{\eta x}^2 \end{bmatrix}$$

The measurement error series, υ_t and the signal series, $\eta_t = (\eta_{yt}, \eta_{xt})'$, are assumed serially uncorrelated

Equations (7) and (8) provide an unobserved components representation (UC) of the structural model, (1)-(2). The next subsection provides two propositions to show that the structural parameters in (3) can be identified in terms of the parameters in the UC representation, and demonstrate how measurement errors and their correlation contribute to the bias in the OLS estimator, which in turn helps clarify how the estimation of the UC form can avoid the measurement error bias and provide consistent estimators of β and α_t.

2.1 Identification of the structural parameters

Proposition 1. *Assume the true model is defined by (1)-(4) and (8), with equivalent UC representation defined by (7)-(8); then, the parameters of the UC representation and the structural parameters β and α_t are identified:*

(a) The two covariance matrices of the UC representation are identified by the first two autocovariances of the observed variable vector,

$$\Gamma(0) = \Sigma_\eta + 2\Sigma_v \tag{9}$$

$$\Gamma(1) = -\Sigma_v \tag{10}$$

$$\Gamma(\tau) = 0, \ \tau \geq 2 \tag{11}$$

(b) The structural parameters are given by

$$\beta = \rho_\eta \frac{\sigma_{\eta y}}{\sigma_{\eta x}} \tag{12}$$

$$\alpha_t = y^*_{t-1} - \beta x^*_{t-1} \tag{13}$$

The Proof is provided in Appendix A.

Remark. The estimation of the structural parameters reduces to the estimation of the covariance parameters and the state vector of the UC representation of the model. The identification of α_t and β is possible even though no prior information about the variance of the measurement error, the reliability ratio or other extra information is used. Identification is achieved because the dynamics of the true underlying processes are different from those of the measurement errors. This would not be possible if (y_t^*, x_t^*) were time-invariant. This is the classical case of measurement errors. Reisersol (1950), among others, established that the normality of the underlying series in a static model where measurement errors are also normally distributed, is what spoils identification unless prior information is available.

Another important point to note is that, unlike the traditional treatment of measurement errors, the UC approach allows for the *measurement errors to exist in both y_t^* and x_t^* and for them to be correlated.* The estimation of the UC model is straight forward as it is in a standard state space representation which allows the use of the standard Kalman filter coupled with maximum likelihood estimation. It will prove useful to consider what the OLS estimator will yield if least squares is applied to the differenced series.

Proposition 2. *Assume y_t and x_t follow the UC model in (7)-(8); the OLS regression of the first differences of y_t on x_t will estimate*

$$\hat{\beta}_{OLS} = \frac{\rho_\eta \sigma_{\eta y} \sigma_{\eta x} + 2\rho_\upsilon \sigma_{\upsilon y} \sigma_{\upsilon x}}{\sigma_{\eta x}^2 + 2\sigma_{\upsilon x}^2} \tag{14}$$

The Proof is provided in Appendix A.

Remark. Granger (1986) and Phillips and Durlauf (1986) established that if y_t^* and x_t^* are cointegrated but are observed with error, then the two observed series, y_t and x_t, will also be cointegrated if all measurement errors are $I(0)$. In the bivariate case above y_t and x_t are cointegrated if $\rho_\eta = \pm 1$ which will result in $\alpha_t = \alpha$ (as $\sigma_\alpha^2 = 0$), and then it is easy to show that y_t and x_t are functions of a common stochastic trend. Hassler and Kuzin (2009) study the effect of measurement error in the cointegrated case and find the rate of convergence of the static OLS, and therefore its consistency, is not affected by EIV. Their work assumes the regression error and the measurement error in the regressor variables are not correlated. Buonaccorsi (2010) remarks that the attenuation bias is only guaranteed if the measurement errors in x and y are uncorrelated; when this is not the case, the direction of bias of the OLS estimator is ambiguous. This is clear in (14); as long as $\rho_v \neq 0$, the OLS bias remains.

There are two special cases where OLS applied to the first differenced series can estimate the true structural parameter consistently.

1. There is no measurement error in x_t, $\sigma_{vx}^2 = 0$. Then the measurement error bias caused by $\rho_v \sigma_{vy} \sigma_{vx}$ and σ_{vx}^2 becomes zero. The expression of $\hat{\beta}_{OLS}$ in (14) becomes:

$$\hat{\beta}_{OLS} = \frac{\rho_\eta \sigma_{\eta y} \sigma_{\eta x} + 2\rho_v \sigma_{vy} \sigma_{vx}}{\sigma_{\eta x}^2 + 2\sigma_{vx}^2} = \rho_\eta \frac{\sigma_{\eta y}}{\sigma_{\eta x}} = \beta \tag{15}$$

   This shows that, in the absence of measurement errors in x_t, OLS applied to first differences can correctly identify β.

2. The UC model is homogenous, $p\Sigma_\eta = \Sigma_v$. Technically, this means that the covariance matrix of the measurement errors is proportional to the covariance matrix of the signals. Intuitively, this means all linear combinations of y_t and x_t have the same stochastic properties (Harvey, 1989). When this is the case, the

quantity in (14) collapses to the one in (12).

$$\hat{\beta}_{OLS} = \frac{\rho_\eta \sigma_{\eta y}\sigma_{\eta x} + 2\rho_\upsilon \sigma_{\upsilon y}\sigma_{\upsilon x}}{\sigma^2_{\eta x} + 2\sigma^2_{\upsilon x}} = \frac{\rho_\eta \sigma_{\eta y}\sigma_{\eta x} + 2p\rho_\eta \sigma_{\eta y}\sigma_{\eta x}}{\sigma^2_{\eta x} + 2p\sigma^2_{\eta x}} \tag{16}$$

$$= \frac{(1+2p)\rho_\eta \sigma_{\eta y}\sigma_{\eta x}}{(1+2p)\sigma^2_{\eta x}} = \rho_\eta \frac{\sigma_{\eta y}}{\sigma_{\eta x}} = \beta \tag{17}$$

This coincidence is not just a mathematical reduction of the formula resulting from the restriction $\Sigma_\eta = p\Sigma_\epsilon$. In fact, statistically, this homogeneity directly implies that x_t is exogenous in the sense that conditioning on x_t does not improve the information about the structural parameters we are interested in (Harvey, 1989).

Finally, comparing (14) with (12), it can be seen that by selecting relevant components (the variances and the correlation of the signals, not those of the measurement errors), the bias due to measurement errors can be extracted and the true structural parameters of interest identified.

This subsection has provided the identification of the structural parameters when the true variables are unobservable and assumed to follow random walk processes. The following subsection exemplifies an extension where cyclicality is incorporated into the underlying true series. Seasonality might be incorporated in a similar fashion.

2.2 Alternative specification of the components model - trend plus cycle model.

An alternative and more general specification from the simple random walk case above is to assume the following specification (which allows for the presence of cycles; seasonality could also be incorporated):

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} y_t^* \\ x_t^* \end{bmatrix} + \begin{bmatrix} v_{yt} \\ v_{xt} \end{bmatrix} \tag{18}$$

$$\begin{bmatrix} y_t^* \\ x_t^* \end{bmatrix} = \begin{bmatrix} \mu_{yt} \\ \mu_{xt} \end{bmatrix} + \begin{bmatrix} c_{yt} \\ c_{xt} \end{bmatrix} \tag{19}$$

$$\begin{bmatrix} \mu_{yt} \\ \mu_{xt} \end{bmatrix} = \begin{bmatrix} \mu_{yt-1} \\ \mu_{xt-1} \end{bmatrix} + \begin{bmatrix} \eta_{yt} \\ \eta_{xt} \end{bmatrix} \tag{20}$$

$$\begin{bmatrix} c_{yt} \\ \tilde{c}_{yt} \\ c_{xt} \\ \tilde{c}_{xt} \end{bmatrix} = \begin{bmatrix} \rho_y cos\lambda_{cy} & \rho_y sin\lambda_{cy} & 0 & 0 \\ -\rho_y sin\lambda_{cy} & \rho_y cos\lambda_{cy} & 0 & 0 \\ 0 & 0 & \rho_x cos\lambda_{cx} & \rho_x sin\lambda_{cx} \\ 0 & 0 & -\rho_x sin\lambda_{cx} & \rho_x cos\lambda_{cx} \end{bmatrix} \begin{bmatrix} c_{yt-1} \\ \tilde{c}_{yt-1} \\ c_{xt-1} \\ \tilde{c}_{xt-1} \end{bmatrix} + \begin{bmatrix} \omega_{yt} \\ \tilde{\omega}_{yt} \\ \omega_{xt} \\ \tilde{\omega}_{xt} \end{bmatrix} \tag{21}$$

where the cycles disturbance vector follows a normal distribution with covariance matrix

$$\Sigma_\omega = \begin{bmatrix} \sigma_{\omega y}^2 & 0 & \rho_\omega \sigma_{\omega y} \sigma_{\omega x} & 0 \\ 0 & \sigma_{\omega y}^2 & 0 & \rho_\omega \sigma_{\omega y} \sigma_{\omega x} \\ \rho_\omega \sigma_{\omega y} \sigma_{\omega x} & 0 & \sigma_{\omega x}^2 & 0 \\ 0 & \rho_\omega \sigma_{\omega y} \sigma_{\omega x} & 0 & \sigma_{\omega x}^2 \end{bmatrix}$$

