HANDLING THE MEASUREMENT ERROR PROBLEM BY MEANS OF PANEL DATA: MOMENT METHODS APPLIED ON FIRM DATA *)

by

Erik Biørn

University of Oslo, Department of Economics,
P.O. Box 1095 Blindern, 0317 Oslo, Norway.
e-mail: erik.biorn@econ.uio.no
phone: +47 22 85 51 20

February 27, 2002

Abstract

The estimation of a linear equation from panel data with measurement errors is considered. The equation is estimated (I) by methods operating on the equation in differenced period means, and (II) by Generalized Method of Moments (GMM) procedures using (a) the equation in differences with instruments in levels and (b) the equation in levels with instruments in differences. Both difference transformations eliminate unobserved individual heterogeneity. Examples illustrating the input response to output changes for materials and capital inputs from an eight year panel of Norwegian manufacturing firms, are given.

Keywords: Panel Data. Errors-in-Variables. GMM Estimation.
Factor Demand. Returns to scale
JEL classification: C13, C23, D24, L60.

---

*) Paper to appear in: B.P. Stigum (ed.): *Econometrics and the Philosophy of Economics*. Princeton: Princeton University Press.

1 Introduction and basic model

Panel data are a valuable source of information for theory-data confrontation in contemporary econometrics. The panel data available are frequently from individuals, firms, or other kinds of micro units. A primary reason for the strongly increasing utilization of panel data during the last three decades seems to be the opportunity which such data offer for 'controlling for' unobserved individual and/or time specific heterogeneity which may be correlated with the included explanatory variables. As is well known, the effect of individual heterogeneity in a panel data set relative to a linear equation can be removed by measuring all variables from their individual means or by operating on suitably differenced data.

Micro data, including panel data, and inferences drawn from such data may, however, have deficiencies following from measurement errors. Not only observation errors in the narrow sense, but also departures between theoretical variable definitions and their observable counterparts in a wider sense may be present. A familiar property of the Ordinary Least Squares (OLS) in the presence of random measurement errors (errors-in-variables, EIV) in the regressors is that the slope coefficient estimator is inconsistent. In the one regressor case (or the multiple regressor case with uncorrelated regressors), under standard assumptions, the estimator is biased towards zero, often denoted as *attenuation*. More seriously, unless some 'extraneous' information is available, *e.g.*, the existence of valid parameter restrictions or valid instruments for the error-ridden regressors, slope coefficients cannot in general be identified from standard data [see Fuller (1987, section 1.1.3)].[1] This *lack of identification* in EIV models, however, relates to *uni-dimensional* data, *i.e.*, pure (single or repeated) cross-sections or pure time-series. If the variables are observed as panel data, exhibiting *two-dimensional* variation, it may be possible to handle jointly the heterogeneity problem and the EIV identification problem and estimate slope coefficients consistently and efficiently without extraneous information, provided that the distribution of the latent regressors and the measurement errors satisfy certain weak conditions.

Briefly, the reason why the existence of variables observed along two dimensions makes the EIV identification problem more manageable, is partly (i) the *repeated measurement* property of panel data – each individual and each period is 'replicated' – so that the effect of measurement errors can be reduced by taking averages, which, in turn, may show sufficient variation to permit consistent estimation, and partly (ii) the larger set of *other linear data transformations* available for estimation. Such transformations may be needed to compensate for uni-dimensional 'nuisance variables' like unobserved individual or period specific *heterogeneity*, which are potentially correlated with the regressor.

From the panel data literature disregarding the EIV problem we know that the effect of, say, additive (fixed or random) individual heterogeneity within a linear model can be eliminated by deducting individual means, taking differences over periods, etc. [see Hsiao (1986, Section 1.1) and Baltagi (2001, Chapter 2)]. Such transformations, however, may magnify the variation in the measurement error component of the observations relative to the variation in the true structural component, *i.e.*, they may increase the 'noise/signal ratio'. Data transformations intended to 'solve' the latent heterogeneity problem may then aggravate the EIV problem. Several familiar estimators for panel data models, including the fixed effects within-group and between-group estimators, and the random effects Generalized Least Squares (GLS) estimators will then be inconsistent, although to a degree depending, *inter alia*, on the way in which the number of individuals and/or periods tend to infinity and on the heterogeneity of the measurement error process. See Griliches and Hausman (1986) and Biørn (1992, 1996) for examples for one regressor models.

If the distribution of the latent regressor vector is *not* time invariant and the second order moments of the measurement errors and disturbances are structured to some extent, several consistent instrumental variables estimators of the coefficient of the latent regressor vector exist. Their consistency is robust to correlation between the individual heterogeneity and the latent regressor. Serial correlation or non-stationarity of the latent regressor is favourable from the point of view of identification and estimability of the coefficient vector. Briefly, there should not be 'too much structure' on the second order moments of the latent exogenous regressors across the panel, and not 'too little structure' on the second order moments of the errors and disturbances; see Biørn (2000, section 2.b).

The focus of this paper is on the estimation of linear, static regression equations from balanced panel data with additive, random measurement errors in the regressors by means of methods utilizing instrumental variables (IV's). We consider a data set with N (≥ 2) individuals observed in T (≥ 2) periods and a relationship between y (observable scalar) and a $(1 \times K)$ vector $\boldsymbol{\xi}$ (latent),

$$(1) \qquad y_{it} = c + \boldsymbol{\xi}_{it}\boldsymbol{\beta} + \alpha_i + u_{it}, \qquad i = 1,\ldots,N;\ t = 1,\ldots,T,$$

where $(y_{it}, \boldsymbol{\xi}_{it})$ is the value of $(y, \boldsymbol{\xi})$ for individual i in period t, c is a scalar constant, $\boldsymbol{\beta}$ is a $(K \times 1)$ vector and α_i is a zero (marginal) mean individual effect, which we consider as random and potentially correlated with $\boldsymbol{\xi}_{it}$, and u_{it} is a zero mean disturbance, which may also contain a measurement error in y_{it}. We observe

$$(2) \qquad \boldsymbol{x}_{it} = \boldsymbol{\xi}_{it} + \boldsymbol{v}_{it}, \qquad i = 1,\ldots,N;\ t = 1,\ldots,T,$$

where $\boldsymbol{v}_{it}$ is a zero mean vector of measurement errors. Hence,

$$(3) \qquad y_{it} = c + \boldsymbol{x}_{it}\boldsymbol{\beta} + \epsilon_{it}, \quad \epsilon_{it} = \alpha_i + u_{it} - \boldsymbol{v}_{it}\boldsymbol{\beta},$$

or in vector form,

$$(4) \quad \boldsymbol{y}_{i\cdot} = \boldsymbol{e}_T c + \boldsymbol{X}_{i\cdot}\boldsymbol{\beta} + \boldsymbol{\epsilon}_{i\cdot}, \quad \boldsymbol{\epsilon}_{i\cdot} = \boldsymbol{e}_T\alpha_i + \boldsymbol{u}_{i\cdot} - \boldsymbol{V}_{i\cdot}\boldsymbol{\beta}, \qquad i = 1,\ldots,N,$$

where $\boldsymbol{y}_{i\cdot} = (y_{i1},\ldots,y_{iT})'$, $\boldsymbol{X}_{i\cdot} = (\boldsymbol{x}_{i1}',\ldots,\boldsymbol{x}_{iT}')'$, etc., and $\boldsymbol{e}_T$ is the $(T \times 1)$ vector of ones. We denote ϵ_{it} as a composite error/disturbance. We assume that $(\boldsymbol{\xi}_{it}, u_{it}, \boldsymbol{v}_{it}, \alpha_i)$ are independent across individuals [which excludes random period specific components in $(\boldsymbol{\xi}_{it}, u_{it}, \boldsymbol{v}_{it})$], and make the following basic *basic orthogonality assumptions*, corresponding to those in the EIV literature for standard situations:[2]

$$\text{ASSUMPTION (A):} \qquad \begin{cases} \mathsf{E}(\boldsymbol{v}_{it}'u_{i\theta}) = \mathbf{0}_{K1}, \\ \mathsf{E}(\boldsymbol{\xi}_{it}'u_{i\theta}) = \mathbf{0}_{K1}, \\ \mathsf{E}(\boldsymbol{\xi}_{i\theta}' \otimes \boldsymbol{v}_{it}) = \mathbf{0}_{KK}, \\ \mathsf{E}(\alpha_i \boldsymbol{v}_{it}) = \mathbf{0}_{1K}, \\ \mathsf{E}(\alpha_i u_{it}) = 0, \end{cases} \qquad \begin{array}{c} i = 1, \ldots, N, \\ t, \theta = 1, \ldots, T, \end{array}$$

where $\mathbf{0}_{mn}$ denotes the $(m \times n)$ zero matrix and $\otimes$ is the Kronecker product operator.

We can eliminate α_i from (3) by taking arbitrary backward differences $\Delta y_{it\theta} = y_{it} - y_{i\theta} = \boldsymbol{d}_{t\theta}\boldsymbol{y}_{i\cdot}$, $\Delta \boldsymbol{x}_{it\theta} = \boldsymbol{x}_{it} - \boldsymbol{x}_{i\theta} = \boldsymbol{d}_{t\theta}\boldsymbol{X}_{i\cdot}$, etc., where $\boldsymbol{d}_{t\theta}$ is the $(1 \times T)$ vector with element t equal to 1, element θ equal to -1 and zero otherwise. Premultiplying (4) by $\boldsymbol{d}_{t\theta}$, we get[3]

$$\Delta y_{it\theta} = \Delta \boldsymbol{x}_{it\theta}\,\boldsymbol{\beta} + \Delta \epsilon_{it\theta}, \qquad t = 2, \ldots, T; \theta = 1, \ldots, t-1. \tag{5}$$

In the next section, we present an interpretation of this model framework based on production theory and a panel of manufacturing firms. In the following sections, we describe the estimation methods and introduce the additional assumptions needed. Two kinds of estimation methods will be in focus: (i) Methods operating on period means, illustrating applications of the repeated measurement property of panel data (Section 3), and (ii) Generalized Method of Moments (GMM) procedures (Sections 4 – 6). The GMM procedures involve *a mixture of level and difference variables* and are of two kinds: (a) The equation is transformed to differences, as in (5), and is estimated by GMM, and as instruments we use level values of the regressors and/or regressands for other periods. (b) The equation is kept in level form, as in (3), and is estimated by GMM, and as instruments we use differenced values of the regressors and/or regressands for other periods. Our (a) procedures extend and modify procedures proposed in Griliches and Hausman (1986), Wansbeek and Koning (1991), Arellano and Bover (1995), Biørn (1992, 1996), Biørn and Klette (1998, 1999), and Wansbeek (2001).

2 Application: Input elasticities in manufacturing

We next present a simple interpretation of (1) with a single regressor ($K = 1$), to be used as basis for our empirical applications. The data are from eight successive annual Norwegian manufacturing censuses for the years 1983 – 1990 ($T = 8$), collected by Statistics Norway, for four two-digit sectors, comprising 1647 firms (plants): *Manufacture of Textiles (ISIC 32)* ($N = 215$), *Manufacture of Wood and Wood Products (ISIC 33)* ($N = 603$), *Manufacture of Paper and Paper Products (ISIC 34)* ($N = 600$), and *Manufacture of Chemicals (ISIC 35)* ($N = 229$). The data base specifies labour, capital, and materials (including energy) inputs, but for our illustrative purposes and in order not to inflate our tables of results, we confine attention on the two latter. This pair of inputs is interesting since capital raises much heavier measurement problems than materials, although both inputs and the output contain potential measurement errors, both in the strict and wide sense. Our measure of capital input is based on deflated fire insurance values, which is a wealth related measure and hence contain potential errors as indicators of the productive capacity of the capital.

Let us, with reference to production theory, describe two alternative interpretations of the model (1) – (5). We do not go deeply into the problems of theory-data confrontation and refer to Stigum (1995) for a thorough discussion.

The first and simplest interpretation is to assume a technology with one output X^*_{it} and *one input* Y^*_{it}, *i.e.*, either capital or materials, both latent. Firm specific differences in technology are represented by the factor e^{ϕ_i}, indicating firm i's departure from the technology of the average firm (characterized by $\phi_i = 0$). We specify the technology as

$$X^*_{it} = e^{\phi_i} F(Y^*_{it}) = A e^{\phi_i} (Y^*_{it})^{\mu}, \tag{6}$$

where A is a positive constant, μ is the scale elasticity for this one factor Cobb-Douglas technology, and $\mathsf{E}(\phi_i) = 0$. We can allow for an unspecified period specific effect by extending (6) to

$$X_{it}^* = e^{\phi_i} F_t(Y_{it}^*) = A e^{\phi_i} e^{\psi_t} (Y_{it}^*)^{\mu}, \tag{7}$$

where $\mathsf{E}(\phi_i) = \mathsf{E}(\psi_t) = 0$.

The second interpretation is to assume a (neo-classical) technology with one output and *several inputs*, of which two are capital and materials, and output constrained *cost minimization*. Let X_{it}^* denote output, $Y_{it}^* = (Y_{it}^{*1}, \ldots, Y_{it}^{*G})$ the vector of G inputs, and $w_t^* = (w_t^{*1}, \ldots, w_t^{*G})$ the vector of input prices, common to all firms – all treated as latent variables. We describe the technology by

$$X_{it}^* = e^{\phi_i} F_t(Y_{it}^*), \tag{8}$$

where F_t is a production function common to all firms, t reflecting that technological changes are allowed for. We interpret ϕ_i as a constant known to firm i, but unobserved by the econometrician. The dual cost function can then be written as $C_{it}^* = G_t(w_t^*, e^{-\phi_i} X_{it}^*)$ [cf., *e.g.*, Jorgenson (1986, section 5)], where $C_{it}^* = \sum_{k=1}^{G} w_t^{*k} Y_{it}^{*k}$. Using Shephard's lemma, we can express firm i's optimal input of factor k in year t as

$$Y_{it}^{*k} = g_t^k(w_t^*, e^{-\phi_i} X_{it}^*), \tag{9}$$

where $g_t^k(\cdot) = \partial G_t(\cdot)/\partial w_t^{*k}$. Assuming that F_t represents a *homothetic* technology, so that G_t can be separated as $G_t(w_t^*, e^{-\phi_i} X_{it}^*) = H_t(w_t^*) K(e^{-\phi_i} X_{it}^*)$, where the functions H_t and K are monotonically increasing, (9) becomes

$$Y_{it}^{*k} = h_t^k(w_t^*)\, K(e^{-\phi_i} X_{it}^*),$$

with $h_t^k(\cdot) = \partial H_t(\cdot)/\partial w_t^{*k}$. If, in particular, (8) has a *constant scale elasticity* μ for all firms and years, then $K(e^{-\phi_i} X_{it}^*) = e^{-\phi_i/\mu} (X_{it}^*)^{1/\mu}$ for all i, t and hence

$$Y_{it}^{*k} = h_t^k(w_t^*) e^{-\phi_i/\mu} (X_{it}^*)^{1/\mu}. \tag{10}$$

Taking logs, we can then write both (7) and (9) in simplified notation as

$$\chi_{it} = c + \alpha_i + \gamma_t + \beta \xi_{it}, \tag{11}$$

where $\chi_{it} = \ln Y_{it}^*$, $\xi_{it} = \ln X_{it}^*$, $\beta = 1/\mu$, $\alpha_i = -(1/\mu)\phi_i$, γ_t is either zero, $-(1/\mu)\psi_t$ or $\ln h_t^k(w_t^*)$, and c is a constant. The observed log-output and log-input are $y_{it} = \chi_{it} + u_{it}$ and $x_{it} = \xi_{it} + v_{it}$, where u_{it} and v_{it} are measurement errors. This gives an equation of the form (3). In the more general case where F_t represents a non-homothetic technology, separability of G_t does not hold. Then the input elasticity $\boldsymbol{\beta}$ will be different for different inputs and hence cannot be interpreted as an inverse scale elasticity.