Equation (19) defines the true underlying vector as comprising of cycles and trends. Equation (20) assumes that the trend vector follows a random walk process with the innovation $(\eta_{yt}, \eta_{xt})'$, which follows a normal distribution with mean zero and covariance matrix Σ_η. Equations (21) specify the dynamics of the cycles. Since the cycle in each series is driven by two disturbances, there are two sets of disturbances and these are assumed to have the same covariance matrix (following the convention described in Harvey (1989) and Koopman et al. (1995)). ρ_y and ρ_x are the damping factors while λ_{cy} and λ_{cx} are the frequency of the cycles in y_t^* and x_t^*, respectively.

Under normality of the system disturbance vectors, the distribution of the true vector is normal with mean vector and covariance matrix given by:

$$
\begin{aligned}
E\left(\begin{bmatrix} y_t^* \\ x_t^* \end{bmatrix} |\Im_{t-1}\right) &= \begin{bmatrix} \mu_{yt-1} \\ \mu_{xt-1} \end{bmatrix} + \begin{bmatrix} \rho_y cos\lambda_{cy} c_{yt-1} + \rho_y sin\lambda_{cy} \tilde{c}_{yt-1} \\ \rho_x cos\lambda_{cx} c_{xt-1} + \rho_x sin\lambda_{cx} \tilde{c}_{xt-1} \end{bmatrix} \\
&= \begin{bmatrix} \mu_{yt-1} \\ \mu_{xt-1} \end{bmatrix} + \begin{bmatrix} c_{yt-1} \\ c_{xt-1} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
Var\left(\begin{bmatrix} y_t^* \\ x_t^* \end{bmatrix} |\Im_{t-1}\right) &= \Sigma_\eta + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} Var\left(\begin{bmatrix} \omega_{yt} \\ \tilde{\omega}_{yt} \\ \omega_{xt} \\ \tilde{\omega}_{xt} \end{bmatrix}\right) \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \\
&= \Sigma_\eta + \begin{bmatrix} \sigma_{\omega y}^2 & \rho_\omega \sigma_{\omega y} \sigma_{\omega x} \\ \rho_\omega \sigma_{\omega y} \sigma_{\omega x} & \sigma_{\omega x}^2 \end{bmatrix} \\
&= \Sigma_\eta + \Sigma_c
\end{aligned}
$$

Therefore, applying the result from Proposition 1, we can identify the structural parameters as:

$$
\begin{aligned}
\beta &= \frac{(\rho_\omega \sigma_{\omega y} \sigma_{\omega x} + \rho_\eta \sigma_{\eta y} \sigma_{\eta x})}{\sigma_{\eta x}^2 + \sigma_{\omega x}^2} && (22) \\
\alpha_t &= (\mu_{yt-1} + c_{yt-1}) - \beta(\mu_{xt-1} + c_{xt-1}) && (23) \\
\alpha_t &= y_{t-1}^* - \beta x_{t-1}^* && (24)
\end{aligned}
$$

Which model (local level trend or trend plus cycle) is to be estimated to obtain the structural parameters is an empirical question. Therefore, model selection must precede estimation. As the model has a standard structural time series form, model selection and estimation can be implemented using many existing commercial softwares.

3 Estimation

The parameters required to obtain an estimate of β (using expression (12)/(22)) require estimation of the covariances Σ_η (Σ_η and Σ_c) and Σ_v. The parameters in these covariances are hyperparameters of the state-space representation in (7)-(8)/(18)-(20); and, the underlying means, $\mu_t = (y_t^*, x_t^*)'$, are the state vector of the representation. Thus, the log-likelihood can be formed and computed by running the standard equations of the Kalman filter (for further issues on specification and estimation of structural time series models the reader is referred to Durbin and Koopman (2012)).

Denote by $\tilde{\beta}$ and $\tilde{\alpha}_t$ the estimators of the parameters of the structural model in (3). To show their form define $\mathbf{y}_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix}$, $\mu_t = \begin{pmatrix} y_t^* \\ x_t^* \end{pmatrix}$, $Z_t = I$, $T_t = I$, $R_t = I$ (I is the identity matrix), and thus the UC model rewritten in the standard state space form is,

$$
\begin{aligned}
\mathbf{y}_t &= Z_t \mu_t + v_t; && (25) \\
\mu_t &= T_t \mu_{t-1} + R_t \eta_t && (26)
\end{aligned}
$$

The Kalman Filter algorithm (KF) is implemented by running through the prediction and the updating steps of the filter given estimates of Σ_v and Σ_η. Let $m_{t|t-1}$ be the optimal estimator of μ_t given information up to and including the period $t-1$ and $P_{t|t-1}$ be the mean square error matrix of $m_{t|t-1}$. Also let m_t be the optimal estimator of μ_t once the t^{th} observation is realized and P_t be the mean square error matrix of this estimator. For the system in (7)-(8), the predicting step is given by:

$$m_{t|t-1} = T_t m_{t-1} \tag{27}$$

$$P_{t|t-1} = T_t P_{t-1} T_t' + R_t \Sigma_\eta R_t' \tag{28}$$

Once a new observation is available, the updating will proceed through the following equations:

$$m_t = m_{t|t-1} + P_{t|t-1} Z_t' F_{t.}^{-1} \epsilon_t \tag{29}$$

$$P_t = P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1} \tag{30}$$

where $F_t = Z_t P_{t|t-1} Z_t' + \Sigma_v$. $\epsilon_t = \mathbf{y}_t - Z_t m_{t|t-1}$ is called innovation, and used as the driver of the updating process. Starting from the assumed distribution of the initial state vector with mean, u_0, and covariance matrix, P_0, the KF delivers the optimal estimate of the state vector μ_t. It can be shown that under the normality assumption of v_t and η_t, $\mathbf{y}_t$ is also normally distributed conditional on past information with mean $E(\mathbf{y}_t) = Z_t m_{t|t-1}$ and variance $F_t = Z_t P_{t|t-1} Z_t' + \Sigma_v$. Accordingly, the log-likelihood function can be easily constructed as

$$\begin{aligned} \mathcal{L}(\psi|\mathbf{y}_t, m_t) &= log\left(\prod_{t=1}^{T} p(\mathbf{y}_t|Y_{t-1})\right) \\ &= \frac{NT}{2}log2\pi - \frac{1}{2}\Sigma_{t=1}^{T} log|F_t| - \frac{1}{2}\Sigma_{t=1}^{T}(\mathbf{y} - Z_t m_{t|t-1})F_t^{-1}(\mathbf{y} - Z_t m_{t|t-1}) && (31) \\ &= \frac{NT}{2}log2\pi - \frac{1}{2}\sum_{t=1}^{T} log|F_t| - \frac{1}{2}\sum_{t=1}^{T} \epsilon_t' F_t^{-1} \epsilon_t && (32) \end{aligned}$$

This form of the log-likelihood function provides a convenient avenue to estimate the hyper-parameters, $\psi = \{\sigma^2_{vy}, \sigma^2_{vx}, \rho_v, \sigma^2_{\eta y}, \sigma^2_{\eta x}, \rho_\eta\}$. Given the maximum likelihood estimates (MLE), $\tilde{\psi}$, the Kalman filter and smoothing algorithms are used to provide estimates of the state vector $m_{t|T} = (\tilde{y}^*_t, \tilde{x}^*_t)'$. The estimators of β and α_t are given by,

$$\tilde{\beta} = \tilde{\rho}_\eta \frac{\tilde{\sigma}_{\eta y}}{\tilde{\sigma}_{\eta x}} \tag{33}$$

$$\tilde{\alpha}_t = \tilde{y}^*_t - \tilde{\beta}\tilde{x}^*_t \tag{34}$$

Similar arguments follow to obtain estimators from the system in (18)-(20). This estimation can be carried out with any commercial software that contains a state-space estimation module, such as STATA or EViews. In the empirical section we use Structural Time Series Analysers, Modeler and Predictor (STAMP).

In the next section three simulation exercises are presented to demonstrate Propositions 1 and 2 and the performance of (33) and (34).

4 Simulations

We present three simulations. The first considers the case when there is no measurement error in x_t, and the true data generation process (DGP) is as (7)-(8). The OLS, $\hat{\beta}_{OLS}$, and UC, $\tilde{\beta}$, based estimates to be identical (by Proposition 2). The second simulation considers the case when there is measurement error in x_t and the true data generation process (DGP) is (7)-(8). By Proposition 1, the components based estimates of the slope and time-varying level parameters will be unbiased. By Proposition 2 the size of the bias in the OLS estimate is known and the simulations confirms this result. In the third case, the data generation process is the structural regression relationship (1)-(4) Each simulation consists of 200 replications. Each data set has 100 time periods (after the first 50 observations are trimmed to ensure the time-series are not affected by the initializations). In all graphs we will denote by "OLS" the estimate of β obtained by running a regression of Δy_t on Δx_t, and by "SF" the estimates of β and α_t using the MLE and standard filtering/smoothing estimates, $\tilde{\psi}$ and $m_{t|T} = (\tilde{y}_t^*, \tilde{x}_t^*)'$ and expressions (33) and (34).

Simulation 1: *No measurement error in x_t - True Data Generating Process (DGP) is the UC model*

Data are generated to follow the UC model in (7)-(8) with the following hyperparameters

$$\Sigma_v = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} ; \ \Sigma_\eta = \begin{pmatrix} 7.5 & 4.5 \\ 4.5 & 3 \end{pmatrix} ; \ \begin{pmatrix} \mu_{y0} \\ \mu_{x0} \end{pmatrix} = \begin{pmatrix} 20 \\ 5 \end{pmatrix}$$

In this specification, x_t is measured correctly. The true values of the structural parameters α_t, β are given by

$$\begin{aligned} \beta &= cov_\eta / \sigma_{\eta x}^2 = 4.5/3 = 1.5; \\ \alpha_t &= y_{t-1}^* - \beta x_{t-1}^* = y_{t-1}^* - 1.5 x_{t-1}^* \end{aligned}$$

Results of Simulation 1

Figure 1a) presents the distributions of the OLS and the estimates of β computed using expressions in (12) and (13). By Proposition 2 we expect that in the absence of measurement error in x_t, both the OLS estimator of the regression in first differences and the SF estimator, $\tilde{\beta}$, can estimate correctly the slope. Both estimators have approximately the same distribution, which is symmetric and centered around the true value, $\beta = 1.5$. The SF can estimate correctly the slope because it disentangles the true structural parameters from the bias due to measurement errors. In this case, no measurement error is present in x_t. The SF estimator, of course, remains valid. In contrast, it is due to the zero measurement error in x_t that the OLS estimator yields the true structural slope.

Confidence intervals for both OLS and SF estimates in each replication are constructed. The rate at which the confidence intervals of the SF estimates include the value true $\beta = 1.5$ is $189/200 = 94.5\%$. The rate of coverage for the OLS estimates is 186/200=93%. This means we can be confident at the 10% level of significance that both the SF estimator and the OLS estimator can estimate correctly the slope in the structural relationship.

OLS in first differences is not able to estimate the time-varying parameter. In contrast, by estimating the UC model, the estimate $\tilde{\alpha}_t$ is obtained. Pointwise confidence intervals for the estimation errors of $\tilde{\alpha}_t$ are constructed. The coverage rate of the confidence intervals over the true values at all points on the series ranges from 93.5% to 97.5%. This means the estimation errors are insignificant at the 10% level of significance at all points of the estimated series. Figure 1b) presents the true α_t together with $\tilde{\alpha}_t$ from the 19th sample.

[Figure 1 here]

Simulation 2: *Measurement error in x_t - True DGP is the UC model*

The hyper-parameters of the UC in this experiment are set to:

$$\Sigma_v = \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix}; \Sigma_\eta = \begin{pmatrix} 7.5 & 4.5 \\ 4.5 & 3 \end{pmatrix}; \begin{pmatrix} y_0^* \\ x_0^* \end{pmatrix} = \begin{pmatrix} 20 \\ 5 \end{pmatrix}$$

Note that these covariance matrices are chosen to be general in the sense that they are not proportional and the measurement error in the input series exists, $\sigma_{vx}^2 = 3 \neq 0$.

With this specification, the true structural parameters α_t, β are

$$\beta = cov_\eta/\sigma^2_{\eta x} = 1.5 \; ; \; \alpha_t = y^*_{t-1} - 1.5x^*_{t-1}$$

In this experiment it is verified that OLS applied to the first differences fails to estimate the true slope. In contrast $\tilde{\beta}$ yields the correct estimate. Furthermore, OLS applied to the first differences will estimate,

$$\hat{\beta}_{OLS} = \frac{2\sigma_{vy}\sigma_{vx}\rho_v + \sigma_{\eta y}\sigma_{\eta x}\rho_\eta}{2\sigma^2_{vx} + \sigma^2_{\eta x}} = \frac{2 \times 2 + 4.5}{2 \times 3 + 3} = 0.94$$

Results of Simulation 2

It is strikingly clear from the kernel density plots in Figure 2a) that OLS estimates are well below the true β. They are centered around a value just below 1. The null hypothesis $H_o : \beta = 0.94$ tested for the OLS estimates is not rejected in 92% of the replications, indicating that the parameter that OLS estimates is not significantly different from 0.94. Hence, the OLS estimator does not return the true structural parameter of interest. The plot also shows that the SF estimates looks normally distributed, centered around the value 1.5, the true value of β. The rate at which the confidence intervals of SF estimates includes 1.5 is 192/200 = 96%. Therefore we can be confident that the SF estimator can estimate the true β at the 5% level of significance.

The pointwise confidence intervals for the estimation errors of $\tilde{\alpha}_t$ are computed. The coverage of those confidence intervals over zero at all points in the estimated series ranges from 93.5% to 97.5%, implying the estimation errors are insignificant at the 10% level of significance at all points in the series. As shown in Figure 2b), the estimated α_t tracks its true series well.

[Figure 2 here]

Simulation 3: *Measurement error in x_t - True DGP is the structural model*

This experiment depart from the UC framework to a different DGP as follows,

$$
\begin{aligned}
y_t &= y_t^* + v_{yt}; \\
x_t &= x_t^* + v_{xt}; \\
\begin{pmatrix} v_{yt} \\ v_{xt} \end{pmatrix} &\sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix} \right) \\
y_t^* &= \alpha_t + \beta x_t^*; \\
\beta &= 1.5 \\
\alpha_t &= 0.08 + \alpha_{t-1} + \epsilon_{\alpha,t};\ \alpha_0 = 10;\ \epsilon_{\alpha,t} \sim N(0, 0.5^2) \\
x_t^* &= x_{t-1}^* + \eta_{xt};\ x_0^* = 5;\ \eta_{xt} \sim N(0, 2.5^2)
\end{aligned} \tag{35}
$$

The x_t and y_t series are measured with errors, $\begin{pmatrix} v_{yt} & v_{xt} \end{pmatrix}'$, which are jointly normally distributed. The true relationship is expressed in terms of the underlying correct measures, y_t^* and x_t^*, in which the slope is constant whereas the intercept varies according to a random walk with drift process. Finally, x_t^* is a random walk process with a normally distributed random noise, η_{xt}, which has variance of 2.5^2 .