Neither of these model interpretations imposes a specific normalization on (11) and (3), as observed input and output are both formally endogenous variables. In the empirical application, two normalizations will be considered: (i) y_{it} and x_{it} are, respectively, the log of an observed factor input and the log of observed gross production, both measured as values at constant prices and β corresponds to $1/\mu$, and (ii) y_{it} and x_{it} have the reverse interpretation and β corresponds to μ.

3 Estimators based on period means

In this section, we consider various estimators of $\boldsymbol{\beta}$ constructed from differenced period means. From (3) we obtain

$$(12)\qquad \Delta_s \bar{y}_{\cdot t} = \Delta_s \bar{\boldsymbol{x}}_{\cdot t}\boldsymbol{\beta} + \Delta_s \bar{\epsilon}_{\cdot t}, \qquad s = 1, \dots, T-1;\ t = s+1, \dots, T,$$

$$(13)\qquad (\bar{y}_{\cdot t} - \bar{y}) = (\bar{\boldsymbol{x}}_{\cdot t} - \bar{\boldsymbol{x}})\boldsymbol{\beta} + (\bar{\epsilon}_{\cdot t} - \bar{\epsilon}), \qquad t = 1, \dots, T,$$

where $\bar{y}_{\cdot t} = \sum_i y_{it}/N$, $\bar{y} = \sum_i \sum_t y_{it}/(NT)$, $\bar{\boldsymbol{x}}_{\cdot t} = \sum_i \boldsymbol{x}_{it}/N$, $\bar{\boldsymbol{x}} = \sum_i \sum_t \boldsymbol{x}_{it}/(NT)$, etc. and Δ_s denotes differencing over s periods.

The (weak) law of the large numbers, when (A) is satisfied, implies under weak conditions [cf. McCabe and Tremayne (1993, section 3.5)],[4] that $\text{plim}(\bar{\epsilon}_{\cdot t}) = 0$, $\text{plim}(\bar{\boldsymbol{x}}_{\cdot t} - \bar{\boldsymbol{\xi}}_{\cdot t}) = \mathbf{0}_{1K}$, so that $\text{plim}[\bar{\boldsymbol{x}}_{\cdot t}'\bar{\epsilon}_{\cdot t}] = \mathbf{0}_{K1}$ even if $\text{plim}[(1/N)\sum_{i=1}^{N} \boldsymbol{x}_{it}'\epsilon_{it}] \neq \mathbf{0}_{K1}$. From (12) and (13) we therefore get

$$(14)\qquad \text{plim}[(\Delta_s \bar{\boldsymbol{x}}_{\cdot t})'(\Delta_s \bar{y}_{\cdot t})] = \text{plim}[(\Delta_s \bar{\boldsymbol{x}}_{\cdot t})'(\Delta_s \bar{\boldsymbol{x}}_{\cdot t})]\boldsymbol{\beta},$$

$$(15)\qquad \text{plim}[(\bar{\boldsymbol{x}}_{\cdot t} - \bar{\boldsymbol{x}})'(\bar{y}_{\cdot t} - \bar{y})] = \text{plim}[(\bar{\boldsymbol{x}}_{\cdot t} - \bar{\boldsymbol{x}})'(\bar{\boldsymbol{x}}_{\cdot t} - \bar{\boldsymbol{x}})]\boldsymbol{\beta}.$$

Hence, provided that $\mathsf{E}[(\Delta_s\bar{\boldsymbol{\xi}}_{\cdot t})'(\Delta_s\bar{\boldsymbol{\xi}}_{\cdot t})]$ and $\mathsf{E}[(\bar{\boldsymbol{\xi}}_{\cdot t} - \bar{\boldsymbol{\xi}})'(\bar{\boldsymbol{\xi}}_{\cdot t} - \bar{\boldsymbol{\xi}})]$ have rank K, consistent estimators of $\boldsymbol{\beta}$ can be obtained by applying OLS on (12) or on (13), which give, respectively,

$$(16)\ \widehat{\boldsymbol{\beta}}_{\Delta s} = \left[\sum_{t=s+1}^{T} (\Delta_s\bar{\boldsymbol{x}}_{\cdot t})'(\Delta_s\bar{\boldsymbol{x}}_{\cdot t})\right]^{-1}\left[\sum_{t=s+1}^{T} (\Delta_s\bar{\boldsymbol{x}}_{\cdot t})'(\Delta_s\bar{y}_{\cdot t})\right], \quad s = 1, \ldots, T-1,$$

$$(17)\ \widehat{\boldsymbol{\beta}}_{BP} = \left[\sum_{t=1}^{T}(\bar{\boldsymbol{x}}_{\cdot t} - \bar{\boldsymbol{x}})'(\bar{\boldsymbol{x}}_{\cdot t} - \bar{\boldsymbol{x}})\right]^{-1}\left[\sum_{t=1}^{T}(\bar{\boldsymbol{x}}_{\cdot t} - \bar{\boldsymbol{x}})'(\bar{y}_{\cdot t} - \bar{y})\right].$$

The latter is the 'between period' (BP) estimator. The consistency of these estimators simply relies on the fact that averages of a large number of repeated measurements of an error-ridden variable give, under weak conditions, an error-free measure of the true average at the limit, *provided that this average shows variation along the remaining dimension, i.e., across periods*. Basic to these conclusions is the assumption that *the measurement error has no period specific component.* If such a component is present, it will not vanish when taking plims of period means, *i.e.*, plim($\bar{\boldsymbol{v}}_{\cdot t}$) will no longer be zero, (14) and (15) will no longer hold, and so $\widehat{\boldsymbol{\beta}}_{\Delta s}$ and $\widehat{\boldsymbol{\beta}}_{BP}$ will be inconsistent.

Table 24.1 reports between period estimates of $\boldsymbol{\beta}$ based on levels (column 2) and on differences (column 5) – the latter removing the effect of technical changes represented by a log-linear trend – as well as seven-period difference estimates (column 7) for the four sectors and the two inputs.[5] Since $K = 1$ in this application, the estimators read

$$\widehat{\boldsymbol{\beta}}_{BP} = \frac{\sum_t(\bar{x}_{\cdot t} - \bar{x})(\bar{y}_{\cdot t} - \bar{y})}{\sum_t(\bar{x}_{\cdot t} - \bar{x})^2},$$

$$\widehat{\boldsymbol{\beta}}_{BPDC} = \frac{\sum_t(\Delta\bar{x}_{\cdot t} - \Delta\bar{x})(\Delta\bar{y}_{\cdot t} - \Delta\bar{y})}{\sum_t(\Delta\bar{x}_{\cdot t} - \Delta\bar{x})^2},$$

$$\widehat{\boldsymbol{\beta}}_{\Delta 7} = \frac{\bar{y}_{\cdot 8} - \bar{y}_{\cdot 1}}{\bar{x}_{\cdot 8} - \bar{x}_{\cdot 1}}.$$

Rows 1 and 3 can be interpreted as estimates of $1/\mu$ and rows 2 and 4 as estimates of μ. This way of running original and reverse regressions in an EIV context can be related to Frisch's *confluence analysis* [Frisch (1934, sections 5, 10, 11, and 14)], in which he proposed taking regressions in different directions, *e.g.*, in the directions of the 'x axis' and of the 'y axis' as a device for handling measurement errors. He did not, however, consider this method in a panel data context.

For *materials*, the between period (BP) estimates on levels for the original and the reverse regression imply virtually the same input elasticity, $1/\mu$, in the range 1.00 – 1.09 for the four sectors considered. They are also very close to the estimates obtained from seven-period differences. The BP estimates based on differences, $\widehat{\boldsymbol{\beta}}_{BPDC}$, show somewhat larger discrepancies. For *capital*, there are substantial deviations between the level BP, the difference BP, and the seven-period difference estimates. For the BP estimators on levels, the reverse regression gives systematically higher estimates of the input elasticity of capital (lower estimates of μ) than the original regressions. This may indicate that the measurement errors in capital have period specific, or strongly serially correlated, components, which make both the between period and all period difference estimators inconsistent. For capital, unlike materials, the results also suggest the presence of period specific heterogeneity in the relationship.

OLS estimates calculated from levels and from one period differences,

$$\widehat{\boldsymbol{\beta}}_{OLS} = \frac{\sum_i \sum_t (x_{it} - \bar{x})(y_{it} - \bar{y})}{\sum_i \sum_t (x_{it} - \bar{x})^2},$$
$$\widehat{\boldsymbol{\beta}}_{OLSDC} = \frac{\sum_i \sum_t (\Delta x_{it} - \Delta \bar{x})(\Delta y_{it} - \Delta \bar{y})}{\sum_i \sum_t (\Delta x_{it} - \Delta \bar{x})^2},$$
$$\widehat{\boldsymbol{\beta}}_{OLSD} = \frac{\sum_i \sum_t (\Delta x_{it})(\Delta y_{it})}{\sum_i \sum_t (\Delta x_{it})^2},$$

are also reported (columns 1, 4, and 8), $\widehat{\boldsymbol{\beta}}_{OLSDC}$ removing the possible effect of linear trends. Columns 3 and 6 contain within firm (WF) estimates calculated from levels and differences,

$$\widehat{\boldsymbol{\beta}}_{WF} = \frac{\sum_i \sum_t (x_{it} - \bar{x}_{i\cdot})(y_{it} - \bar{y}_{i\cdot})}{\sum_i \sum_t (x_{it} - \bar{x}_{i\cdot})^2},$$
$$\widehat{\boldsymbol{\beta}}_{WFDC} = \frac{\sum_i \sum_t (\Delta x_{it} - \Delta \bar{x}_{i\cdot})(\Delta y_{it} - \Delta \bar{y}_{i\cdot})}{\sum_i \sum_t (\Delta x_{it} - \Delta \bar{x}_{i\cdot})^2},$$

$\widehat{\boldsymbol{\beta}}_{WFDC}$ removing the possible effect of linear trends. These three OLS and the two WF estimates, all of which are inconsistent in the presence of measurement errors, clearly illustrate the attenuation effect. They have, however, different degree of robustness. While $\boldsymbol{\beta}_{OLS}$ is neither robust to firm specific heterogeneity nor to trend effects, $\boldsymbol{\beta}_{WF}$ and $\boldsymbol{\beta}_{OLSD}$ are robust to firm specific heterogeneity (which is potentially correlated with the regressand or the regressor), but not robust to trend effects, and $\boldsymbol{\beta}_{OLSDC}$ and $\boldsymbol{\beta}_{WFDC}$ are robust to both firm specific heterogeneity and a linear trend. For materials, unlike capital, $\boldsymbol{\beta}_{OLSD}$, $\boldsymbol{\beta}_{OLSDC}$, and $\boldsymbol{\beta}_{WFDC}$ give fairly equal estimates in all sectors.

Although these examples show that it is possible to construct consistent estimators, which give estimates of reasonable size (at least for materials), from period means, their efficiency may be low, since they do not exploit any inter-individual variation in the data, and the latter often tends to dominate. Therefore there is a potential to improve the estimation by considering methods which utilizes this inter-individual variation. One such method is the GMM.

4 The principle of GMM estimation

Before elaborating the GMM procedures for our panel data situation, we describe some generalities of this procedure, referring to, *e.g*, Davidson and MacKinnon (1993, Chapter 17) and Harris and Mátyás (1999) for more detailed expositions. Assume, in general, that we want to estimate the $(K \times 1)$ coefficient vector $\boldsymbol{\beta}$ in the equation

$$y = \boldsymbol{x}\boldsymbol{\beta} + \epsilon, \tag{18}$$

where y and ϵ are scalars and $\boldsymbol{x}$ is a $(1 \times K)$ regressor vector. There exists an instrument vector $\boldsymbol{z}$, of dimension $(1 \times G)$, for $\boldsymbol{x}$ $(G \geq K)$, satisfying the orthogonality conditions

$$E(\boldsymbol{z}'\epsilon) = E[\boldsymbol{z}'(y - \boldsymbol{x}\boldsymbol{\beta})] = \mathbf{0}_{G,1}. \tag{19}$$

These conditions are assumed to be derived from the economic theory and the statistical auxiliary hypotheses (*e.g.*, about disturbance/error autocorrelation) underlying our model. We have n observations on $(y, \boldsymbol{x}, \boldsymbol{z})$, denoted as $(y_j, \boldsymbol{x}_j, \boldsymbol{z}_j)$, $j =$

$1, \ldots, n$, and define the vector valued $(G \times 1)$ function of corresponding means taken over all available observations,

$$\boldsymbol{g}_n(y, \boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\beta}) = (1/n) \sum_{j=1}^{n} \boldsymbol{z}_j'(y_j - \boldsymbol{x}_j \boldsymbol{\beta}). \tag{20}$$

It may be considered the empirical counterpart to $E[\boldsymbol{z}'(y - \boldsymbol{x}\boldsymbol{\beta})]$ based on the sample. The *essence of GMM* is to choose as an estimator for $\boldsymbol{\beta}$ the value which brings the value of $\boldsymbol{g}_n(y, \boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\beta})$ as close to its theoretical counterpart, the zero vector $\mathbf{0}_{G,1}$, as possible. If $G = K$, an exact solution to the equation $\boldsymbol{g}_n(y, \boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\beta}) = \mathbf{0}_{G,1}$ exists and is the simple IV estimator

$$\boldsymbol{\beta}^* = [\textstyle\sum_j \boldsymbol{z}_j' \boldsymbol{x}_j]^{-1} [\textstyle\sum_j \boldsymbol{z}_j' y_j]. \tag{21}$$

If $G > K$, which is the most common situation, the GMM procedure solves the estimation problem by *minimizing a distance measure represented by a quadratic form in* $g_n(y, \boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\beta})$ *for a suitably chosen positive definit* $(G \times G)$ *weighting matrix* $\boldsymbol{W}_n$, *i.e.*,

$$\boldsymbol{\beta}^*_{GMM} = \boldsymbol{\beta}^*_{GMM}(\boldsymbol{W}_n) = \text{argmin}_{\boldsymbol{\beta}} [\boldsymbol{g}_n(y, \boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\beta})' \boldsymbol{W}_n \boldsymbol{g}_n(y, \boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\beta})]. \tag{22}$$

All estimators obtained in this way are consistent. The choice of $\boldsymbol{W}_n$ determines the efficiency of the method. A choice which leads to an asymptotically efficient estimator of $\boldsymbol{\beta}$, is to set this weighting matrix equal (or proportional) to the inverse of (an estimate of) the (asymptotic) covariance matrix of $(1/n) \sum_{j=1}^{n} \boldsymbol{z}_j' \epsilon_j$; see, *e.g.*, Davidson and MacKinnon (1993, Theorem 17.3) and Harris and Mátyás (1999, section 1.3.3).