By (35),

$$
\begin{aligned}
x_t^* | \Im_{t-1} &\sim N(x_{t-1}^*, 2.5^2) \\
\alpha_t | \Im_{t-1} &\sim N(0.08 + \alpha_{t-1}, 0.5^2)
\end{aligned}
$$

Because x_t^* and α_t are independent, $y_t^* | \Im_{t-1} = \alpha_t + \beta x_{t-1}^* | \Im_{t-1} \sim N(0.08 + \alpha_{t-1} + \beta x_{t-1}^*, 2.5^2 + \beta^2 \times 0.5^2)$. Since $y_{t-1}^* = \alpha_{t-1} + \beta x_{t-1}^*$, we can write the marginal distribution of y_t^* (conditional on past information) as

$$y_t^* | \Im_{t-1} \sim N(0.08 + y_{t-1}^*, 2.5^2 + \beta^2 \times 0.5^2)$$

Results of Simulation 3

Figure 3a) represents the distributions of the SF and OLS estimates of β. Clearly, OLS estimates are below the true value, $\beta = 1.5$, whereas the distribution of SF estimates are symmetric, centered around this true value. It is also calculated that the cover rate of confidence intervals constructed from the SF estimates of the UC model over this true value is 96.3%. In contrast, no confidence intervals constructed from OLS estimates includes the true value.

Moreover, OLS cannot be used to estimate α_t. In Figure 3b) an example of the estimates of α_t obtained from the UC representation for the 1st replication in the experiment are provided.

[Figure 3 here]

5 Empirical Illustration

The empirical application estimates returns to scale and technological change for four US industrial sectors and the US aggregate manufacturing sector. Two recent papers that have considered and modeled technical progress in US industrial sectors are those by Diewert and Fox (2008)-D&F and Jin and Jorgenson (2010)-J&J. D&F first use an index number approach to construct two indexes, one of inputs and one of outputs for each sector. X_t and Y_t are Törnquist indexes of the growth in inputs and outputs at period t, respectively, with base period being $t-1$, i.e., $X_t = \ln Q_t(w^{t-1}, w^t, x^{t-1}, x^t)$ and $Y_t = \ln Q_t(p^{t-1}, p^t, y^{t-1}, y^t)$ where y^t and x^t are the output and input vectors and w^t and p^t are the output and input prices, respectively. Five inputs, including capital, labor, energy, materials and purchased business services inputs are included. Outputs are deflated added-value of production. They then show that these two indexes are theoretically related through two parameters, returns to scale and change in technical change under the assumptions of cost reducing technical change and constant markup factors within each period across commodities, $X_t = -\tau + \rho Y_t$, with ρ the reciprocal of returns to scale parameter and τ is a technical change parameter, defined as $\theta_t - \theta_{t-1}$ where θ_t is the period t productivity parameter. These two parameters can be estimated econometrically. An EIV form of the model is derived (pp 182) and the argument is made that τ should be time-varying; however, the results presented are based on a fixed parameters using equation (36) and least squares. The model is implemented for eighteen US industrial sectors using data from BLS.

$$Y_t = \tau/\rho + (1/\rho)X_t \tag{36}$$

J&J use a fully econometric approach to estimate technical change which is modeled as an unobserved trend. Their model is a fully specified cost function approach. The value of output is equal to the value of four inputs, capital, labor, energy and non-energy materials. A translog price function, the input share equations and a number of economic constraints (e.g. homogeneity, symmetry of shared elasticities and local concavity) are imposed. The model is written in state-space form and technical change is assumed to evolve as a stochastic trend. Therefore, the change in technical change is stationary. Their model uses prices as explanatory variables which induces endogeneity. The estimation approach is a standard Kalman filtering/smoothing approach; however, it requires two main modifications. First, replacing endogenous variables by instrumented price predictions and second a two-step maximum likelihood of the parameters in the unknown matrices of the state-space system. The model is implemented for 35 industrial sectors using data from Jorgenson et al. (2007).

In this paper we estimate a revised form of the D&F relationship by allowing change in technical change to be time-varying as in J&J's work. The model is estimated using a revised vintage of D&F dataset released by BLS in 2015, which covers the period 1949 - 2001 (Appendix B provides details on data preparation). Sectoral models are estimated for four industrial sectors that are in both the J&J and D&F studies, as well as the aggregate US manufacturing sector, which is only in D&F. The four industrial sectors are, Textile Mill Production (SIC22), Apparel & Related Production (SIC23), Industrial and Commercial Machinery And Computer Equipment (SIC35), Electrical and Electronic Equipment (SIC36)[1].

5.1 Model, Estimation and Results

Let $\alpha_t = \tau_t/\rho$ and $\beta = 1/\rho$; thus, $\alpha_t > 0$ indicates positive technological change, and $\beta > 1$ implies increasing returns to scale. The economic specification is of the form

$$Y_t = \alpha_t + \beta X_t \tag{37}$$

We treat the output and input indexes as a jointly distributed bivariate vector (Y_t, X_t) with elements measured with errors and fit three forms of the UC model, a local level trend model (LLT), a trend plus cycle (the trend is a random walk with drift) model (TC), and a level plus cycle model (LC), where the trend is a random walk without drift. The last two versions will imply technological change also has some cyclical characteristics. Through a statistical model selection process, the data can fit α_t without imposing strong assumptions about its stochastic properties.

The LLT model is given by

[1]SIC stands for Standard Industrial Classification

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} y_t^* \\ x_t^* \end{bmatrix} + \begin{bmatrix} \upsilon_{yt} \\ \upsilon_{xt} \end{bmatrix} \tag{38}$$

$$\begin{bmatrix} y_t^* \\ x_t^* \end{bmatrix} = \begin{bmatrix} y_{t-1}^* \\ x_{t-1}^* \end{bmatrix} + \begin{pmatrix} \varsigma_y \\ \varsigma_x \end{pmatrix} + \begin{pmatrix} \eta_{yt} \\ \eta_{xt} \end{pmatrix} \tag{39}$$

where $\varsigma = (\varsigma_y, \varsigma_x)'$ is the vector of drifts.

The TC model is given by

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} y_t^* \\ x_t^* \end{bmatrix} + \begin{bmatrix} \upsilon_{yt} \\ \upsilon_{xt} \end{bmatrix} \tag{40}$$

$$\begin{bmatrix} y_t^* \\ x_t^* \end{bmatrix} = \begin{bmatrix} \mu_{yt} \\ \mu_{xt} \end{bmatrix} + \begin{bmatrix} c_{yt} \\ c_{xt} \end{bmatrix} \tag{41}$$

$$\begin{bmatrix} \mu_{yt} \\ \mu_{xt} \end{bmatrix} = \begin{bmatrix} \mu_{yt-1} \\ \mu_{xt-1} \end{bmatrix} + \begin{pmatrix} \varsigma_y \\ \varsigma_x \end{pmatrix} + \begin{bmatrix} \eta_{yt} \\ \eta_{xt} \end{bmatrix} \tag{42}$$

$$\begin{bmatrix} c_{yt} \\ \tilde{c}_{yt} \\ c_{xt} \\ \tilde{c}_{xt} \end{bmatrix} = \begin{bmatrix} \rho_y cos\lambda_{cy} & \rho_y sin\lambda_{cy} & 0 & 0 \\ -\rho_y sin\lambda_{cy} & \rho_y cos\lambda_{cy} & 0 & 0 \\ 0 & 0 & \rho_x cos\lambda_{cx} & \rho_x sin\lambda_{cx} \\ 0 & 0 & -\rho_x sin\lambda_{cx} & \rho_x cos\lambda_{cx} \end{bmatrix} \begin{bmatrix} c_{yt-1} \\ \tilde{c}_{yt-1} \\ c_{xt-1} \\ \tilde{c}_{xt-1} \end{bmatrix} + \begin{bmatrix} \omega_{yt} \\ \tilde{\omega}_{yt} \\ \omega_{xt} \\ \tilde{\omega}_{xt} \end{bmatrix} \tag{43}$$

The model labelled LC is equivalent to TC but without the drift component, ς.

This empirical analysis consists of two parts. First a model selection is carried out to determine whether the UC model should be LLT, TC or LC. Second, the chosen model is estimated for each sector and the parameters of (37) are estimated using the relevant form of $\tilde{\beta}$ and $\tilde{\alpha}_t$ given the included components. STAMP is used here (although any standard commercial software where estimation by state-space is available can be used).

Model selection

Figure 4 shows, as an example, the plot of the aggregate manufacturing sector's input growth index, X_t, and output growth index Y_t over the sample of 49 years. The plot suggests the trend components, μ_t, are likely to be fairly stable, while there might be significant cycle components, c_t.

The selection criteria are computed for all five sectors, and results are presented in Table 1. The LC specification has the smallest value of both criteria in all cases except for SIC23 where the minimum value of both AIC and BIC is for the specification TC.

[Figure 4 and Table 1]

Parameters estimates and discussion of results

The model chosen by the criteria for each sector is fitted. From the estimated hyper-parameters in Σ_η and Σ_c, and the state vector, $(\mu_{yt}, \mu_{xt})'$ and $(c_{yt}, c_{xt})'$, the estimates of the structural parameter as defined in expressions (22) and (23) are computed using the smoothed estimates ($\tilde{y}_t^*$ and $\tilde{x}_t^*$). The hyper-parameters estimates and the estimates of the regression parameters, β, are reported in Table 2. To illustrate the unobserved components, Figure 5 presents the plots of the estimated components for the Electric & Electrical Equipment sector's X_t and Y_t series. The estimates of the elements of μ_t are labelled 'level' and those of c_t 'cycle'. These are the components that enter the computation of the structural parameters, $\tilde{\beta}$ and $\tilde{\alpha}_t (= \tau_t/\rho)$. The OLS estimates of β from D&F's paper and the OLS estimates using the new vintage data are presented on the last two rows of the table. The data vintage effect appears to be negligible as expected. The top rows contain the estimates of the variances and correlations of the components (trend, cycle and measurement error).

[Table 2 and Figure 5]

For some sectors the cycle variances are larger than those of the trend component, and thus cycles appear to be important contributors to the total variation of the observed series, which supports the choice of the model that allows for cyclicality. In the aggregate manufacturing case, the trend component is constant (estimated $\sigma^2_{\eta y}$ and $\sigma^2_{\eta x}$ are effectively zero), while the stochastic cycle dominates the signal component. All other sectors have both stochastic trends and cycles; however, the cycles in SIC22 and SIC23 have very small variances compared to those of the other sectors. Notably SIC35 shows a negative correlation between the output and input indexes' trend components.

It is evident that measurement errors are important components of the observed series and there is strong evidence of a correlation between the measurement errors in X_t and Y_t in all sectoral models. As Proposition 2 indicates, this correlation leads to an ambiguous size in the bias of the OLS estimator caused by the correlated measurement errors in the observed variables. The empirical evidence is that the difference between $\tilde{\beta}$ and $\hat{\beta}_{OLS}$ across sector/industries varies. For aggregate manufacturing and SIC35 the estimate of β is lower than that obtained by OLS, it is substantially higher for SIC22 and SIC36 and remains unchanged for SIC23. In terms of conclusions on sectoral returns to scale, these results agree with the least squares based conclusion that SIC23 was characterized by decreasing returns and SIC35 and SIC36 by increasing returns to scale. The conclusions differ on the findings for aggregate manufacturing and SIC22. The OLS estimate is below one for SIC22; however, once measurement error is taken into account SIC22 showed increasing returns to scale over the period (estimate is 1.278, substantially higher than 1). For aggregate manufacturing the estimate is very close to 1, around 24% lower than that obtained using OLS suggesting evidence for constant returns to scale instead of increasing returns to scale for the sector.

The $\tilde{\alpha}_t$ and $\tilde{\beta}$ are used to obtain the technical change parameters τ_t (since $\alpha_t = \tau_t/\rho$

and $\rho = {}^{1}/_{\beta}$), which are graphed in Figure 6. J&J's model estimated the level of technical change, θ_t in our notation. The paper reports what they label rate of autonomous technical change which is defined as the difference between the levels of technical change in periods 1 and T (1960 and 2005 in their case), we refer to this as Δ (refer to their Figure 16). They find Δ to be the largest for Industrial Machinery & Computer Equipment, and Electric & Electrical Equipment. Apparel & Related Products and Textile Mill Production have smaller Δ over the sample period; however, they are both positive and larger than most other sectors considered in their study. D&F find the estimates of technological change (fixed τ in their model) to be very modest and mostly not significantly different from zero. The highest estimate is 2.2% technical progress over the period 1949-2000 for the sector Textile Mills Productions (SIC22), for Industrial Machinery & Computer Equipment (SIC35) is 0.7%, for Electric & Electrical Equipment (SIC36) is 1% and for Apparel and Related Products (SIC23) is 0.8%.

[Figure 6]

Allowing for time-varying technological change parameters provides an opportunity to study the effects of important events as well as trends within each sector. Our estimate of τ_t for Textile Mills Production, Figure 6, is mostly constant and just above 0.02. This is the closest of results between D&F and our estimates. For Aggregate Manufacturing, D&F cannot reject the null of no technical progress. The time-varying estimates show why this might have been the case. Significantly, it picks up a large drop in technical change in 1974, to coincide with the oil crisis, and a number of other periods when technical change was slightly negative, which can explain why on average the estimate was found to be zero. In general, the changes in technical change seem fairly constant for the earlier period up to 1980. However, from 1980 onwards some trends are clear. Overall the results are more aligned with those of J&J in that Industrial Machinery & Computer Equipment (SIC35) and Electric & Electrical Equipment

(SIC36) exhibit sustained and significant progress, this is the well known US IT boom and is clear after 1980, placing these two sectors at the top of the rankings in technological change. The Electronic and Electrical Equipment industry produces semiconductor components for computers and other electronic equipment, while Industrial Machinery and Equipment includes computers. Jin and Jorgenson (2010) show these two sectors were at the top in terms of reduction of relative output prices, i.e., a dramatic fall in the prices of outputs relative to the materials (inputs) they consume, and are also ranked at the top in terms of the rate of autonomous technical change (Δ) observed over the period of their study. This increasing trend in technical progress is very clear in our estimates as well (see Figure 6).

6 Conclusions

The paper considers the estimation of structural parameters of interest when all observed variables are measured with errors and these are correlated. Contaminated variables are modeled as jointly distributed in an unobserved components framework. This provides a venue for identification of the parameters of interest and a way of modeling them as time-varying. The theoretical results are presented through two propositions and a simulation study. The empirical application estimates technical change and returns to scale in US industrial sectors. The example is chosen as the results can be compare to those obtained in recent works in the literature using alternative modeling strategies.

Acknowledgment

The first and third authors would like to thank Andrew Harvey for the valuable discussions.

A Proof of Propositions

Proof of Proposition 1

Proof. (a) The identifiability of the UC representation can be seen by evoking the stochastic properties of the vector $\triangle Y_t = (\triangle y_t, \triangle x_t)$ (using Harvey 1989, p. 431) since $\triangle Y_t = \eta_t + v_t - v_{t-1}$ is a stationary process. The autocovariance matrices are given by (9)-(10).

(b) Since y_t and x_t are measured with error, the structural relationship between them should be derived from their underlying means y_t^* and x_t^*. From (8) and the normality assumption of $(\eta_{yt}, \eta_{xt})'$, the joint distribution of the underlying means of y_t and x_t given past information $\Im_{t-1} = \{(y_i, x_i), i = 1, ...t-1\}$ is:

$$\begin{pmatrix} \mu_{yt} \\ \mu_{xt} \end{pmatrix} |\Im_{t-1} \sim N\left(\begin{pmatrix} y_{t-1}^* \\ x_{t-1}^* \end{pmatrix}; \begin{pmatrix} \sigma_{\eta y}^2 & \rho_\eta \sigma_{\eta y} \sigma_{\eta x} \\ \rho_\eta \sigma_{\eta y} \sigma_{\eta x} & \sigma_{\eta x}^2 \end{pmatrix} \right) \tag{44}$$

Since a multivariate normal distribution of a vector of variables implies normal distribution of each variable conditional on the other variables, (see Greene (2012) for instance), y_t^* is normally distributed conditional on x_t^*,

$$y_t^* | x_t^*, \Im_{t-1} = \left(y_{t-1}^* - \rho_\eta \frac{\sigma_{\eta y}}{\sigma_{\eta x}} x_{t-1}^* \right) + \rho_\eta \frac{\sigma_{\eta y}}{\sigma_{\eta x}} x_t^* + \xi_t \tag{45}$$

where $\xi_t \sim N(0, \sigma_\xi^2)$ and $\sigma_\xi^2 = \sigma_{\eta y}^2 - \rho_\eta \sigma_{\eta x}^2$. Matching this relationship between the conditional mean of y_t^* given x_t^* and the exact relationship equation (3), we can identify the structural parameters as having the form:

$$\begin{aligned} \beta &= \rho_\eta \frac{\sigma_{\eta y}}{\sigma_{\eta x}} \\ \alpha_t &= y_{t-1}^* - \beta x_{t-1}^* \end{aligned}$$