If ϵ is serially uncorrelated and homoskedastic, with variance σ_ϵ^2, the appropriate choice is simply $\boldsymbol{W}_n = [n^{-2} \sigma_\epsilon^2 \sum_{j=1}^{n} \boldsymbol{z}_j' \boldsymbol{z}_j]^{-1}$. The resulting estimator obtained from (22) is

$$\widehat{\boldsymbol{\beta}}_{GMM} = [(\textstyle\sum_j \boldsymbol{x}_j' \boldsymbol{z}_j)(\sum_j \boldsymbol{z}_j' \boldsymbol{z}_j)^{-1}(\sum_j \boldsymbol{z}_j' \boldsymbol{x}_j)]^{-1} [(\sum_j \boldsymbol{x}_j' \boldsymbol{z}_j)(\sum_j \boldsymbol{z}_j' \boldsymbol{z}_j)^{-1}(\sum_j \boldsymbol{z}_j' y_j)], \tag{23}$$

which is the standard Two-Stage Least Squares (2SLS) estimator. The method can also be fruitfully applied if ϵ_j has a heteroskedasticity of unspecified (or unknown) form. It can also take account of disturbance/error autocorrelation more or less strictly specified, by reformulating the orthogonality conditions in an appropriate way, as will be examplified below. This flexibility with respect to the imposition of restrictions on the second order moments of disturbances/errors is one of the primary virtues of GMM as compared with classical 2SLS. To operationalize the latter method in the presence of unknown heteroskedasticity, we then first construct consistent residuals $\widehat{\epsilon}_j$, usually from (23), which we consider as a *first step GMM estimator*, and estimate $\boldsymbol{W}_n$ by $\widehat{\boldsymbol{W}}_n = [n^{-2}\sum_j \boldsymbol{z}_j'\widehat{\epsilon}_j^2\boldsymbol{z}_j]^{-1}$; see White (1984, sections IV.3 and VI.2). Inserting this into (22) gives

$$\widetilde{\boldsymbol{\beta}}_{GMM} = [(\textstyle\sum_j \boldsymbol{x}_j'\boldsymbol{z}_j)(\sum_j \boldsymbol{z}_j'\widehat{\epsilon}_j^2\boldsymbol{z}_j)^{-1}(\sum_j \boldsymbol{z}_j'\boldsymbol{x}_j)]^{-1} \times [(\sum_j \boldsymbol{x}_j'\boldsymbol{z}_j)(\sum_j \boldsymbol{z}_j'\widehat{\epsilon}_j^2\boldsymbol{z}_j)^{-1}(\sum_j \boldsymbol{z}_j' y_j)]. \tag{24}$$

The latter, *second step GMM estimator*, is in a sense an optimal GMM estimator in the presence of unspecified error/disturbance heteroskedasticity. Both will be considered in our empirical application below.

5 Simple GMM estimators combining differences and levels

As explained in Section 4, the orthogonality conditions (OC's) derived from economic theory, (19), their empirical counterparts (20), and other restrictions imposed on second order moments of observed variables and errors and disturbances play an essential rôle in GMM procedures. We have already made Assumption (A). Before presenting the specific estimators for our panel data measurement error situation, we state the additional assumptions we will need.

5.a Additional assumptions

Our additional assumptions with respect to the errors and disturbances are the *non-autocorrelation* assumptions:

Assumption (B1):	$\mathsf{E}(\boldsymbol{v}_{it}'\boldsymbol{v}_{i\theta}) = \mathbf{0}_{KK}$,	$t \neq \theta$,
Assumption (C1):	$\mathsf{E}(u_{it}u_{i\theta}) = 0$,	$t \neq \theta$.

Sometimes, the following weaker assumptions, allowing for some autocorrelation, will be sufficient:

Assumption (B2):	$\mathsf{E}(\boldsymbol{v}_{it}'\boldsymbol{v}_{i\theta}) = \mathbf{0}_{KK}$,	$\lvert t-\theta \rvert > \tau$,
Assumption (B3):	$\mathsf{E}(\boldsymbol{v}_{it}'\boldsymbol{v}_{i\theta})$ is invariant to t, θ,	$t \neq \theta$,
Assumption (C2):	$\mathsf{E}(u_{it}u_{i\theta}) = 0$,	$\lvert t-\theta \rvert > \tau$,
Assumption (C3):	$\mathsf{E}(u_{it}u_{i\theta})$ is invariant to t, θ,	$t \neq \theta$,

of which (B2) and (C2) allow for a (vector) moving average (MA) structure up to order τ (≥ 1), and (B3) and (C3) allow for time invariance of the autocorrelation. The latter will, for example, be satisfied if the measurement errors and the disturbances have individual components, say $\boldsymbol{v}_{it} = \boldsymbol{v}_{1i} + \boldsymbol{v}_{2it}$, $u_{it} = u_{1i} + u_{2it}$, where $\boldsymbol{v}_{1i}$, $\boldsymbol{v}_{2it}$, u_{1i}, and u_{2it} are independent IID processes.

Our additional assumptions with respect to the *distribution of the latent regressor vector* $\boldsymbol{\xi}_{it}$ are:

Assumption (D1):	$\mathsf{E}(\boldsymbol{\xi}_{it})$	is invariant to t,
Assumption (D2):	$\mathsf{E}(\alpha_i \boldsymbol{\xi}_{it})$	is invariant to t,
Assumption (E):	$\text{rank}(\mathsf{E}[\boldsymbol{\xi}_{ip}'(\Delta\boldsymbol{\xi}_{it\theta})]) = K$	for some p, t, θ different.

Assumptions (D1) and (D2) hold when $\boldsymbol{\xi}_{it}$ is mean stationary for all i. Assumption (E) *imposes non-*IID *and some form of autocorrelation or (covariance) non-stationarity on* $\boldsymbol{\xi}_{it}$. It excludes, for example, the case where $\boldsymbol{\xi}_{it}$ has an individual component, so that $\boldsymbol{\xi}_{it} = \boldsymbol{\xi}_{1i} + \boldsymbol{\xi}_{2it}$, where $\boldsymbol{\xi}_{1i}$ and $\boldsymbol{\xi}_{2it}$ are independent (vector) IID processes.

Assumptions (A) – (E) do not go very far in structuring the distributions of the variables of the model. This has both its pros and cons. It may be possible to impose more structure on the first and second order moments of the u_{it}'s, $\boldsymbol{v}_{it}$'s, $\boldsymbol{\xi}_{it}$'s, and α_i's – confer the 'structural approach' to EIV modelling. In this way we might obtain more efficient estimators by operating on the full covariance matrix of the y_{it}'s and the $\boldsymbol{x}_{it}$'s, and possibly higher order moments, rather than eliminating the α_i's by differencing, as elaborated below. Such estimators, however, may be less robust to specification errors.

5.b Moment equations and orthogonality conditions

A substantial number of moment conditions involving second order moments in y_{it}, $\boldsymbol{x}_{it}$, and ϵ_{it} can be derived from Assumptions (A) – (E).

From (1) – (3) and Assumption (A) we obtain the following moment equations involving observable variables in levels and differences:

$$\text{(25)} \qquad \mathsf{E}[\boldsymbol{x}_{ip}'(\Delta\boldsymbol{x}_{it\theta})] = \mathsf{E}[\boldsymbol{\xi}_{ip}'(\Delta\boldsymbol{\xi}_{it\theta})] + \mathsf{E}[\boldsymbol{v}_{ip}'(\Delta\boldsymbol{v}_{it\theta})],$$

$$\text{(26)} \qquad \mathsf{E}[\boldsymbol{x}_{ip}'(\Delta y_{it\theta})] = \mathsf{E}[\boldsymbol{\xi}_{ip}'(\Delta\boldsymbol{\xi}_{it\theta})]\boldsymbol{\beta},$$

$$\text{(27)} \qquad \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'y_{it}] = \mathsf{E}[(\Delta\boldsymbol{\xi}_{ipq})'\boldsymbol{\xi}_{it}]\boldsymbol{\beta} + \mathsf{E}[(\Delta\boldsymbol{\xi}_{ipq})'(\alpha_i + c)]$$

and involving observable variables and errors/disturbances:

$$\text{(28)} \qquad \mathsf{E}[\boldsymbol{x}_{ip}'(\Delta\epsilon_{it\theta})] = -\mathsf{E}[\boldsymbol{v}_{ip}'(\Delta\boldsymbol{v}_{it\theta})]\boldsymbol{\beta},$$

$$\text{(29)} \qquad \mathsf{E}[y_{ip}(\Delta\epsilon_{it\theta})] = \mathsf{E}[u_{ip}(\Delta u_{it\theta})],$$

$$\text{(30)} \qquad \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'\epsilon_{it}] = \mathsf{E}[(\Delta\boldsymbol{\xi}_{ipq})'\alpha_i] - \mathsf{E}[(\Delta\boldsymbol{v}_{ipq})'\boldsymbol{v}_{it}]\boldsymbol{\beta},$$

$$\text{(31)} \qquad \mathsf{E}[(\Delta y_{ipq})\epsilon_{it}] = \boldsymbol{\beta}'\mathsf{E}[(\Delta\boldsymbol{\xi}_{ipq})'\alpha_i] + \mathsf{E}[(\Delta u_{ipq})u_{it}], \qquad t,\theta,p,q = 1,\dots,T.$$

The moments on the left hand side of (25) – (27) are structured by Assumptions (D) and (E). The moments at the left hand side of (28) – (31) are structured by Assumptions (B) – (D). Depending on which assumptions are valid, some of the terms on the right hand side of (28) – (31), or all, vanish. Provided that $T > 2$, (3), (5), and (28) – (31) imply

$$\text{(32)} \qquad \begin{cases} \textit{When either (B1) holds and } t,\ \theta,\ p \textit{ are different,} \\ \textit{or (B2) holds and } |t-p|, |\theta-p| > \tau, \textit{ then} \\ \mathsf{E}[\boldsymbol{x}_{ip}'(\Delta\epsilon_{it\theta})] = \mathsf{E}[\boldsymbol{x}_{ip}'(\Delta y_{it\theta})] - \mathsf{E}[\boldsymbol{x}_{ip}'(\Delta\boldsymbol{x}_{it\theta})]\boldsymbol{\beta} = \mathbf{0}_{K1}. \end{cases}$$

$$(33) \quad \begin{cases} \textit{When either (C1) holds and } t,\, \theta,\, p \textit{ are different,} \\ \textit{or (C2) holds and } |t-p|, |\theta-p| > \tau, \textit{ then} \\ \mathsf{E}[y_{ip}(\Delta\epsilon_{it\theta})] = \mathsf{E}[y_{ip}(\Delta y_{it\theta})] - \mathsf{E}[y_{ip}(\Delta \boldsymbol{x}_{it\theta})]\boldsymbol{\beta} = 0. \end{cases}$$

$$(34) \quad \begin{cases} \textit{When either (B1), (D1), and (D2) hold and } t,\, p,\, q \textit{ are different,} \\ \textit{or (B2), (D1), and (D2) hold and } |t-p|, |t-q| > \tau, \textit{ then} \\ \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'\epsilon_{it}] = \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'y_{it}] - \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'\boldsymbol{x}_{it}]\boldsymbol{\beta} = \mathbf{0}_{K1}. \end{cases}$$

$$(35) \quad \begin{cases} \textit{When either (C1), (D1), and (D2) hold and } t,\, p,\, q \textit{ are different,} \\ \textit{or (C2), (D1), and (D2) hold and } |t-p|, |t-q| > \tau, \textit{ then} \\ \mathsf{E}[(\Delta y_{ipq})\epsilon_{it}] = \mathsf{E}[(\Delta y_{ipq})y_{it}] - \mathsf{E}[(\Delta y_{ipq})\boldsymbol{x}_{it}]\boldsymbol{\beta} = 0. \end{cases}$$

The intercept c needs a comment. When mean stationarity of the latent regressor, (D1), holds, then $\mathsf{E}(\Delta\boldsymbol{x}_{ipq}) = \mathbf{0}_{1K}$ and $\mathsf{E}(\Delta y_{ipq}) = 0$. If we relax (D1), which cannot be assumed to hold in many situations due to non-stationarity, we get

$$\mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'\epsilon_{it}] = \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'y_{it}] - \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})']c - \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'\boldsymbol{x}_{it}]\boldsymbol{\beta} = \mathbf{0}_{K1},$$

$$\mathsf{E}[(\Delta y_{ipq})\epsilon_{it}] = \mathsf{E}[(\Delta y_{ipq})y_{it}] - \mathsf{E}[(\Delta y_{ipq})]c - \mathsf{E}[(\Delta y_{ipq})\boldsymbol{x}_{it}]\boldsymbol{\beta} = 0.$$

Eliminating c by means of $\mathsf{E}(\epsilon_{it}) = \mathsf{E}(y_{it}) - c - E(\boldsymbol{x}_{it})\boldsymbol{\beta} = 0$ leads to the following modifications of (34) and (35):

$$(36) \begin{cases} \textit{When either (B1) and (D2) hold and } t,\, p,\, q \textit{ are different,} \\ \textit{or (B2) and (D2) hold and } |t-p|, |t-q| > \tau, \textit{ then} \\ \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'\epsilon_{it}] = \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'(y_{it} - \mathsf{E}(y_{it}))] - \mathsf{E}[(\Delta\boldsymbol{x}_{ipq})'(\boldsymbol{x}_{it} - \mathsf{E}(\boldsymbol{x}_{it}))]\boldsymbol{\beta} = \mathbf{0}_{K1}. \end{cases}$$

$$(37) \begin{cases} \textit{When either (C1) and (D2) hold and } t,\, p,\, q \textit{ are different,} \\ \textit{or (C2) and (D2) hold and } |t-p|, |t-q| > \tau, \textit{ then} \\ \mathsf{E}[(\Delta y_{ipq})\epsilon_{it}] = \mathsf{E}[(\Delta y_{ipq})(y_{it} - \mathsf{E}(y_{it}))] - \mathsf{E}[(\Delta y_{ipq})(\boldsymbol{x}_{it} - \mathsf{E}(\boldsymbol{x}_{it}))]\boldsymbol{\beta} = 0. \end{cases}$$

The OC's (32) – (37), corresponding to (19) in the general exposition of the GMM, will be instrumental in constructing our GMM estimators. Not all of these OC's, whose number is substantial even for small T, are, of course, independent. Let us examine the relationships between the OC's in (32) – (33) and between the OC's in (34) – (35). Some of these conditions are redundant, *i.e.*, linearly dependent of other conditions. Confining attention to the OC's relating to the $\boldsymbol{x}$'s, we have[6]

(α) *Assume that (B1) and (C1) are satisfied. Then: (i) All OC's (32) are linearly dependent on all admissible OC's relating to equations differenced over one period and a subset of the OC's relating to two-period differences. (ii) All OC's (34) are linearly dependent on all admissible OC's relating to IV's differenced over one period and a subset of the IV's differenced over two periods.*

(β) *Assume that (B2) and (C2) are satisfied. Then: (i) All OC's (33) are linearly dependent on all admissible OC's relating to equations differenced over one period and a subset of the OC's relating to differences over* $2(\tau+1)$ *periods. (ii) All OC's (35) are linearly dependent on all admissible OC's relating to IV's differenced over one period and a subset of the IV's differenced over* $2(\tau+1)$ *periods.*

We denote the non-redundant conditions defined by (α) – (β) as *essential* OC's. The following propositions are shown in Biørn (2000, section 2.d):

Proposition 1: *Assume that (B1) and (C1) are satisfied. Then*

(a) $\mathsf{E}[\boldsymbol{x}'_{ip}(\Delta\epsilon_{it,t-1})] = \mathbf{0}_{K,1}$ *for* $p = 1,\dots,t-2,t+1,\dots,T;\ t = 2,\dots,T$ *are* $K(T-1)(T-2)$ *essential OC's for equations differenced over one period.*

(b) $\mathsf{E}[\boldsymbol{x}'_{it}(\Delta\epsilon_{it+1,t-1})] = \mathbf{0}_{K,1}$ *for* $t = 2,\dots,T-1$ *are* $K(T-2)$ *essential OC's for equations differenced over two periods.*

(c) The other OC's are redundant: among the $\frac{1}{2}KT(T-1)(T-2)$ *conditions in (32) when* $T > 2$*, only* $KT(T-2)$ *are essential.*

Proposition 2: *Assume that (B1) and (C1) are satisfied. Then*

(a) $\mathsf{E}[(\Delta\boldsymbol{x}_{ip,p-1})'\epsilon_{it}] = \mathbf{0}_{K,1}$ *for* $t = 1,\dots,p-2,p+1,\dots,T;\ p = 2,\dots,T$ *are* $K(T-1)(T-2)$ *essential OC's for equations in levels, with IV's differenced over one period.*

(b) $\mathsf{E}[(\Delta \boldsymbol{x}_{it+1,t-1})'\epsilon_{it}] = \mathbf{0}_{K,1}$ *for* $t = 2,\dots,T-1$ *are* $K(T-2)$ *essential OC's for equations in levels, with IV's differenced over two periods.*

(c) The other OC's are redundant: among the $\frac{1}{2}KT(T-1)(T-2)$ *conditions in (33) when* $T > 2$*, only* $KT(T-2)$ *are essential.*

For generalizations to the case where ϵ_{it} is a *MA(*τ*) process*, see Biørn (2000, section 2.d). These propositions can be (trivially) modified to include also the essential and redundant OC's in the y's or the Δy's, given in (33) and (35).