$\square$

Proof of Proposition 2

Proof. From (7) - (8), write the first differences of the observed variables vector as a sum of the signal and the measurement error vector:

$$\begin{pmatrix} \triangle y_t \\ \triangle x_t \end{pmatrix} = \begin{pmatrix} \eta_{yt} \\ \eta_{xt} \end{pmatrix} + \begin{pmatrix} \upsilon_{yt} \\ \upsilon_{xt} \end{pmatrix} - \begin{pmatrix} \upsilon_{yt-1} \\ \upsilon_{xt-1} \end{pmatrix}$$

Using the assumptions made on the UC representation, the distribution of the first differenced series is given by

$$\begin{pmatrix} \triangle y_t \\ \triangle x_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 2\sigma^2_{\upsilon y} + \sigma^2_{\eta y} & 2\rho_\upsilon \sigma_{\upsilon y}\sigma_{\upsilon x} + \rho_\eta \sigma_{\eta y}\sigma_{\eta x} \\ 2\rho_\upsilon \sigma_{\upsilon y}\sigma_{\upsilon x} + \rho_\eta \sigma_{\eta y}\sigma_{\eta x} & 2\sigma^2_{\upsilon x} + \sigma^2_{\eta x} \end{pmatrix}\right)$$

The regression that corresponds to this bivariate normal distribution is given by

$$\triangle y_t = \frac{2\rho_\upsilon \sigma_{\upsilon y}\sigma_{\upsilon x} + \rho_\eta \sigma_{\eta y}\sigma_{\eta x}}{2\sigma^2_{\upsilon x} + \sigma^2_{\eta x}} \triangle x_t + \xi_t \tag{46}$$

By construction, $\triangle x_t$ is uncorrelated to ξ_t (see, for instance, Greene (2012)). OLS applied to the first differenced series will provide an estimate of the parameter:

$$\hat{\beta}_{OLS} = \frac{\rho_\eta \sigma_{\eta y}\sigma_{\eta x} + 2\rho_\upsilon \sigma_{\upsilon y}\sigma_{\upsilon x}}{\sigma^2_{\eta x} + 2\sigma^2_{\upsilon x}}$$

□

A Data preparation

There are some small differences between the data used in D&F's paper and the data used here. The data available from BLS now covers the period of 1949 to 2001 (one more year from that used by D&F); however, the years 1950 to 1952 are missing. Therefore, the data used for estimation is for the period 1953-2001.

X^t, the Törnquist input quantity index, is an aggregate of five inputs, including capital, labor, energy, materials and purchased business services inputs.

$$X^t = \frac{1}{2}\sum_{n=1}^{5}(s_n^t + s_n^{t-1})(lnx_n^t - lnx_n^{t-1}) = \frac{1}{2}\sum_{n=1}^{5}(s_n^t + s_n^{t-1})ln(x_n^t/x_n^{t-1})$$

where s_n^t is the cost share of the input n at time t, $s_n^t = cost\ of\ x_n^t/total\ cost^t$. The cost of each input and the total cost of production are available in the data set. However, the data set does not contain the level of each input used. Only the annual percentage change and the index for each input (the base year is $t = 1996$) are presented. Using these two series, x_n^t/x_n^{t-1} can be calculated in two ways: $\frac{x_n^t}{x_n^{t-1}} = \frac{x_n^t/x_n^{1996}}{x_n^{t-1}/x_n^{1996}} = \frac{index\ year\ t}{index\ year\ t-1}$ or $\frac{x_n^t}{x_n^{t-1}} = \left(\frac{x_n^t - x_n^{t-1}}{x_n^{t-1}} + 1\right) = \left(\frac{\%annual\ change}{100} + 1\right)$.

The output quantity index, Y^t can be calculated as

$$Y^t = \frac{1}{2}\sum_{m=1}^{M}(\sigma_m^t + \sigma_m^{t-1})(lny_m^t - lny_m^{t-1}) = \frac{1}{2}\sum_{m=1}^{M}(\sigma_m^t + \sigma_m^{t-1})ln(y_m^t/y_m^{t-1})$$

where σ_n^t is the share of output m. Output in these data is based on the deflated added-value of production. In this sense, these output data are aggregate data. So in applying the Törnquist formula above, we let $m = 1$ and thus, $\sigma_m^t = \sigma_m^{t-1} = 1$. y_m^t/y_m^{t-1} is calculated in the same way as x_n^t/x_n^{t-1} described above.

References

Anderson BDO, Deistler M. 1984. Identifiability in dynamic errors-in-variables models. *Journal of Time Series Analysis* **1**: 1–13.

Anderson T. 1976. Estimation of linear fufunction relationships: Approximate distribution and connections with simultaneous equations in econometrics. *Journal of Royal Statistical Society, Series B 38* : 1–36.

BLS. 2015. Superceded historical sic measures for manufacturing sectors and 2-digit sic manufacturing industries, 1949-2001.
URL `http://www.bls.gov/mfp/`

Buonaccorsi JP. 2010. *Measurement error: Models, Methods, and Applications*. Taylor and Francis.

Devereux PJ. 2001. The cyclicality of real wages with employers-employee matches. *Industrial and Labor Relations Review* **54**: 835–850.

Diewert WE, Fox KJ. 2008. On the estimation of returns to scale, technical progress and monopolistic markups. *Journal of Econometrics* **145**: 174 –193. ISSN 0304-4076. The use of econometrics in informing public policy makers.

Durbin J, Koopman SJ. 2012. *Time Series Analysis by State Space Methods*, volume 38. Oxford Statistical science series, 2nd edition.

Fuller WA. 1987. *Measurement Error Models*. Wiley Series in Probability and Mathematical Statistics.

Granger CWJ. 1986. Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics* **48**: 213–228.

Greene W. 2012. *Econometric Analysis*. Pearson, seventh edition.

Harvey AC. 1989. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

Harvey AC. 2011. Modelling the phillips curve with unobserved components. *Applied financial economics* **21**: 7–17.

Harvey AC. 2013. *Dynamic models for volatility and heavy tails with application to Financial and economics time series*. ESM52 - Cambridge University Press.

Hassler U, Kuzin V. 2009. Cointegration analysis under measurement errors. *In Measurement Error: Consequences, Applications and Solutions* : 131–150.

Jin H, Jorgenson DW. 2010. Econometric modeling of technical change. *Journal of Econometrics* **157**: 205–219. ISSN 0304-4076.

Jorgenson D, Ho MS, Samuels JD, Stiroh KJ. 2007. Industry origins of the american productivity resurgence. *Economic Systems Research* **19**: 229–252.

Kapteyn A, Wansbeek T. 1983. Identification in th linear errors in variables model. *Econometrica* **51**: 1847–1849.

Kendall SM, Stuart A. 1977. *The Advanced theory of statistics*, volume 2 Inference and Relationship. Charles Griffin & Company Limited, 4 edition.

Klette T, Griliches Z. 1996. The inconsistency of common scale eestimator when output prices are unobserved and endogenous. *Journal of applied econometrics* **11**: 343–361.

Komumjer I, Ng S. 2014. Measurement errors in dynamic models. *Econometric Theory* **30**: 150–175.

Koopman SJ, Harvey AC, Doornik JA, Shephard N. 1995. *STAMP 5.0 Structural Time Series Analyser, Modeller and Predictor*. Chapman&Hall.

Maravall A, Aigner DJ. 1977. Identification of the dynamic shock-error model : The case of dynamic regression. *Latent Variables in Socio-economic models Editted D. J. Aigner and A. S. Goldberger* .

Marschak J, Andrews J William H. 1944. Random simultaneous equations and the theory of production. *Econometrica* **12**: pp. 143–205.

Nowak E. 1993. The identification of multivariate linear dynamic errors- in- variables models. *Journal of Econometrics* **59**: 213–227.

Ornaghi C. 2006. Assessing the effects of measurement errors on the estimation of production functions. *Journal of Applied Econometrics* **21**: pp. 879–891.