5.c The estimators

We are now in a position to specialize (23) and (24) to define (i) consistent GMM estimators of $\boldsymbol{\beta}$ in (5) *for one pair of periods* (t,θ), utilizing as IV's for $\Delta \boldsymbol{x}_{it\theta}$ *all admissible* $\boldsymbol{x}_{ip}$*'s*, and (ii) consistent GMM estimators of $\boldsymbol{\beta}$ in (3), *i.e.*, *for one period* (t), utilizing as IV's for $\boldsymbol{x}_{it}$ *all admissible* $\Delta \boldsymbol{x}_{ipq}$*'s.* This is a preliminary to Section 6, in which we combine on the one hand (i) the differenced equations for all pairs of periods, and on the other hand (ii) the level equations for all periods, respectively, in one equation system.

We let $\boldsymbol{P}_{t\theta}$ denote the $((T-2)\times T)$ selection matrix obtained by deleting from $\boldsymbol{I}_T$ rows t and θ, and introduce the $[(T-2)\times T]$ matrix

$$\boldsymbol{D}_t = \begin{bmatrix} \boldsymbol{d}_{21} \\ \vdots \\ \boldsymbol{d}_{t-1,t-2} \\ \boldsymbol{d}_{t+1,t-1} \\ \boldsymbol{d}_{t+2,t+1} \\ \vdots \\ \boldsymbol{d}_{T,T-1} \end{bmatrix}, \qquad t = 1,\dots,T,$$

which is a one-period differencing matrix, except that $\boldsymbol{d}_{t,t-1}$ and $\boldsymbol{d}_{t+1,t}$ are replaced by their sum, $\boldsymbol{d}_{t+1,t-1}$, the two-period difference being effective only for $t = 2,\dots,T-1$, and use the notation

$$\boldsymbol{y}_{i(t\theta)} = \boldsymbol{P}_{t\theta}\boldsymbol{y}_{i\cdot}, \quad \boldsymbol{X}_{i(t\theta)} = \boldsymbol{P}_{t\theta}\boldsymbol{X}_{i\cdot}, \quad \boldsymbol{x}_{i(t\theta)} = \mathrm{vec}(\boldsymbol{X}_{i(t\theta)})',$$
$$\Delta\boldsymbol{y}_{i(t)} = \boldsymbol{D}_t\boldsymbol{y}_{i\cdot}, \quad \Delta\boldsymbol{X}_{i(t)} = \boldsymbol{D}_t\boldsymbol{X}_{i\cdot}, \quad \Delta\boldsymbol{x}_{i(t)} = \mathrm{vec}(\Delta\boldsymbol{X}_{i(t)})',$$

etc. Here $\boldsymbol{X}_{i(t\theta)}$ denotes the $[(T-2)\times K]$ matrix of $\boldsymbol{x}$ levels obtained by *deleting* rows t and θ from $\boldsymbol{X}_{i\cdot}$, and $\Delta\boldsymbol{X}_{i(t)}$ denotes the $[(T-2)\times K]$ matrix of $\boldsymbol{x}$ differences obtained by stacking all one-period differences between rows of $\boldsymbol{X}_{i\cdot}$ *not including period* t and the single two-period difference between the columns for periods $t+1$ and $t-1$. The vectors $\boldsymbol{y}_{i(t\theta)}$ and $\Delta\boldsymbol{y}_{i(t)}$ are constructed from $\boldsymbol{y}_{i\cdot}$ in a similar way. In general, we let subscripts $(t\theta)$ and (t) on a matrix or vector denote deletion of $(t\theta)$ differences and t levels, respectively. Stacking $\boldsymbol{y}'_{i(t\theta)}$, $\Delta\boldsymbol{y}'_{i(t)}$, $\boldsymbol{x}_{i(t\theta)}$, and $\Delta\boldsymbol{x}_{i(t)}$ by individuals, we get

$$\boldsymbol{Y}_{(t\theta)} = \begin{bmatrix} \boldsymbol{y}'_{1(t\theta)} \\ \vdots \\ \boldsymbol{y}'_{N(t\theta)} \end{bmatrix}, \Delta\boldsymbol{Y}_{(t)} = \begin{bmatrix} \Delta\boldsymbol{y}'_{1(t)} \\ \vdots \\ \Delta\boldsymbol{y}'_{N(t)} \end{bmatrix}, \boldsymbol{X}_{(t\theta)} = \begin{bmatrix} \boldsymbol{x}_{1(t\theta)} \\ \vdots \\ \boldsymbol{x}_{N(t\theta)} \end{bmatrix}, \Delta\boldsymbol{X}_{(t)} = \begin{bmatrix} \Delta\boldsymbol{x}_{1(t)} \\ \vdots \\ \Delta\boldsymbol{x}_{N(t)} \end{bmatrix},$$

which have dimensions $(N\times(T-2))$, $(N\times(T-2))$, $(N\times(T-2)K)$, and $(N\times(T-2)K)$, respectively. *These four matrices contain the alternative IV sets in the GMM procedures to be considered below.*

Equation in differences, IV's in levels. Using $\boldsymbol{X}_{(t\theta)}$ as IV matrix for $\Delta\boldsymbol{X}_{t\theta}$, we obtain the following estimator of $\boldsymbol{\beta}$, *specific to period* (t,θ) *differences and utilizing all admissible* $\boldsymbol{x}$ *level IV's*,

$$\begin{aligned}
(38)\qquad \widehat{\boldsymbol{\beta}}_{Dx(t\theta)} &= \Big[(\Delta\boldsymbol{X}_{t\theta})'\boldsymbol{X}_{(t\theta)}\left(\boldsymbol{X}'_{(t\theta)}\boldsymbol{X}_{(t\theta)}\right)^{-1}\boldsymbol{X}'_{(t\theta)}(\Delta\boldsymbol{X}_{t\theta})\Big]^{-1} \\
&\qquad\times\Big[(\Delta\boldsymbol{X}_{t\theta})'\boldsymbol{X}_{(t\theta)}\left(\boldsymbol{X}'_{(t\theta)}\boldsymbol{X}_{(t\theta)}\right)^{-1}\boldsymbol{X}'_{(t\theta)}(\Delta\boldsymbol{y}_{t\theta})\Big] \\
&= \Big[\big[\textstyle\sum_i(\Delta\boldsymbol{x}_{it\theta})'\boldsymbol{x}_{i(t\theta)}\big]\big[\textstyle\sum_i\boldsymbol{x}'_{i(t\theta)}\boldsymbol{x}_{i(t\theta)}\big]^{-1}\big[\textstyle\sum_i\boldsymbol{x}'_{i(t\theta)}(\Delta\boldsymbol{x}_{it\theta})\big]\Big]^{-1} \\
&\qquad\times\Big[\big[\textstyle\sum_i(\Delta\boldsymbol{x}_{it\theta})'\boldsymbol{x}_{i(t\theta)}\big]\big[\textstyle\sum_i\boldsymbol{x}'_{i(t\theta)}\boldsymbol{x}_{i(t\theta)}\big]^{-1}\big[\textstyle\sum_i\boldsymbol{x}'_{i(t\theta)}(\Delta y_{it\theta})\big]\Big].
\end{aligned}$$

It exists if $\boldsymbol{X}'_{(t\theta)}\boldsymbol{X}_{(t\theta)}$ has rank $(T-2)K$, which requires $N\geq(T-2)K$. This estimator examplifies (23), utilizes the OC $\mathsf{E}[\boldsymbol{x}'_{i(t\theta)}(\Delta\epsilon_{it\theta})] = \mathbf{0}_{(T-2)K,1}$ – which follows from (32) – and minimizes the quadratic form

$$[N^{-1}\boldsymbol{X}'_{(t\theta)}\Delta\boldsymbol{\epsilon}_{t\theta}]'[N^{-2}\boldsymbol{X}'_{(t\theta)}\boldsymbol{X}_{(t\theta)}]^{-1}[N^{-1}\boldsymbol{X}'_{(t\theta)}\Delta\boldsymbol{\epsilon}_{t\theta}].$$

The weight matrix $(N^{-2}\boldsymbol{X}'_{(t\theta)}\boldsymbol{X}_{(t\theta)})^{-1}$ is proportional to the inverse of the (asymptotic) covariance matrix of $N^{-1}\boldsymbol{X}'_{(t\theta)}\Delta\boldsymbol{\epsilon}_{t\theta}$ when $\Delta\boldsymbol{\epsilon}_{it\theta}$ is IID across i. *The consistency of $\widehat{\boldsymbol{\beta}}_{Dx(t\theta)}$ relies on Assumptions (A), (B1), and (E).*

Two *modifications* of $\widehat{\boldsymbol{\beta}}_{Dx(t\theta)}$ exist: First, if $\text{var}(\Delta\epsilon_{it\theta})$ varies with i, we can increase the efficiency of (38) by replacing $\boldsymbol{x}'_{i(t\theta)}\boldsymbol{x}_{i(t\theta)}$ by $\boldsymbol{x}'_{i(t\theta)}(\widehat{\Delta\epsilon_{it\theta}})^2\boldsymbol{x}_{i(t\theta)}$, which gives an asymptotically optimal GMM estimator of the form (24). Second, instead of using $\boldsymbol{X}_{(t\theta)}$ as IV matrix for $\Delta\boldsymbol{X}_{t\theta}$, we may either, if $K=1$, use $\boldsymbol{Y}_{(t\theta)}$, or, for arbitrary K, $(\boldsymbol{X}_{(t\theta)} : \boldsymbol{Y}_{(t\theta)})$, provided that also (C1) is satisfied.

Equation in levels, IV's in differences. Using $\Delta\boldsymbol{X}_{(t)}$ as IV matrix for $\boldsymbol{X}_t$ (for notational simplicity we omit the 'dot' subscript on $\boldsymbol{X}_{\cdot t}$ and $\boldsymbol{y}_{\cdot t}$), we get the following estimator of $\boldsymbol{\beta}$, *specific to period t levels, utilizing all admissible $\boldsymbol{x}$ difference IV's,*

$$
\begin{aligned}
(39)\qquad \widehat{\boldsymbol{\beta}}_{Lx(t)} &= \Big[\boldsymbol{X}'_t(\Delta\boldsymbol{X}_{(t)})\left((\Delta\boldsymbol{X}_{(t)})'(\Delta\boldsymbol{X}_{(t)})\right)^{-1}(\Delta\boldsymbol{X}_{(t)})'\boldsymbol{X}_t\Big]^{-1} \\
&\qquad \times \Big[\boldsymbol{X}'_t(\Delta\boldsymbol{X}_{(t)})\left((\Delta\boldsymbol{X}_{(t)})'(\Delta\boldsymbol{X}_{(t)})\right)^{-1}(\Delta\boldsymbol{X}_{(t)})'\boldsymbol{y}_t\Big] \\
&= \Big[\Big[\textstyle\sum_i \boldsymbol{x}'_{it}(\Delta\boldsymbol{x}_{i(t)})\Big]\Big[\textstyle\sum_i(\Delta\boldsymbol{x}_{i(t)})'(\Delta\boldsymbol{x}_{i(t)})\Big]^{-1}\Big[\textstyle\sum_i(\Delta\boldsymbol{x}_{i(t)})'\boldsymbol{x}_{it}\Big]\Big]^{-1} \\
&\qquad \times \Big[\Big[\textstyle\sum_i \boldsymbol{x}'_{it}(\Delta\boldsymbol{x}_{i(t)})\Big]\Big[\textstyle\sum_i(\Delta\boldsymbol{x}_{i(t)})'(\Delta\boldsymbol{x}_{i(t)})\Big]^{-1}\Big[\textstyle\sum_i(\Delta\boldsymbol{x}_{i(t)})'y_{it}\Big]\Big].
\end{aligned}
$$

It exists if $(\Delta\boldsymbol{X}_{(t)})'(\Delta\boldsymbol{X}_{(t)})$ has rank $(T-2)K$, which requires $N \geq (T-2)K$. This estimator examplifies (23), utilizes the OC $\mathsf{E}[(\Delta\boldsymbol{x}_{i(t)})'\epsilon_{it}] = \mathbf{0}_{(T-2)K,1}$ – which follows from (34) – and minimizes the quadratic form

$$[N^{-1}(\Delta\boldsymbol{X}_{(t)})'\boldsymbol{\epsilon}_t]'[N^{-2}(\Delta\boldsymbol{X}_{(t)})'(\Delta\boldsymbol{X}_{(t)})]^{-1}[N^{-1}(\Delta\boldsymbol{X}_{(t)})'\boldsymbol{\epsilon}_t].$$

The weight matrix $[N^{-2}(\Delta\boldsymbol{X}_{(t)})'(\Delta\boldsymbol{X}_{(t)})]^{-1}$ is proportional to the inverse of the (asymptotic) covariance matrix of $N^{-1}(\Delta\boldsymbol{X}_{(t)})'\epsilon_t$ when ϵ_{it} is IID across i. *The consistency of $\widehat{\boldsymbol{\beta}}_{Lx(t)}$ relies on Assumptions (A), (B1), (D1), (D2), and (E).*

Three *modifications* of $\widehat{\boldsymbol{\beta}}_{Lx(t)}$ exist: First, if $\text{var}(\epsilon_{it})$ varies with i, we can increase the efficiency of (39) by replacing $(\Delta\boldsymbol{x}_{i(t)})'(\Delta\boldsymbol{x}_{i(t)})$ by $(\Delta\boldsymbol{x}_{i(t)})'(\widehat{\epsilon}_{it})^2(\Delta\boldsymbol{x}_{i(t)})$, which gives an asymptotically optimal GMM estimator of the form (24). Second, instead of using $\Delta \boldsymbol{X}_{(t)}$ as IV matrix for $\boldsymbol{X}_t$, we may either, if $K = 1$, use $\Delta \boldsymbol{Y}_{(t)}$, or, for arbitrary K, $(\Delta \boldsymbol{X}_{(t)} : \Delta \boldsymbol{Y}_{(t)})$, provided that also (C1) is satisfied. Third, we can deduct period means from $\boldsymbol{x}_{it}$ and y_{it} and relax the stationarity in mean assumption of the latent regressor, (D1); cf. (36) – (37).

If we relax Assumptions (B1) or (C1) and replace them by (B2) or (C2), we must reconstruct the OC's underlying (38) and (39) to ensure that the variables in the IV matrix have a lead or lag of at least $\tau + 1$ periods to the regressor, to 'get clear of' the τ period memory of the MA(τ) process. The IV sets will then be reduced.

6 Composite GMM estimators combining differences and levels

We now take the single equation GMM estimators (38) and (39) and their heteroskedasticity robust modifications one step further and construct GMM estimators of the common coefficient vector $\boldsymbol{\beta}$ when we combine the essential OC's for all periods, *i.e.*, for all differences or for all levels. This gives *multi-equation, or overall, GMM estimators* for panel data with measurement errors, still belonging to the general framework described in Section 4. The procedures to be described in this section, like the single-equation procedures in section 5.c, may be modified to be applicable to situations with disturbance/error autocorrelation.

Equation in differences, IV's in levels. Consider the differenced equation (5) for *all* $\theta = t - 1$ and $\theta = t - 2$. These $(T-1) + (T-2)$ equations stacked for individual i read

$$\begin{bmatrix} \Delta y_{i21} \\ \vdots \\ \Delta y_{i,T,T-1} \\ \Delta y_{i31} \\ \vdots \\ \Delta y_{i,T,T-2} \end{bmatrix} = \begin{bmatrix} \Delta \boldsymbol{x}_{i21} \\ \vdots \\ \Delta \boldsymbol{x}_{i,T,T-1} \\ \Delta \boldsymbol{x}_{i31} \\ \vdots \\ \Delta \boldsymbol{x}_{i,T,T-2} \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \Delta \epsilon_{i21} \\ \vdots \\ \Delta \epsilon_{i,T,T-1} \\ \Delta \epsilon_{i31} \\ \vdots \\ \Delta \epsilon_{i,T,T-2} \end{bmatrix}, \tag{40}$$

or, compactly,

$$\Delta \boldsymbol{y}_i = (\Delta \boldsymbol{X}_i)\boldsymbol{\beta} + \Delta \boldsymbol{\epsilon}_i.$$

The IV matrix (cf. Proposition 1) is the $((2T-3) \times KT(T-2))$ matrix

$$\boldsymbol{Z}_i = \begin{bmatrix} \boldsymbol{x}_{i(21)} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{x}_{i(T,T-1)} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{x}_{i2} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{x}_{i,T-1} \end{bmatrix}. \tag{41}$$

We here use different IV's for the $(T-1)+(T-2)$ equations in (40), with $\boldsymbol{\beta}$ as a common slope coefficient. Let

$$\begin{aligned} \Delta \boldsymbol{y} &= [(\Delta \boldsymbol{y}_1)', \dots, (\Delta \boldsymbol{y}_N)']', \qquad \Delta \boldsymbol{\epsilon} = [(\Delta \boldsymbol{\epsilon}_1)', \dots, (\Delta \boldsymbol{\epsilon}_N)']', \\ \Delta \boldsymbol{X} &= [(\Delta \boldsymbol{X}_1)', \dots, (\Delta \boldsymbol{X}_N)']', \quad \boldsymbol{Z} = [\boldsymbol{Z}_1', \dots, \boldsymbol{Z}_N']'. \end{aligned}$$

The overall GMM estimator corresponding to (32), which we now write as $\mathsf{E}[\boldsymbol{Z}_i'(\Delta\epsilon_i)] = \boldsymbol{0}_{T(T-2)K,1}$, minimizing $[N^{-1}(\Delta\boldsymbol{\epsilon})'\boldsymbol{Z}](N^{-2}\boldsymbol{V})^{-1}[N^{-1}\boldsymbol{Z}'(\Delta\boldsymbol{\epsilon})]$ for $\boldsymbol{V} = \boldsymbol{Z}'\boldsymbol{Z}$, can be written as

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_{Dx} &= [(\Delta\boldsymbol{X})'\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'(\Delta\boldsymbol{X})]^{-1}[(\Delta\boldsymbol{X})'\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'(\Delta\boldsymbol{y})] \\ &= \Big[[\textstyle\sum_i(\Delta\boldsymbol{X}_i)'\boldsymbol{Z}_i]\,[\sum_i \boldsymbol{Z}_i'\boldsymbol{Z}_i]^{-1}\,[\sum_i \boldsymbol{Z}_i'(\Delta\boldsymbol{X}_i)]\Big]^{-1} \\ &\qquad \times \Big[[\textstyle\sum_i(\Delta\boldsymbol{X}_i)'\boldsymbol{Z}_i]\,[\sum_i \boldsymbol{Z}_i'\boldsymbol{Z}_i]^{-1}\,[\sum_i \boldsymbol{Z}_i'(\Delta\boldsymbol{y}_i)]\Big]. \end{aligned} \tag{42}$$

This estimator examplifies (23). *The consistency of* $\boldsymbol{\beta}_{Dx}$ *relies on Assumptions (A), (B1), and (E).* If $\Delta\boldsymbol{\epsilon}$ has a non-scalar covariance matrix, a more efficient estimator is obtained for $\boldsymbol{V} = \boldsymbol{V}_{Z(\Delta\epsilon)} = \mathsf{E}[\boldsymbol{Z}'(\Delta\boldsymbol{\epsilon})(\Delta\boldsymbol{\epsilon})'\boldsymbol{Z}]$, which gives

$$\widetilde{\boldsymbol{\beta}}_{Dx} = \Big[(\Delta\boldsymbol{X})'\boldsymbol{Z}\boldsymbol{V}_{Z(\Delta\epsilon)}^{-1}\boldsymbol{Z}'(\Delta\boldsymbol{X})\Big]^{-1}\Big[(\Delta\boldsymbol{X})'\boldsymbol{Z}\boldsymbol{V}_{Z(\Delta\epsilon)}^{-1}\boldsymbol{Z}'(\Delta\boldsymbol{y})\Big].$$

We can estimate $\boldsymbol{V}_{Z(\Delta\epsilon)}/N$ consistently from the residuals obtained from (42), $\widehat{\Delta\boldsymbol{\epsilon}}_i = \Delta\boldsymbol{y}_i - (\Delta\boldsymbol{X}_i)\widehat{\boldsymbol{\beta}}_{Dx}$, by $\widehat{\boldsymbol{V}}_{Z(\Delta\epsilon)}/N = (1/N)\sum_{i=1}^N \boldsymbol{Z}_i'(\widehat{\Delta\boldsymbol{\epsilon}}_i)(\widehat{\Delta\boldsymbol{\epsilon}}_i)'\boldsymbol{Z}_i$. The resulting asymptotically optimal GMM estimator, which examplifies (24), is

$$\begin{aligned} \widetilde{\boldsymbol{\beta}}_{Dx} &= \Big[[\textstyle\sum_i(\Delta\boldsymbol{X}_i)'\boldsymbol{Z}_i][\sum_i \boldsymbol{Z}_i'\widehat{\Delta\boldsymbol{\epsilon}}_i\widehat{\Delta\boldsymbol{\epsilon}}_i'\boldsymbol{Z}_i]^{-1}[\sum_i \boldsymbol{Z}_i'(\Delta\boldsymbol{X}_i)]\Big]^{-1} \\ &\qquad \times \Big[[\textstyle\sum_i(\Delta\boldsymbol{X}_i)'\boldsymbol{Z}_i][\sum_i \boldsymbol{Z}_i'\widehat{\Delta\boldsymbol{\epsilon}}_i\widehat{\Delta\boldsymbol{\epsilon}}_i'\boldsymbol{Z}_i]^{-1}[\sum_i \boldsymbol{Z}_i'(\Delta\boldsymbol{y}_i)]\Big]. \end{aligned} \tag{43}$$

The estimators $\widehat{\boldsymbol{\beta}}_{Dx}$ and $\widetilde{\boldsymbol{\beta}}_{Dx}$ can be modified by extending $\boldsymbol{x}_{i(t,t-1)}$ to $(\boldsymbol{x}_{i(t,t-1)} : \boldsymbol{y}'_{i(t,t-1)})$ and $\boldsymbol{x}_{it}$ to $(\boldsymbol{x}_{it} : \boldsymbol{y}_{it})$ in (41), also exploiting Assumption (C1) and the OC's in the y's. This is indicated be replacing subscript Dx by Dy or Dxy on the estimator symbols.

Table 24.2 contains, for the four manufacturing sectors and the two inputs, the overall GMM estimates obtained from the complete set of *differenced equations*. The standard deviation estimates are computed as described in the Appendix.[7] The estimated input-output elasticities (column 1, rows 1 and 3) are always lower than the inverse output-input elasticities (column 2, rows 2 and 4). This 'attenuation effect', also found for the OLS estimates (cf. Table 24.1), agrees with the fact that $\widehat{\boldsymbol{\beta}}_{Dx}$ and $\widehat{\boldsymbol{\beta}}_{Dy}$ can be interpreted as obtained by running standard 2SLS on the 'original' and on the 'reverse regression' version of (40), respectively. Under both normalizations, the estimates utilizing the y instruments (column 2) tend to exceed those based on the x instruments (column 1). Using the optimal weighting (columns 4 and 5), we find that the estimates are more precise, according to the standard deviation estimates, than those in columns 1 and 2, as they should be. The standard deviation estimates for capital are substantially higher than for materials.

Sargan-Hansen orthogonality test statistics, which are asymptotically distributed as χ^2 with a number of degrees of freedom equal to the number of OC's imposed less the number of coefficients estimated (one in this case) under the null hypothesis of orthogonality [cf. Hansen (1982), Newey (1985), and Arellano and Bond (1991)], corresponding to the asymptotically efficient estimates in columns 4 and 5, are reported in columns 6 and 7. For *materials*, these statistics indicate non-rejection of the full set of OC's when using the x's as IV's for the original regression (rows 1) and the y's as IV's for the reverse regression (rows 2) – *i.e.*, the output variable in both cases – with p values exceeding 5 %. The OC's when using the y's as IV's for the original regression and the x's as IV's for the reverse regression – *i.e.*, the material input variable in both cases – is however rejected. For *capital* the tests come out with very low p values in all cases, indicating rejection of the OC's. This may be due to lagged response, autocorrelated measurement errors or disturbances and/or (deterministic or stochastic) trends in the capital input relationship. The latter would violate, for example, the stationarity assumption for capital. Owing to the short time span of our data, we have not, however, performed a cointegration analysis.

All the results in Table 24.2 uniformly indicate *a marginal input elasticity of materials,* $1/\mu$, *larger than one*; $\widehat{\boldsymbol{\beta}}_{Dx}$ and $\widetilde{\boldsymbol{\beta}}_{Dx}$ are, however, lower than the (inconsistent) estimate obtained by running OLS regression on differences (cf. $\widehat{\boldsymbol{\beta}}_{OLSD}$ for the materials-output regression in Table 24.1), and $\widehat{\boldsymbol{\beta}}_{Dy}$ and $\widetilde{\boldsymbol{\beta}}_{Dy}$ are higher than the (inconsistent) estimate obtained by running reverse OLS regression on differences (cf. $\widehat{\boldsymbol{\beta}}_{OLSD}$ for the output-materials regression in Table 24.1).

Equation in levels, IV's in differences. We next consider the procedures for estimating all the level equations (3) with the IV's in differences. The T stacked level equations for individual i are

$$(44)\qquad \begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} c \\ \vdots \\ c \end{bmatrix} + \begin{bmatrix} \boldsymbol{x}_{i1} \\ \vdots \\ \boldsymbol{x}_{iT} \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \epsilon_{i1} \\ \vdots \\ \epsilon_{iT} \end{bmatrix},$$

or compactly, omitting the 'dot' subscript [cf. (4)],

$$\boldsymbol{y}_i = \boldsymbol{e}_T c + \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i.$$

The IV matrix (cf. Proposition 2) is the $(T \times T(T-2)K)$ matrix

$$(45)\qquad \Delta \boldsymbol{Z}_i = \begin{bmatrix} \Delta \boldsymbol{x}_{i(1)} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \Delta \boldsymbol{x}_{i(T)} \end{bmatrix}.$$

Again, we use different IV's for different equations, considering (44) as T equations with $\boldsymbol{\beta}$ as a common slope coefficient. Let

$$\boldsymbol{y} = [\boldsymbol{y}_1', \ldots, \boldsymbol{y}_N']', \qquad \boldsymbol{\epsilon} = [\boldsymbol{\epsilon}_1', \ldots, \boldsymbol{\epsilon}_N']',$$
$$\boldsymbol{X} = [\boldsymbol{X}_1', \ldots, \boldsymbol{X}_N']', \quad \Delta \boldsymbol{Z} = [(\Delta \boldsymbol{Z}_1)', \ldots, (\Delta \boldsymbol{Z}_N)']'.$$

The overall GMM estimator corresponding to (34), which we now write as $\mathsf{E}[(\Delta \boldsymbol{Z}_i)'\epsilon_i] = \mathbf{0}_{T(T-2)K,1}$, minimizing $[N^{-1}\boldsymbol{\epsilon}'(\Delta \boldsymbol{Z})](N^{-2}\boldsymbol{V}_{\Delta})^{-1}[N^{-1}(\Delta \boldsymbol{Z})'\boldsymbol{\epsilon}]$ for $\boldsymbol{V}_{\Delta} = (\Delta \boldsymbol{Z})'(\Delta \boldsymbol{Z})$, can be written as

$$
\begin{aligned}
(46)\ \widehat{\boldsymbol{\beta}}_{Lx} &= [\boldsymbol{X}'(\Delta \boldsymbol{Z})[(\Delta \boldsymbol{Z})'(\Delta \boldsymbol{Z})]^{-1}(\Delta \boldsymbol{Z})'\boldsymbol{X}]^{-1}[\boldsymbol{X}'(\Delta \boldsymbol{Z})[(\Delta \boldsymbol{Z})'(\Delta \boldsymbol{Z})]^{-1}(\Delta \boldsymbol{Z})'\boldsymbol{y}] \\
&= \Big[[\textstyle\sum_i \boldsymbol{X}_i'(\Delta \boldsymbol{Z}_i)]\,[\sum_i(\Delta \boldsymbol{Z}_i)'(\Delta \boldsymbol{Z}_i)]^{-1}\,[\sum_i(\Delta \boldsymbol{Z}_i)'\boldsymbol{X}_i]\Big]^{-1} \\
&\qquad \times \Big[[\textstyle\sum_i \boldsymbol{X}_i'(\Delta \boldsymbol{Z}_i)]\,[\sum_i(\Delta \boldsymbol{Z}_i)'(\Delta \boldsymbol{Z}_i)]^{-1}\,[\sum_i(\Delta \boldsymbol{Z}_i)'\boldsymbol{y}_i]\Big].
\end{aligned}
$$