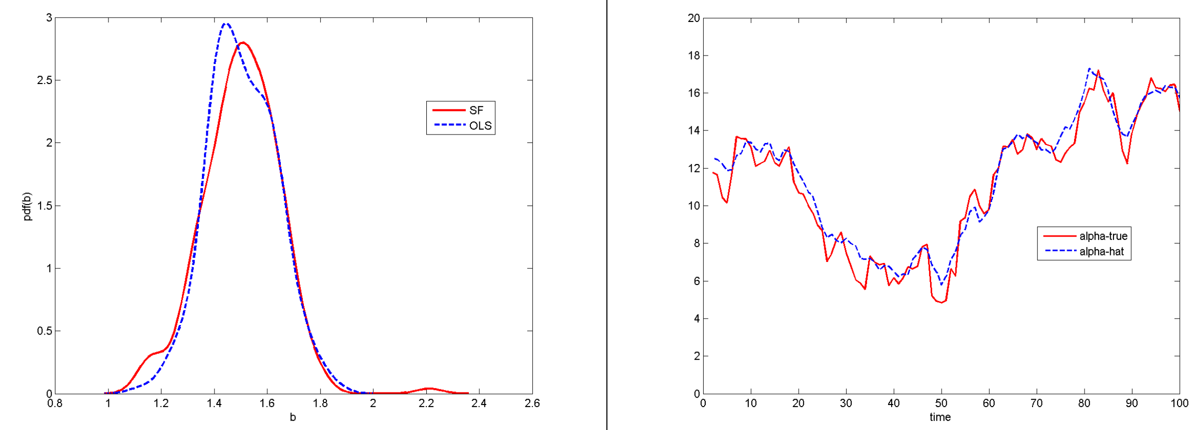
Phillips PC, Durlauf SN. 1986. Multiple time series regression with integrated processes. *Review of Economic Studies* **53**: 473–495.

Reisersol O. 1950. Identifiability of a linear relation between variables which are subject to error. *Econometrica* **18**: 375–389.

Sargent TJ. 1989. Two models of measurements and the investment accelerator. *Journal of Political Economy* **97**: 251–287.

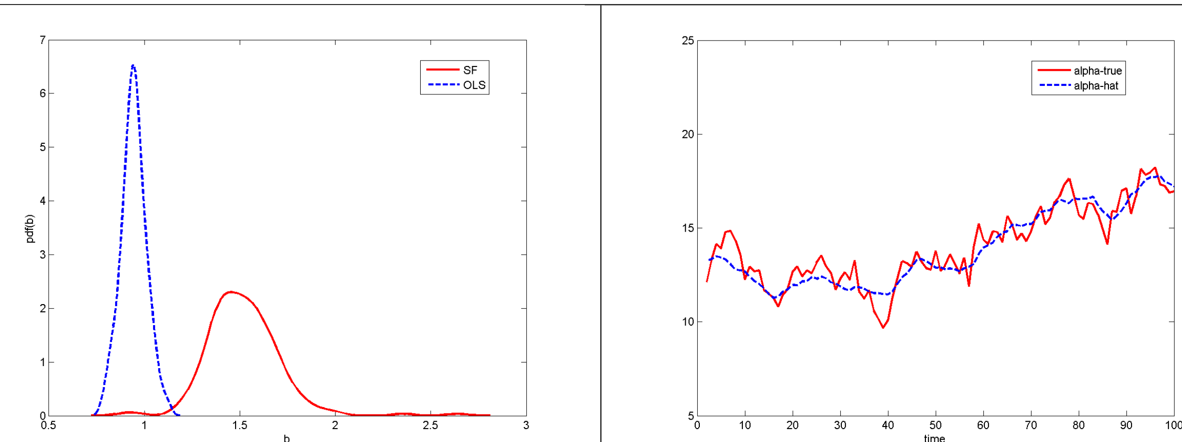
Solo V. 1986. Identifiability of time series models with errors in variables. *Journal of Applied Probability* **23**: 63–71.

Van Beveren I. 2012. Total factor productivity estimation: A practical review. *Journal of Economic Surveys* **26**: 98 – 128. ISSN 09500804.



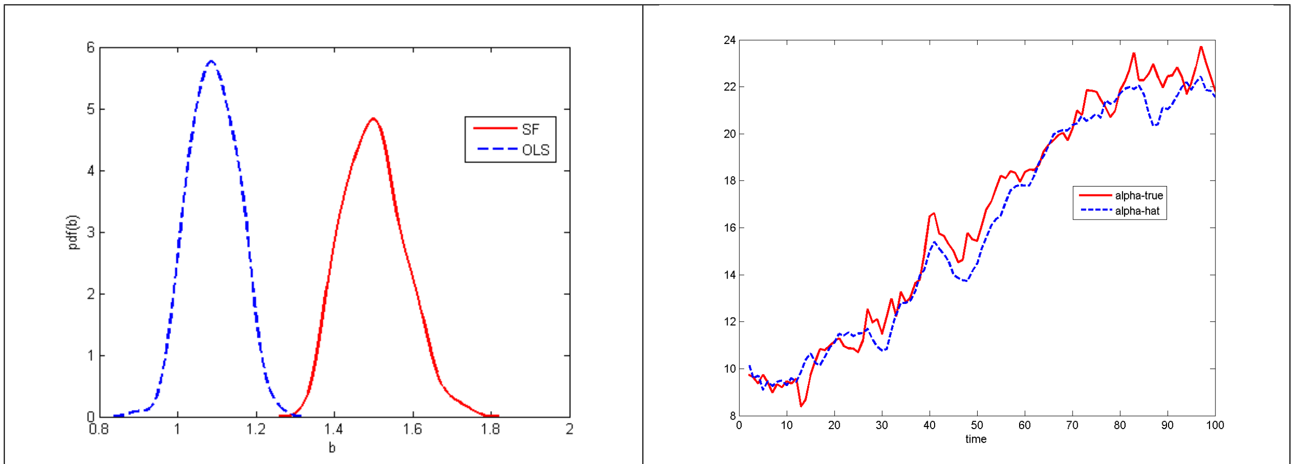
a) Distributions of $\hat{\beta}_{OLS}$ and $\tilde{\beta}$ b) $\tilde{\alpha}_t$ from the 19th sample

Figure 1: Simulation Experiment 1



a) Distributions of $\hat{\beta}_{OLS}$ and $\tilde{\beta}$ b) $\tilde{\alpha}_t$ from the 19th sample

Figure 2: Simulation Experiment 2



a) Distributions of $\hat{\beta}_{OLS}$ and $\tilde{\beta}$ b) $\tilde{\alpha}_t$ from the 1st sample

Figure 3: Simulation Experiment 3

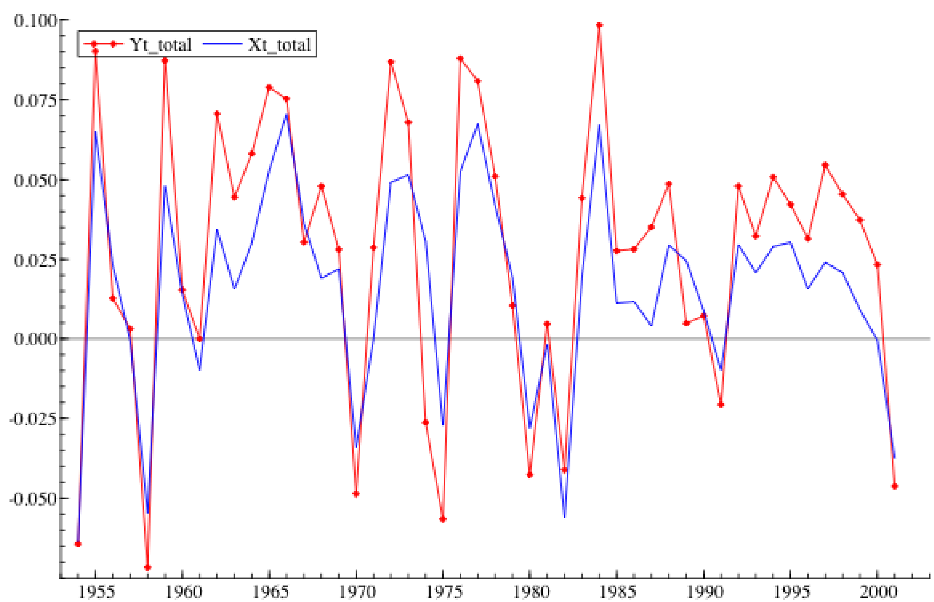


Figure 4: Törnquist indexes of the input and output of the U.S aggregate manufacturing sector, 1954-2001