This estimator examplifies (23). *The consistency of* $\boldsymbol{\beta}_{Lx}$ *relies on Assumptions (A), (B1), (D1), (D2), and (E).* If $\boldsymbol{\epsilon}$ has a non-scalar covariance matrix, a more efficient estimator is obtained for $\boldsymbol{V}_{\Delta} = \boldsymbol{V}_{(\Delta Z)\epsilon} = \mathsf{E}[(\Delta \boldsymbol{Z})'\boldsymbol{\epsilon}\boldsymbol{\epsilon}'(\Delta \boldsymbol{Z})]$, which gives

$$
\widetilde{\boldsymbol{\beta}}_{Lx} = \Big[\boldsymbol{X}'(\Delta \boldsymbol{Z})\boldsymbol{V}_{(\Delta Z)\epsilon}^{-1}(\Delta \boldsymbol{Z})'\boldsymbol{X}\Big]^{-1}\Big[\boldsymbol{X}'(\Delta \boldsymbol{Z})\boldsymbol{V}_{(\Delta Z)\epsilon}^{-1}(\Delta \boldsymbol{Z})'\boldsymbol{y}\Big].
$$

We can estimate $\boldsymbol{V}_{(\Delta Z)\epsilon}/N$ consistently from the residuals obtained from (46) $\widehat{\boldsymbol{\epsilon}}_i = \boldsymbol{y}_i - \boldsymbol{X}_i\widehat{\boldsymbol{\beta}}_{Lx}$, by $\widehat{\boldsymbol{V}}_{(\Delta Z)\epsilon}/N = (1/N)\sum_{i=1}^{N}(\Delta \boldsymbol{Z}_i)'\,\widehat{\boldsymbol{\epsilon}}_i\widehat{\boldsymbol{\epsilon}}_i'(\Delta \boldsymbol{Z}_i)$. We can here omit the intercept c; see Section 5.b. The resulting asymptotically optimal GMM estimator, which examplifies (24), is

$$
\begin{aligned}
(47)\qquad \widetilde{\boldsymbol{\beta}}_{Lx} &= \Big[[\textstyle\sum_i \boldsymbol{X}_i'(\Delta \boldsymbol{Z}_i)]\,[\sum_i(\Delta \boldsymbol{Z}_i)'\,\widehat{\boldsymbol{\epsilon}}_i\widehat{\boldsymbol{\epsilon}}_i'(\Delta \boldsymbol{Z}_i)]^{-1}\,[\sum_i(\Delta \boldsymbol{Z}_i)'\boldsymbol{X}_i]\Big]^{-1} \\
&\qquad \times \Big[[\textstyle\sum_i \boldsymbol{X}_i'(\Delta \boldsymbol{Z}_i)]\,[\sum_i(\Delta \boldsymbol{Z}_i)'\,\widehat{\boldsymbol{\epsilon}}_i\widehat{\boldsymbol{\epsilon}}_i'(\Delta \boldsymbol{Z}_i)]^{-1}\,[\sum_i(\Delta \boldsymbol{Z}_i)'\boldsymbol{y}_i]\Big].
\end{aligned}
$$

The estimators $\widehat{\boldsymbol{\beta}}_{Lx}$ and $\widetilde{\boldsymbol{\beta}}_{Lx}$ can be modified by extending $\Delta \boldsymbol{x}_{i(t)}$ to $(\Delta \boldsymbol{x}_{i(t)} : \Delta \boldsymbol{y}_{i(t)}')$ in (45), also exploiting Assumption (C1) and the OC's in the Δy's. This is indicated be replacing subscript Lx by Ly or Lxy on the estimator symbols. We can also deduct period means from the level variables in (44) to take account of possible non-stationarity of these variables and relax (D1) [cf. (36) – (37)].

Tables 24.3 and 24.4 contain the overall GMM estimates obtained from the complete set of *level equations*, the first using the untransformed observations and the second based on observations measured from their year means. The orthogonality test statistics (columns 6 and 7) give for *materials* conclusions similar to those for the differenced equation in Table 24.2 for Textiles and Chemicals (which have the fewer observations): Non-rejection of the OC's when using the x's as IV's (cf. $\chi^2(\widetilde{\boldsymbol{\beta}}_{Lx})$ in rows 1) and the y's as IV's (cf. $\chi^2(\widetilde{\boldsymbol{\beta}}_{Ly})$ in rows 2) – *i.e.*, the output variable in both cases – and rejection when using the y's as IV's in the materials-output regression and the x's as IV's in the output-materials regression – *i.e.*, the material input variable in both cases. For *capital*, the orthogonality test statistics once again come out with very low p values in all cases, which may again reflect mis-specified dynamics or trend effects. There is, however, a striking difference between Tables 24.3 and 24.4. In Table 24.3 – in which we make no adjustment for non-stationarity in means and impose (D1) – we find uniform rejection of the OC's for capital in all sectors and for Wood Products and Paper Products for materials. In Table 24.4 – in which we make adjustment for non-stationarity in means by deducting period means from the level variables and relax (D1) – we find non-rejection when using output as instrument for all sectors for materials (p values exceeding 5 %), and for capital in all sectors except Textiles and Wood Products (p values exceeding 1 %). Note that the set of orthogonality conditions under test in Tables 24.3 and 24.4 is larger than in Table 24.2, since it also includes Assumption (D2), time invariance of the covariance between the firm specific effect α_i and the latent regressor $\boldsymbol{\xi}_{it}$.

These estimates for the level equation, unlike those for the differenced equation in Table 24.2, however, do not uniformly give marginal input elasticity estimates of materials greater than one. Using level observations measured from year means (Table 24.4) and relaxing mean stationarity of the latent regressor, we get estimates exceeding one, while using untransformed observations and imposing mean stationarity, we get estimates less than one. There are also substantial differences for capital.

A tentative conclusion we can draw from the examples in Tables 24.2 – 24.4 is that overall GMM estimates of the input elasticity of *materials* with respect to output tend to be larger than one if we use either the equation in differences with IV's in levels or the equation in levels, measuring the observations from their year means, with IV's in differences. If we use the non-adjusted equation in levels with IV's in differences, the GMM estimates tend to be less than one. For capital, the picture is less clear. Overall, there is a considerable difference between the elasticity estimates of materials and those of capital. An interpretation we may give of this difference is that the underlying production technology is *non-homothetic*; cf. Section 2.

7 Concluding remarks

In this paper, we have constructed and illustrated several estimators which may handle jointly the heterogeneity problem and the measurement error problem in panel data. These problems may be untractable when only pure (single or repeated) cross section data or pure time series data are available. The estimators considered are estimators operating on period specific means, *inter alia*, the between period (BP) estimator, and Generalized Method of Moments (GMM) estimators. The GMM estimators use either equations in differences with level values as instruments, or equations in levels with differenced values as instruments. In both cases, the differences may be taken over one period or more.

In GMM estimation, not only instruments constructed from the observed regressors (x's), but also instruments constructed from the observed regressands (y's) may be useful, even if both are, formally, endogenous variables. Our empirical examples – using materials and capital input data and output data for firms in a single regressor case – indicate that for both normalizations of the equation, GMM estimates using y instruments tend to exceed those using x instruments. Even if the GMM estimates, unlike the OLS estimates, are consistent, they seem to some extent to be affected by the 'attenuation' known for the OLS in errors-in-variables situations. Using levels as instruments for differences or *vice versa* as a general estimation strategy within a GMM framework, however, may raise problems related to 'weak instruments'; cf. Nelson and Startz (1990) and Staiger and Stock (1997). It is left for future research to explore these problems, *e.g.*, by means of Monte

Carlo experiments.

The between period (BP) estimates on levels for the original and the reverse regression give virtually the same input elasticity for materials. For capital, we find substantial deviations between the two sets of BP estimators, which may indicate that measurement errors or disturbances have period specific, or strongly serially correlated, components.

Finally, we find that GMM estimates based on the equation in levels are more precise than those based on the equation in differences. Deducting period means from levels to compensate for non-stationarity of the latent regressor, give estimates for the level equation which are less precise and more sensitive to the choice of instrument set than those operating on untransformed levels. On the other hand, this kind of transformations of level variables may be needed to compensate for period effects, mis-specified dynamics, or non-stationarity of the variables, in particular for the capital input variable. It should come as no surprise that the adjustment of material input is far easier to model within the framework considered than is capital.

Acknowledgements: I thank Harald Goldstein and Bernt P. Stigum for helpful comments, which led to improvement in the organization of the paper. The usual disclaimer applies.

Notes

[1]Identification under non-normality of the true regressor is, however, possible by utilizing moments of the distribution of the observable variables of order higher than the second [see Reiersøl (1950)]. Even under non-identification, bounds on the parameters can be established from the distribution of the observable variables [see Fuller (1987, p. 11)]. These bounds may be wide or narrow, depending on the covariance structure of the variables; see, *e.g.*, Klepper and Leamer (1984) and Bekker *et al.* (1987).

[2]The last two assumptions are stronger than strictly needed; time invariance of $\mathsf{E}(\alpha_i \boldsymbol{v}_{it})$ and $\mathsf{E}(\alpha_i u_{it})$ is sufficient. A modification to this effect will be of minor practical importance, however.

[3]Premultiplication of (4) by $\boldsymbol{d}_{t\theta}$ is not the only way of eliminating α_i. Any $(1 \times T)$ vector $\boldsymbol{c}_{t\theta}$ such that $\boldsymbol{c}_{t\theta}\boldsymbol{e}_T = 0$, for example the rows of the within individual transformation matrix $\boldsymbol{I}_T - \boldsymbol{e}_T\boldsymbol{e}_T'/T$, where $\boldsymbol{I}_T$ is the T dimensional identity matrix, has this property.

[4]Here and in the following plim always denotes probability limits when N goes to infinity and T is finite.

[5]We report no standard error estimates in Table 24.1, since some of the methods are inconsistent.

[6]The OC's involving y's can be treated similarly. Essential and redundant moment conditions in the context of AR models for panel data are discussed in, *inter alia*, Ahn and Schmidt (1995), Arellano and Bover (1995), and Blundell and Bond (2000). This problem resembles, in some respects, the problem for static measurement error models discussed here.

[7]All numerical calculations are performed by means of procedures constructed by the author in the GAUSS software code.

Appendix

In this appendix, we elaborate the procedures for estimating asymptotic covariance matrices of the GMM estimators. All models in the main text, with suitable interpretations of $\boldsymbol{y}$, $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{\epsilon}$, and $\boldsymbol{\Omega}$, have the form:

$$\text{(A.1)} \qquad \boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \qquad \mathsf{E}(\boldsymbol{\epsilon}) = \boldsymbol{0}, \qquad \mathsf{E}(\boldsymbol{Z}'\boldsymbol{\epsilon}) = \boldsymbol{0}, \qquad \mathsf{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = \boldsymbol{\Omega},$$

where $\boldsymbol{y} = (\boldsymbol{y}_1', \dots, \boldsymbol{y}_N')'$, $\boldsymbol{X} = (\boldsymbol{X}_1', \dots, \boldsymbol{X}_N')'$, $\boldsymbol{Z} = (\boldsymbol{Z}_1', \dots, \boldsymbol{Z}_N')'$, and $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1', \dots, \boldsymbol{\epsilon}_N')'$, $\boldsymbol{Z}_i$ being the IV matrix of $\boldsymbol{X}_i$. The two generic GMM estimators considered are

$$\text{(A.2)} \qquad \widehat{\boldsymbol{\beta}} = [\boldsymbol{X}'\boldsymbol{P}_Z\boldsymbol{X}]^{-1}[\boldsymbol{X}'\boldsymbol{P}_Z\boldsymbol{y}], \qquad \boldsymbol{P}_Z = \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}',$$

$$\text{(A.3)} \qquad \widetilde{\boldsymbol{\beta}} = [\boldsymbol{X}'\boldsymbol{P}_Z(\Omega)\boldsymbol{X}]^{-1}[\boldsymbol{X}'\boldsymbol{P}_Z(\Omega)\boldsymbol{y}], \qquad \boldsymbol{P}_Z(\Omega) = \boldsymbol{Z}(\boldsymbol{Z}'\Omega\boldsymbol{Z})^{-1}\boldsymbol{Z}'.$$

Let the residual vector calculated from $\widehat{\boldsymbol{\beta}}$ be $\widehat{\boldsymbol{\epsilon}} = \boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}$, and use the notation

$$\boldsymbol{S}_{XZ} = \frac{\boldsymbol{X}'\boldsymbol{Z}}{N}, \quad \boldsymbol{S}_{ZX} = \frac{\boldsymbol{Z}'\boldsymbol{X}}{N}, \quad \boldsymbol{S}_{ZZ} = \frac{\boldsymbol{Z}'\boldsymbol{Z}}{N}, \quad \boldsymbol{S}_{\epsilon Z} = \frac{\boldsymbol{\epsilon}'\boldsymbol{Z}}{N}, \quad \boldsymbol{S}_{Z\epsilon} = \frac{\boldsymbol{Z}'\boldsymbol{\epsilon}}{N},$$
$$\boldsymbol{S}_{Z\boldsymbol{\Omega}Z} = \frac{\boldsymbol{Z}'\Omega\boldsymbol{Z}}{N}, \quad \boldsymbol{S}_{Z\boldsymbol{\epsilon}\boldsymbol{\epsilon}Z} = \frac{\boldsymbol{Z}'\boldsymbol{\epsilon}\boldsymbol{\epsilon}'\boldsymbol{Z}}{N}, \quad \boldsymbol{S}_{Z\widehat{\boldsymbol{\epsilon}}\widehat{\boldsymbol{\epsilon}}Z} = \frac{\boldsymbol{Z}'\widehat{\boldsymbol{\epsilon}}\widehat{\boldsymbol{\epsilon}}'\boldsymbol{Z}}{N}.$$

Inserting for $\boldsymbol{y}$ from (A.1) in (A.2) and (A.3), we get

$$\sqrt{N}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sqrt{N}[\boldsymbol{X}'\boldsymbol{P}_Z\boldsymbol{X}]^{-1}[\boldsymbol{X}'\boldsymbol{P}_Z\boldsymbol{\epsilon}] = [\boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}]^{-1}\left[\boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\frac{\boldsymbol{Z}'\boldsymbol{\epsilon}}{\sqrt{N}}\right],$$
$$\sqrt{N}(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sqrt{N}[\boldsymbol{X}'\boldsymbol{P}_Z(\boldsymbol{\Omega})\boldsymbol{X}]^{-1}[\boldsymbol{X}'\boldsymbol{P}_Z(\boldsymbol{\Omega})\boldsymbol{\epsilon}] = [\boldsymbol{S}_{XZ}\boldsymbol{S}_{Z\Omega Z}^{-1}\boldsymbol{S}_{ZX}]^{-1}\left[\boldsymbol{S}_{XZ}\boldsymbol{S}_{Z\Omega Z}^{-1}\frac{\boldsymbol{Z}'\boldsymbol{\epsilon}}{\sqrt{N}}\right],$$

and hence,

$$N(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})' = [\boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}]^{-1}[\boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{Z\boldsymbol{\epsilon}\boldsymbol{\epsilon}Z}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}][\boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}]^{-1},$$
$$N(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})' = [\boldsymbol{S}_{XZ}\boldsymbol{S}_{Z\Omega Z}^{-1}\boldsymbol{S}_{ZX}]^{-1}[\boldsymbol{S}_{XZ}\boldsymbol{S}_{Z\Omega Z}^{-1}\boldsymbol{S}_{Z\boldsymbol{\epsilon}\boldsymbol{\epsilon}Z}\boldsymbol{S}_{Z\Omega Z}^{-1}\boldsymbol{S}_{ZX}][\boldsymbol{S}_{XZ}\boldsymbol{S}_{Z\Omega Z}^{-1}\boldsymbol{S}_{ZX}]^{-1}.$$

The asymptotic covariance matrices of $\sqrt{N}\widehat{\boldsymbol{\beta}}$ and $\sqrt{N}\widetilde{\boldsymbol{\beta}}$ can then, under suitable regularity conditions, be written as [see Bowden and Turkington (1984, pp. 26, 69)]

$$a\mathsf{V}(\sqrt{N}\,\widehat{\boldsymbol{\beta}}) = \lim \mathsf{E}[N(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'] = \text{plim}[N(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'],$$
$$a\mathsf{V}(\sqrt{N}\,\widetilde{\boldsymbol{\beta}}) = \lim \mathsf{E}[N(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'] = \text{plim}[N(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'].$$

Since $\boldsymbol{S}_{Z\boldsymbol{\epsilon\epsilon}Z}$ and $\boldsymbol{S}_{Z\Omega Z}$ coincide asymptotically, we get, using bars to denote plims,

$$\text{(A.4)}\quad a\mathsf{V}(\sqrt{N}\,\widehat{\boldsymbol{\beta}}) = [\bar{\boldsymbol{S}}_{XZ}\bar{\boldsymbol{S}}_{ZZ}^{-1}\bar{\boldsymbol{S}}_{ZX}]^{-1}[\bar{\boldsymbol{S}}_{XZ}\bar{\boldsymbol{S}}_{ZZ}^{-1}\bar{\boldsymbol{S}}_{Z\Omega Z}\bar{\boldsymbol{S}}_{ZZ}^{-1}\bar{\boldsymbol{S}}_{ZX}][\bar{\boldsymbol{S}}_{XZ}\bar{\boldsymbol{S}}_{ZZ}^{-1}\bar{\boldsymbol{S}}_{ZX}]^{-1},$$

$$\text{(A.5)}\quad a\mathsf{V}(\sqrt{N}\,\widetilde{\boldsymbol{\beta}}) = [\bar{\boldsymbol{S}}_{XZ}\bar{\boldsymbol{S}}_{Z\Omega Z}^{-1}\bar{\boldsymbol{S}}_{ZX}]^{-1}.$$

Replacing the plims $\bar{\boldsymbol{S}}_{XZ}$, $\bar{\boldsymbol{S}}_{ZX}$, $\bar{\boldsymbol{S}}_{ZZ}$ and $\bar{\boldsymbol{S}}_{Z\Omega Z}$ by their sample counterparts, $\boldsymbol{S}_{XZ}$, $\boldsymbol{S}_{ZX}$, $\boldsymbol{S}_{ZZ}$ and $\boldsymbol{S}_{Z\hat{\epsilon}\hat{\epsilon}Z}$ and dividing by N, we get from (A.4) and (A.5) the following estimators of the asymptotic covariance matrices of $\widehat{\boldsymbol{\beta}}$ and $\widetilde{\boldsymbol{\beta}}$:

$$\begin{aligned}\text{(A.6)}\quad \widehat{\mathsf{V}(\widehat{\boldsymbol{\beta}})} &= \frac{1}{N}[\boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}]^{-1}[\boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{Z\hat{\boldsymbol{\epsilon}}\hat{\boldsymbol{\epsilon}}Z}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}][\boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}]^{-1}\\ &= [\boldsymbol{X}'\boldsymbol{P}_Z\boldsymbol{X}]^{-1}[\boldsymbol{X}'\boldsymbol{P}_Z\widehat{\boldsymbol{\epsilon}}\widehat{\boldsymbol{\epsilon}}'\boldsymbol{P}_Z\boldsymbol{X}][\boldsymbol{X}'\boldsymbol{P}_Z\boldsymbol{X}]^{-1},\\ \text{(A.7)}\quad \widehat{\mathsf{V}(\widetilde{\boldsymbol{\beta}})} &= \frac{1}{N}[\boldsymbol{S}_{XZ}\boldsymbol{S}_{Z\hat{\boldsymbol{\epsilon}}\hat{\boldsymbol{\epsilon}}Z}^{-1}\boldsymbol{S}_{ZX}]^{-1}\\ &= [\boldsymbol{X}'\boldsymbol{Z}(\boldsymbol{Z}'\widehat{\boldsymbol{\epsilon}}\widehat{\boldsymbol{\epsilon}}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{X}]^{-1} = [\boldsymbol{X}'\boldsymbol{P}_Z(\widehat{\boldsymbol{\epsilon}}\widehat{\boldsymbol{\epsilon}}')\boldsymbol{X}]^{-1}.\end{aligned}$$

These are the generic expressions which we use for estimating variances and covariances of the GMM estimators considered.

When calculating $\widetilde{\boldsymbol{\beta}}$ from (A.3) in practice, we replace $\boldsymbol{P}_Z(\boldsymbol{\Omega})$ by $\boldsymbol{P}_Z(\widehat{\boldsymbol{\epsilon}}\widehat{\boldsymbol{\epsilon}}') = \boldsymbol{Z}(\boldsymbol{Z}'\widehat{\boldsymbol{\epsilon}}\widehat{\boldsymbol{\epsilon}}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'$ [see White (1982, 1984)].

References

Ahn, S.C., and Schmidt, P. (1995): Efficient Estimation of Models for Dynamic Panel Data. *Journal of Econometrics*, 68, 5-27.

Arellano, M., and Bond, S. (1991): Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58, 277-297.

Arellano, M., and Bover, O. (1995): Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics*, 68, 29-51.

Baltagi, B.H. (2001): *Econometric Analysis of Panel Data*, second edition. Chichester: Wiley.

Bekker, P., Kapteyn, A., and Wansbeek, T. (1987): Consistent Sets of Estimates for Regressions with Correlated or Uncorrelated Measurement Errors in Arbitrary Subsets of All Variables. *Econometrica*, 55, 1223-1230.

Biørn, E. (1992): The Bias of Some Estimators for Panel Data Models with Measurement Errors. *Empirical Economics*, 17, 51-66.

Biørn, E. (1996): Panel Data with Measurement Errors. Chapter 10 in *The Econometrics of Panel Data. Handbook of the Theory with Applications*, ed. by L. Mátyás and P. Sevestre. Dordrecht: Kluwer.

Biørn, E. (2000): Panel Data with Measurement Errors. Instrumental Variables and GMM Estimators Combining Levels and Differences. *Econometric Reviews*, 19, 391-424.

Biørn, E., and Klette, T.J. (1998): Panel Data with Errors-in-Variables: Essential and Redundant Orthogonality Conditions in GMM-Estimation. *Economics Letters*, 59, 275-282.

Biørn, E., and Klette, T.J. (1999): The Labour Input Response to Permanent Changes in Output: An Errors in Variables Analysis Based on Panel Data. *Scandinavian Journal of Economics*, 101, 379-404.

Blundell, R., and Bond, S. (1998): Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*, 87, 115-143.

Bowden, R.J., and Turkington, D.A. (1984): *Instrumental Variables*. Cambridge: Cambridge University Press.

Davidson, R., and MacKinnon, J.G. (1993): *Estimation and Inference in Econometrics*. Oxford: Oxford University Press.

Frisch, R. (1934): *Statistical Confluence Analysis by Means of Complete Regression Systems*. Oslo: Universitetets Økonomiske Institutt.

Fuller, W.A. (1987): *Measurement Error Models*. New York: Wiley.

Griliches, Z., and Hausman, J.A. (1986): Errors in Variables in Panel Data. *Journal of Econometrics*, 31, 93-118.

Hansen, L.P. (1982): Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50, 1029-1054.

Harris, D. and Mátyás, L. (1999): Introduction to the Generalized Method of Moments Estimation. Chapter 1 in *Generalized Method of Moments Estimation*, ed. by L. Mátyás. Cambridge: Cambridge University Press.

Hsiao, C. (1986): *Analysis of Panel Data*. Cambridge: Cambridge University Press.

Jorgenson, D.W. (1986): Econometric Methods for Modeling Producer Behavior. Chapter 31 in *Handbook of Econometrics, vol. III*, ed. by Z. Griliches and M.D. Intriligator. Amsterdam, North-Holland.

Klepper, S., and Leamer, E. (1984): Consistent Sets of Estimates for Regressions with Errors in All Variables. *Econometrica*, 52, 163-183.

McCabe, B., and Tremayne, A. (1993): *Elements of Modern Asymptotic Theory with Statistical Applications*. Manchester: Manchester University Press.

Nelson, C.R., and Startz, R. (1990): Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator. *Econometrica*, 58, 967-976.

Newey, W.K. (1985): Generalized Method of Moments Specification Testing. *Journal of Econometrics*, 29, 229-256.

Reiersøl, O. (1950): Identifiability of a Linear Relation Between Variables which are Subject to Error. *Econometrica*, 18, 375-389.

Staiger, D., and Stock, J.H. (1997): Instrumental Variables Regression With Weak Instruments. *Econometrica*, 65, 557-586.

Stigum, B. (1995): Theory-Data Confrontations in Economics. *Dialogue*, 34, 581-604.

Wansbeek, T.J. (2001): GMM Estimation in Panel Data Models with Measurement Error. *Journal of Econometrics*, 104, 259-268.

Wansbeek, T.J., and Koning, R.H. (1991): Measurement Error and Panel Data. *Statistica Neerlandica*, 45, 85-92.

White, H. (1982): Instrumental Variables Regression with Independent Observations. *Econometrica*, 50, 483-499.

White, H. (1984): *Asymptotic Theory for Econometricians*. Orlando: Academic Press.

Table 24.1:

INPUT ELASTICITIES AND INVERSE INPUT ELASTICITIES.

STANDARD OLS, BETWEEN PERIOD, AND WITHIN FIRM ESTIMATES

Q = output, M = materials, K = capital

Cols. 1 – 3: Equation in levels.

Cols. 4 – 6: Equation in differences, with intercept.

Cols. 7 – 8: Equation in differences, without intercept

Textiles: $N = 215, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{OLS}$	$\widehat{\boldsymbol{\beta}}_{BP}$	$\widehat{\boldsymbol{\beta}}_{WF}$	$\widehat{\boldsymbol{\beta}}_{OLSDC}$	$\widehat{\boldsymbol{\beta}}_{BPDC}$	$\widehat{\boldsymbol{\beta}}_{WFDC}$	$\widehat{\boldsymbol{\beta}}_{\Delta 7}$	$\widehat{\boldsymbol{\beta}}_{OLSD}$
$\ln M,\ \ln Q$	1.1450	1.0028	1.1033	1.1683	1.0935	1.1750	1.0204	1.1608
$\ln Q,\ \ln M$	0.7889	0.9859	0.7005	0.5786	0.8742	0.5696	0.9800	0.5894
$\ln K,\ \ln Q$	0.9899	1.0351	0.6081	0.1621	-0.0170	0.1099	1.5176	0.2313
$\ln Q,\ \ln K$	0.6751	0.6018	0.3281	0.0852	-0.1584	0.0563	0.6589	0.1203

Wood and Wood Products: $N = 603, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{OLS}$	$\widehat{\boldsymbol{\beta}}_{BP}$	$\widehat{\boldsymbol{\beta}}_{WF}$	$\widehat{\boldsymbol{\beta}}_{OLSDC}$	$\widehat{\boldsymbol{\beta}}_{BPDC}$	$\widehat{\boldsymbol{\beta}}_{WFDC}$	$\widehat{\boldsymbol{\beta}}_{\Delta 7}$	$\widehat{\boldsymbol{\beta}}_{OLSD}$
$\ln M,\ \ln Q$	1.0940	1.0535	1.0747	1.1062	1.0635	1.1106	1.0779	1.1046
$\ln Q,\ \ln M$	0.8940	0.9477	0.8127	0.6981	0.9290	0.6869	0.9277	0.7078
$\ln K,\ \ln Q$	0.9843	1.3865	0.6858	0.2111	0.1081	0.1855	1.7914	0.3004
$\ln Q,\ \ln K$	0.7816	0.5566	0.3476	0.1272	0.8143	0.1089	0.5582	0.1719

Paper and Paper Products: $N = 600, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{OLS}$	$\widehat{\boldsymbol{\beta}}_{BP}$	$\widehat{\boldsymbol{\beta}}_{WF}$	$\widehat{\boldsymbol{\beta}}_{OLSDC}$	$\widehat{\boldsymbol{\beta}}_{BPDC}$	$\widehat{\boldsymbol{\beta}}_{WFDC}$	$\widehat{\boldsymbol{\beta}}_{\Delta 7}$	$\widehat{\boldsymbol{\beta}}_{OLSD}$
$\ln M,\ \ln Q$	1.0809	1.0867	1.0759	1.0687	1.0630	1.0664	1.0964	1.0728
$\ln Q,\ \ln M$	0.8935	0.9194	0.7656	0.5560	0.9088	0.5410	0.9120	0.5907
$\ln K,\ \ln Q$	0.9711	1.4169	0.9815	0.3001	0.3790	0.2611	1.5207	0.4801
$\ln Q,\ \ln K$	0.8141	0.6527	0.3757	0.0957	1.0722	0.0812	0.6576	0.1593

Chemicals: $N = 229, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{OLS}$	$\widehat{\boldsymbol{\beta}}_{BP}$	$\widehat{\boldsymbol{\beta}}_{WF}$	$\widehat{\boldsymbol{\beta}}_{OLSDC}$	$\widehat{\boldsymbol{\beta}}_{BPDC}$	$\widehat{\boldsymbol{\beta}}_{WFDC}$	$\widehat{\boldsymbol{\beta}}_{\Delta 7}$	$\widehat{\boldsymbol{\beta}}_{OLSD}$
$\ln M,\ \ln Q$	1.0337	1.0228	1.0275	1.0522	0.9922	1.0573	1.0167	1.0488
$\ln Q,\ \ln M$	0.9484	0.9764	0.8443	0.6922	0.9644	0.6770	0.9836	0.7121
$\ln K,\ \ln Q$	1.0499	1.3520	0.8164	0.1929	0.5105	0.1456	1.3626	0.3069
$\ln Q,\ \ln K$	0.8175	0.7071	0.4447	0.1186	0.8560	0.0883	0.7339	0.1861

Table 24.2:

INPUT ELASTICITIES AND INVERSE INPUT ELASTICITIES.