Table 1: Alternative UC models - Computed Selection Criteria

MODEL	LLT				TC				LC			
EQUATION	Y_t		X_t		Y_t		X_t		Y_t		X_t	
CRITERIA	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
AGGREGATE	-6.1382	-6.0213	-6.7558	-6.6388	-6.1911	-6.0742	-6.8156	-6.6986	**-6.2188**	**-6.1408**	**-6.8739**	**-6.796**
TEXT MILL PROD. (SIC22)	-5.7332	-5.6162	-5.6721	-5.5551	-5.7332	-5.6162	-5.6721	-5.5551	**-5.8856**	**-5.8076**	**-5.8076**	**-5.7296**
APPAREL & RELATED PROD. (SIC23)	-6.0586	-5.9416	-6.1931	-6.0761	**-6.2099**	**-6.0929**	**-6.3034**	**-6.1865**	-6.1283	-6.0504	-6.1975	-6.1196
IND. MACHINERY & COMP. EQ. (SIC35)	-4.7954	-4.6785	-5.1445	-5.0276	-4.8797	-4.7628	-5.2398	-5.1228	**-4.8875**	**-4.8095**	**-5.3171**	**-5.2391**
ELECTRICAL & ELECTR. EQ. (SIC36)	-4.879	-4.7621	-5.3615	-5.2445	-4.9856	-4.8687	-5.4146	-5.2976	**-5.0848**	**-5.0068**	**-5.5371**	**-5.4591**

Table 2: Hyper-parameters of the trend plus cycle model, and the estimates of returns to scale for the aggregate US manufacturing sector and four industries

	AGGREGATE	TEXT MILL PROD. (SIC22)	APPAREL & RELATED PROD. (SIC23)	IND. MACHINERY & COMP. EQ. (SIC35)	ELECTRICAL & ELECTR. EQ. (SIC36)
$\hat{\sigma}^2_{\eta y}$	≈0	3.09E-05	2.04E-06	2.00E-05	9.92E-05
$\hat{\sigma}^2_{\eta x}$	≈0	1.89E-05	3.05E-06	1.13E-05	8.56E-06
$\hat{\rho}_{\eta}$	0	≈1	≈1	≈-1	≈1
$\hat{\sigma}^2_{cy}$	0.0006696	3.75E-11	1.15E-09	0.00435	0.002733
$\hat{\sigma}^2_{cx}$	0.000406	1.55E-11	8.58E-10	0.003425	0.001261
$\hat{\rho}_c$	0.8452	≈1	0.9963	0.993	≈1
$\hat{\sigma}^2_{vy}$	0.000949	0.002364	0.001799	0.001461	0.001114
$\hat{\sigma}^2_{vx}$	0.000426	0.002672	0.001612	0.0002111	0.001701
$\hat{\rho}_v$	≈1	0.9418	0.7286	≈1	≈1
$\tilde{\beta}$	1.085	1.278	0.819	1.111	1.485
$\hat{\beta}_{OLS}$	1.256	0.93	0.815	1.137	1.243
$\hat{\beta}_{D\&F}$	1.248	0.89	0.934	1.147	1.228

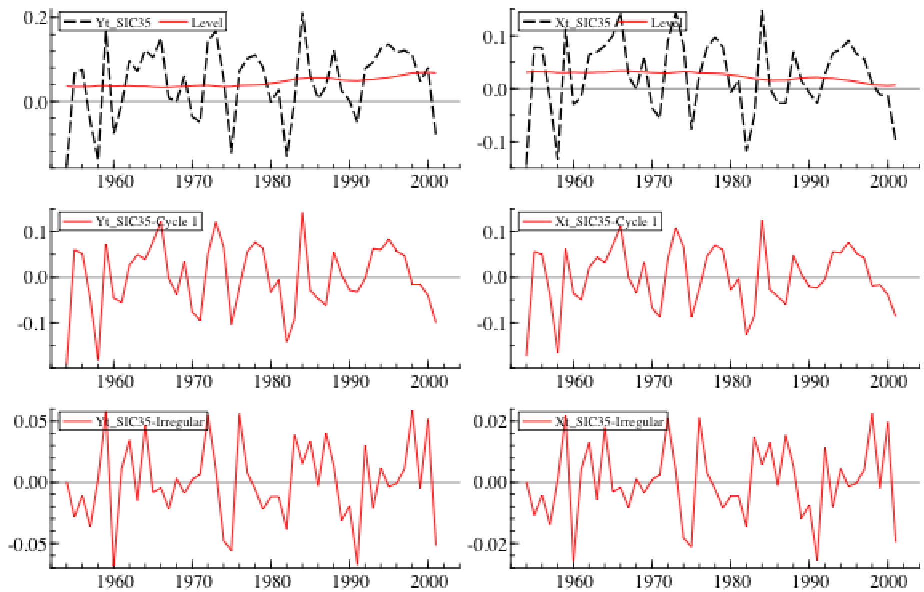


Figure 5: Trends, Cycles and Irregular components - Electric & Electrical Equipment (SIC35) Sector

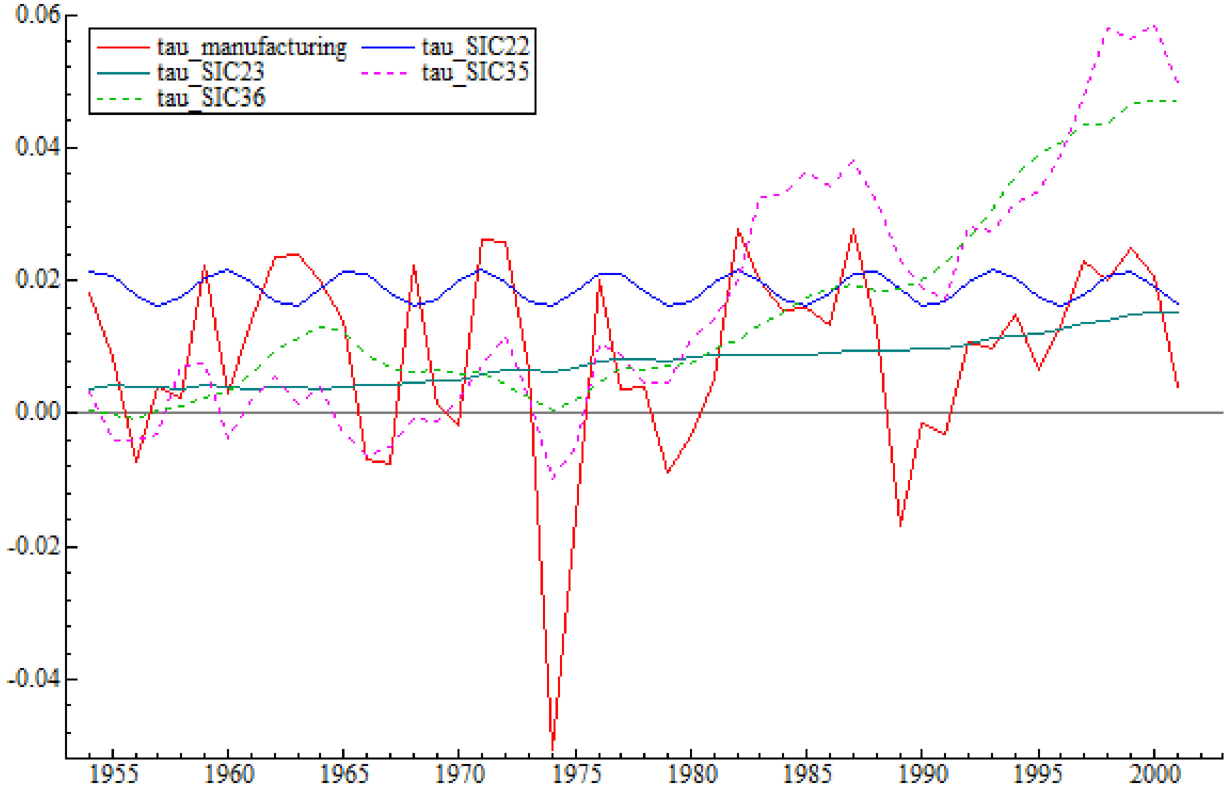


Figure 6: Estimated Technical Change - Selected Sectors, US, 1954-2001