GMM ESTIMATES OF DIFFERENCED EQUATIONS, WITH ALL IV'S IN LEVELS

Q = output, M = materials, K = capital

In parenthesis: Cols. 1 – 5: Standard deviation estimates. Cols. 6 – 7: p values

Textiles: $N = 215, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Dx}$	$\widehat{\boldsymbol{\beta}}_{Dy}$	$\widehat{\boldsymbol{\beta}}_{Dxy}$	$\widetilde{\boldsymbol{\beta}}_{Dx}$	$\widetilde{\boldsymbol{\beta}}_{Dy}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Dx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Dy})$
$\ln M, \ln Q$	1.0821	1.1275	1.0900	1.0546	1.0825	51.71	70.39
	(0.0331)	(0.0346)	(0.0350)	(0.0173)	(0.0169)	(0.2950)	(0.0152)
$\ln Q, \ln M$	0.8404	0.8931	0.8064	0.8917	0.9244	86.55	59.08
	(0.0283)	(0.0283)	(0.0363)	(0.0143)	(0.0148)	(0.0004)	(0.1112)
$\ln K, \ln Q$	0.5095	0.6425	0.5004	0.5239	0.6092	115.68	121.29
	(0.0735)	(0.0700)	(0.0745)	(0.0407)	(0.0314)	(0.0000)	(0.0000)
$\ln Q, \ln K$	0.4170	0.6391	0.4021	0.4499	0.6495	130.50	133.94
	(0.0409)	(0.0561)	(0.0382)	(0.0248)	(0.0330)	(0.0000)	(0.0000)

Wood and Wood Products: $N = 603, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Dx}$	$\widehat{\boldsymbol{\beta}}_{Dy}$	$\widehat{\boldsymbol{\beta}}_{Dxy}$	$\widetilde{\boldsymbol{\beta}}_{Dx}$	$\widetilde{\boldsymbol{\beta}}_{Dy}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Dx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Dy})$
$\ln M, \ln Q$	1.0604	1.0784	1.0632	1.0615	1.0772	63.97	90.28
	(0.0123)	(0.0124)	(0.0128)	(0.0089)	(0.0098)	(0.0502)	(0.0002)
$\ln Q, \ln M$	0.9171	0.9362	0.9117	0.9195	0.9370	91.40	64.13
	(0.0106)	(0.0108)	(0.0115)	(0.0083)	(0.0078)	(0.0001)	(0.0489)
$\ln K, \ln Q$	0.7454	0.8906	0.7494	0.8094	0.9398	290.60	281.57
	(0.0409)	(0.0439)	(0.0425)	(0.0305)	(0.0310)	(0.0000)	(0.0000)
$\ln Q, \ln K$	0.4862	0.6003	0.4806	0.5261	0.6377	283.25	280.65
	(0.0229)	(0.0258)	(0.0223)	(0.0189)	(0.0212)	(0.0000)	(0.0000)

Paper and Paper Products: $N = 600, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Dx}$	$\widehat{\boldsymbol{\beta}}_{Dy}$	$\widehat{\boldsymbol{\beta}}_{Dxy}$	$\widetilde{\boldsymbol{\beta}}_{Dx}$	$\widetilde{\boldsymbol{\beta}}_{Dy}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Dx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Dy})$
$\ln M, \ln Q$	1.0766	1.1102	1.0726	1.0680	1.0820	43.12	81.97
	(0.0150)	(0.0162)	(0.0155)	(0.0119)	(0.0123)	(0.6340)	(0.0012)
$\ln Q, \ln M$	0.8847	0.9204	0.8853	0.9119	0.9301	90.18	44.50
	(0.0140)	(0.0131)	(0.0145)	(0.0101)	(0.0102)	(0.0002)	(0.5769)
$\ln K, \ln Q$	1.0713	1.2134	1.0818	1.0854	1.2543	193.21	220.93
	(0.0430)	(0.0477)	(0.0435)	(0.0324)	(0.0398)	(0.0000)	(0.0000)
$\ln Q, \ln K$	0.5591	0.7048	0.5559	0.5377	0.7075	225.95	193.33
	(0.0198)	(0.0243)	(0.0198)	(0.0170)	(0.0198)	(0.0000)	(0.0000)

Chemicals: $N = 229, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Dx}$	$\widehat{\boldsymbol{\beta}}_{Dy}$	$\widehat{\boldsymbol{\beta}}_{Dxy}$	$\widetilde{\boldsymbol{\beta}}_{Dx}$	$\widetilde{\boldsymbol{\beta}}_{Dy}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Dx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Dy})$
$\ln M, \ln Q$	1.0166	1.0540	1.0263	1.0009	1.0394	54.29	81.64
	(0.0245)	(0.0241)	(0.0251)	(0.0135)	(0.0138)	(0.2166)	(0.0013)
$\ln Q, \ln M$	0.9205	0.9609	0.8972	0.9323	0.9815	87.10	57.90
	(0.0230)	(0.0239)	(0.0231)	(0.0122)	(0.0130)	(0.0003)	(0.1324)
$\ln K, \ln Q$	0.9706	1.2497	0.9579	1.0051	1.2672	90.42	85.36
	(0.0583)	(0.0633)	(0.0582)	(0.0336)	(0.0489)	(0.0001)	(0.0005)
$\ln Q, \ln K$	0.5550	0.7459	0.5637	0.5700	0.7762	96.70	89.57
	(0.0317)	(0.0374)	(0.0314)	(0.0236)	(0.0273)	(0.0000)	(0.0002)

Table 24.3:

INPUT ELASTICITIES AND INVERSE INPUT ELASTICITIES.

GMM ESTIMATES OF LEVEL EQUATIONS, WITH ALL IV'S IN DIFFERENCES.

NO MEAN DEDUCTION

Q = output, M = materials, K = capital

In parenthesis: Cols. 1 – 5: Standard deviation estimates. Cols. 6 – 7: p values

Textiles: $N = 215, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Lx}$	$\widehat{\boldsymbol{\beta}}_{Ly}$	$\widehat{\boldsymbol{\beta}}_{Lxy}$	$\widetilde{\boldsymbol{\beta}}_{Lx}$	$\widetilde{\boldsymbol{\beta}}_{Ly}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Lx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Ly})$
$\ln M, \ln Q$	0.9308	0.9325	0.9274	0.9351	0.9404	56.76	81.49
	(0.0031)	(0.0052)	(0.0036)	(0.0024)	(0.0022)	(0.1557)	(0.0013)
$\ln Q, \ln M$	1.0718	1.0743	1.0772	1.0628	1.0690	80.64	56.69
	(0.0060)	(0.0035)	(0.0039)	(0.0025)	(0.0028)	(0.0016)	(0.1572)
$\ln K, \ln Q$	0.7408	0.7355	0.7381	0.7505	0.7502	107.05	116.19
	(0.0079)	(0.0079)	(0.0072)	(0.0059)	(0.0055)	(0.0000)	(0.0000)
$\ln Q, \ln K$	1.3533	1.3483	1.3490	1.3211	1.3231	115.18	106.84
	(0.0145)	(0.0144)	(0.0129)	(0.0097)	(0.0105)	(0.0000)	(0.0000)

Wood and Wood Products: $N = 603, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Lx}$	$\widehat{\boldsymbol{\beta}}_{Ly}$	$\widehat{\boldsymbol{\beta}}_{Lxy}$	$\widetilde{\boldsymbol{\beta}}_{Lx}$	$\widetilde{\boldsymbol{\beta}}_{Ly}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Lx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Ly})$
$\ln M, \ln Q$	0.9473	0.9469	0.9471	0.9484	0.9496	141.10	159.95
	(0.0011)	(0.0011)	(0.0011)	(0.0010)	(0.0010)	(0.0000)	(0.0000)
$\ln Q, \ln M$	1.0561	1.0557	1.0558	1.0529	1.0543	159.80	141.07
	(0.0013)	(0.0012)	(0.0012)	(0.0011)	(0.0011)	(0.0000)	(0.0000)
$\ln K, \ln Q$	0.7545	0.7560	0.7546	0.7598	0.7699	207.64	272.33
	(0.0030)	(0.0033)	(0.0029)	(0.0027)	(0.0029)	(0.0000)	(0.0000)
$\ln Q, \ln K$	1.3197	1.3244	1.3221	1.2927	1.3124	270.29	207.00
	(0.0056)	(0.0053)	(0.0050)	(0.0049)	(0.0046)	(0.0000)	(0.0000)

Paper and Paper Products: $N = 600, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Lx}$	$\widehat{\boldsymbol{\beta}}_{Ly}$	$\widehat{\boldsymbol{\beta}}_{Lxy}$	$\widetilde{\boldsymbol{\beta}}_{Lx}$	$\widetilde{\boldsymbol{\beta}}_{Ly}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Lx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Ly})$
$\ln M, \ln Q$	0.9301	0.9300	0.9301	0.9304	0.9347	140.14	150.10
	(0.0015)	(0.0016)	(0.0015)	(0.0013)	(0.0013)	(0.0000)	(0.0000)
$\ln Q, \ln M$	1.0751	1.0751	1.0749	1.0695	1.0744	149.82	140.18
	(0.0019)	(0.0017)	(0.0017)	(0.0015)	(0.0015)	(0.0000)	(0.0000)
$\ln K, \ln Q$	0.7703	0.7658	0.7692	0.7761	0.7745	196.22	254.48
	(0.0033)	(0.0034)	(0.0031)	(0.0028)	(0.0029)	(0.0000)	(0.0000)
$\ln Q, \ln K$	1.3025	1.2974	1.2970	1.2848	1.2850	252.95	195.79
	(0.0057)	(0.0055)	(0.0051)	(0.0048)	(0.0046)	(0.0000)	(0.0000)

Chemicals: $N = 229, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Lx}$	$\widehat{\boldsymbol{\beta}}_{Ly}$	$\widehat{\boldsymbol{\beta}}_{Lxy}$	$\widetilde{\boldsymbol{\beta}}_{Lx}$	$\widetilde{\boldsymbol{\beta}}_{Ly}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Lx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Ly})$
$\ln M, \ln Q$	0.9521	0.9518	0.9520	0.9532	0.9535	53.01	87.76
	(0.0015)	(0.0015)	(0.0015)	(0.0012)	(0.0013)	(0.2537)	(0.0003)
$\ln Q, \ln M$	1.0506	1.0503	1.0503	1.0486	1.0490	87.69	52.98
	(0.0017)	(0.0017)	(0.0016)	(0.0014)	(0.0014)	(0.0003)	(0.2544)
$\ln K, \ln Q$	0.7877	0.7886	0.7884	0.7881	0.7994	96.57	117.54
	(0.0046)	(0.0048)	(0.0045)	(0.0040)	(0.0037)	(0.0000)	(0.0000)
$\ln Q, \ln K$	1.2662	1.2686	1.2659	1.2470	1.2652	117.00	96.55
	(0.0077)	(0.0074)	(0.0072)	(0.0058)	(0.0064)	(0.0000)	(0.0000)

Table 24.4:

INPUT ELASTICITIES AND INVERSE INPUT ELASTICITIES.

GMM ESTIMATES OF LEVEL EQUATIONS, WITH ALL IV'S IN DIFFERENCES.

WITH MEAN DEDUCTION

Q = output, M = materials, K = capital

In parenthesis: Cols. 1 – 5: Standard deviation estimates. Cols. 6 – 7: p values

Textiles: $N = 215, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Lx}$	$\widehat{\boldsymbol{\beta}}_{Ly}$	$\widehat{\boldsymbol{\beta}}_{Lxy}$	$\widetilde{\boldsymbol{\beta}}_{Lx}$	$\widetilde{\boldsymbol{\beta}}_{Ly}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Lx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Ly})$
$\ln M, \ln Q$	1.0219	1.2148	1.1881	1.0739	1.1749	54.66	73.56
	(0.0644)	(0.1202)	(0.0786)	(0.0289)	(0.0316)	(0.2065)	(0.0079)
$\ln Q, \ln M$	0.7345	0.9392	0.7048	0.7428	0.8834	64.48	52.42
	(0.0730)	(0.0559)	(0.0621)	(0.0225)	(0.0242)	(0.0460)	(0.2720)
$\ln K, \ln Q$	1.0348	1.2201	1.0776	0.7504	1.3279	84.43	76.36
	(0.1471)	(0.1514)	(0.1153)	(0.0703)	(0.0808)	(0.0007)	(0.0043)
$\ln Q, \ln K$	0.5967	0.7045	0.5902	0.4599	0.6675	69.04	94.75
	(0.0755)	(0.1190)	(0.0682)	(0.0322)	(0.0546)	(0.0198)	(0.0000)

Wood and Wood Products: $N = 603, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Lx}$	$\widehat{\boldsymbol{\beta}}_{Ly}$	$\widehat{\boldsymbol{\beta}}_{Lxy}$	$\widetilde{\boldsymbol{\beta}}_{Lx}$	$\widetilde{\boldsymbol{\beta}}_{Ly}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Lx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Ly})$
$\ln M, \ln Q$	1.0501	1.1174	1.0813	1.0646	1.1328	63.26	65.06
	(0.0219)	(0.0245)	(0.0235)	(0.0140)	(0.0188)	(0.0567)	(0.0415)
$\ln Q, \ln M$	0.8740	0.9425	0.8888	0.8644	0.9277	62.69	62.27
	(0.0192)	(0.0189)	(0.0194)	(0.0145)	(0.0123)	(0.0625)	(0.0671)
$\ln K, \ln Q$	0.6696	1.4487	0.8460	0.4414	1.4470	100.40	126.86
	(0.0927)	(0.1615)	(0.0794)	(0.0489)	(0.1093)	(0.0000)	(0.0000)
$\ln Q, \ln K$	0.5188	0.7927	0.5363	0.3165	0.9208	102.10	149.90
	(0.0655)	(0.0905)	(0.0546)	(0.0339)	(0.0617)	(0.0000)	(0.0000)

Paper and Paper Products: $N = 600, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Lx}$	$\widehat{\boldsymbol{\beta}}_{Ly}$	$\widehat{\boldsymbol{\beta}}_{Lxy}$	$\widetilde{\boldsymbol{\beta}}_{Lx}$	$\widetilde{\boldsymbol{\beta}}_{Ly}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Lx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Ly})$
$\ln M, \ln Q$	1.0797	1.0883	1.0766	1.0799	1.1376	42.95	83.36
	(0.0242)	(0.0410)	(0.0216)	(0.0185)	(0.0301)	(0.6408)	(0.0009)
$\ln Q, \ln M$	0.8911	0.9172	0.8911	0.8271	0.9124	79.18	43.00
	(0.0334)	(0.0209)	(0.0185)	(0.0233)	(0.0158)	(0.0023)	(0.6388)
$\ln K, \ln Q$	0.9242	1.2121	0.9624	0.8171	1.2018	59.67	158.86
	(0.0641)	(0.1117)	(0.0540)	(0.0427)	(0.0791)	(0.1017)	(0.0000)
$\ln Q, \ln K$	0.5953	1.0319	0.7451	0.3715	1.0560	141.11	62.95
	(0.0635)	(0.0711)	(0.0444)	(0.0321)	(0.0506)	(0.0000)	(0.0598)

Chemicals: $N = 229, T = 8$

y, x	$\widehat{\boldsymbol{\beta}}_{Lx}$	$\widehat{\boldsymbol{\beta}}_{Ly}$	$\widehat{\boldsymbol{\beta}}_{Lxy}$	$\widetilde{\boldsymbol{\beta}}_{Lx}$	$\widetilde{\boldsymbol{\beta}}_{Ly}$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Lx})$	$\chi^2(\widetilde{\boldsymbol{\beta}}_{Ly})$
$\ln M, \ln Q$	0.9721	1.0217	0.9950	0.9805	1.0253	55.85	83.10
	(0.0269)	(0.0278)	(0.0225)	(0.0179)	(0.0179)	(0.1765)	(0.0009)
$\ln Q, \ln M$	0.9619	1.0196	0.9760	0.9429	0.9992	81.75	55.71
	(0.0262)	(0.0279)	(0.0214)	(0.0159)	(0.0179)	(0.0013)	(0.1798)
$\ln K, \ln Q$	1.1013	1.4280	1.1151	0.9795	1.4408	68.96	69.82
	(0.0692)	(0.1429)	(0.0623)	(0.0465)	(0.0838)	(0.0201)	(0.0170)
$\ln Q, \ln K$	0.6348	0.8281	0.7261	0.5150	0.8536	67.88	71.83
	(0.0680)	(0.0550)	(0.0428)	(0.0355)	(0.0390)	(0.0247)	(0.0113)