OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 67, 3 (2005) 0305-9049

PRACTITIONERS' CORNER

Correcting Standard Errors in Two-stage Estimation Procedures with Generated Regressands*

MICHEL DUMONT†, GLENN RAYP‡, OLIVIER THAS§ and PETER WILLEMÉ¶

†*Faculty of Applied Economics, University of Antwerp, Antwerp, Belgium (e-mail: michel.dumont@ua.ac.be)*
‡*Department of Economics, Faculty of Economics and Business Administration, Ghent University, SHERPPA, Ghent, Belgium (e-mail: glenn.rayp@ugent.be)*
§*Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Ghent, Belgium (e-mail: olivier.thas@ugent.be)*
¶*Faculty of Applied Economics, University of Antwerp and the Federal Planning Bureau, Belgium (e-mail: peter.willeme@ua.ac.be)*

Abstract

Feenstra and Hanson [NBER Working Paper No. 6052 (1997)] propose a procedure to correct the standard errors in a two-stage regression with generated dependent variables. Their method has subsequently been used in two-stage mandated wage models [Feenstra and Hanson, *Quarterly Journal of Economics* (1999) Vol. 114, pp. 907–940; Haskel and Slaughter, *The Economic Journal* (2001) Vol. 111, pp. 163–187; *Review of International Economics* (2003) Vol. 11, pp. 630–650] and for the estimation of the sector bias of skill-biased technological change [Haskel and Slaughter, *European Economic Review* (2002) Vol. 46, pp. 1757–1783]. Unfortunately, the proposed correction is negatively biased (sometimes even resulting in negative estimated variances) and therefore leads to overestimation of the inferred significance. We

*The authors gratefully acknowledge the useful comments of Adrian Pagan, an anonymous referee and the editor. Any remaining errors are ours.

JEL Classification numbers: C12, C13.

present an unbiased correction procedure and apply it to the models reported by Feenstra and Hanson (1999) and Haskel and Slaughter (2002).

I. Introduction

In recent economic studies two-stage estimation procedures are increasingly being used when the dependent variable of interest cannot be measured but can be estimated in a first-stage regression. For instance, Gaston and Trefler (1995) estimate the impact of trade flows and trade policy on sector wages in two stages. In a first stage, they regress the wages of individual workers on a number of characteristics of the individuals and industry dummies. The coefficients of the industry dummies are considered as an estimate of the inter-industry wage differentials which are, in a second stage, regressed on a vector of explanatory variables reflecting international trade and protection. Bird (2001), in his verification of Sinn's (1995) welfare state theory, estimates 'risk-taking' (of investment in physical and human capital) as the standard deviation of the residuals of the regression of annual personal income on the determinants of permanent income. In a second-stage regression, this estimate of 'risk-taking' is, amongst other variables, related to country-level measures of redistribution.

Haskel and Slaughter (2002) use a two-stage regression to determine whether skill-biased technological change (SBTC) is subject to a sector bias. As SBTC cannot be observed directly, they estimate it as the residual of a first-stage regression of the change in the wage bill share of non-production workers on the change in the relative wages of non-production workers and the change in the capital–output ratio. In a second stage, this estimate of SBTC is regressed on the start-of-decade sector skill intensity.

Perhaps the best-known recent application of a two-stage estimation procedure is the two-stage mandated wage regression (Feenstra and Hanson, 1997, 1999; Haskel and Slaugher, 2001, 2003), used to verify the relevance of the Stolper–Samuelson theorem in the academic debate on the determinants of rising wage inequality. The mandated wage approach was proposed by Leamer (1996), following the early lead of Baldwin and Hilton (1984). From the zero-profit condition (assuming perfect competition), a relationship between changes in domestic product prices, changes in factor rewards and changes in total factor productivity (TFP) can be derived. Regressing the change in domestic product prices on the cost shares of the production factors, the estimated coefficients are considered to be estimates of changes in factor prices that are 'mandated' by changes in the product prices. However, if domestic product prices can change independent of international competition, the effect of international import price competition on domestic prices cannot be observed directly and has to be estimated. This can be carried out by regressing, in a first stage, domestic product price changes on structural (exogenous) determinants (e.g. international trade and technology variables). In the second stage, the estimated contribution of each structural variable to changes in domestic product prices is then regressed on the factor shares to decompose the changes in factor prices into portions explained by each respective structural determinant (Feenstra and Hanson, 1999, p. 915).

Feenstra and Hanson (1997) (henceforth FH97) rightly argue that, as for each structural determinant the estimated coefficient from the first-stage regression is used to generate the dependent variables for the second-stage regressions, the standard errors of the second-stage coefficient estimates need to be corrected to account for the additional variance of the first-stage estimation. The second-stage regressions are indeed conditional estimates of the residuals that do not account for the additional variance that is induced by using estimated rather than effectively measured variables. To test the significance of the second-stage coefficients unconditional estimates of the standard errors, accounting for this additional variance, have to be computed. However, the correction procedure that Feenstra and Hanson (1997) propose does not warrant that the computed variances will be positive. Therefore, in some cases, corrected standard errors cannot be computed and no inference on the statistical significance of the estimates can be made (see Feenstra and Hanson, 1997, p. 48, Table 7; p. 49, Table 8 and Feenstra and Hanson, 1999, p. 934, Table VI). Haskel and Slaughter (2001), applying the FH97 procedure to a two-stage mandated wage estimation for the UK, also run across 'negative variances' (Haskel and Slaughter, 2001, p. 180, Table 7; p. 182, Table 9; p. 185, Table A1). Haskel and Slaughter (2003), performing a two-stage mandated wage regression for the US with a large number of structural determinants (e.g. tariffs, transport costs, exchange rates), also encounter 'negative variances' when correcting the standard errors of the second stage estimates following FH97 (Haskel and Slaughter, 2003, p. 645, Table 6).

The occurrence of 'negative variances' indicates a negative bias in the proposed correction procedure. This implies that the significance tests based on the corrected variances following FH97 underestimate size and overestimate power, i.e. the significance of the estimates will be overestimated.

In section II, we propose a consistent correction procedure which does guarantee positive variances of the second-stage estimated coefficients. In section III, we report the results of a re-estimation of the mandated wage regression by Feenstra and Hanson (1999). We compare the standard errors computed with our correction procedure with the results reported by Feenstra and Hanson (1999). We also report the results of a re-estimation of the sector bias of skill-biased technological change by Haskel and Slaughter (2002), again comparing the standard errors corrected with the FH97 procedure, as applied by the authors, to the standard errors corrected with the procedure that we propose.

II. The correction procedure

In this section, we present a general formulation of the correction of the standard errors in the second stage of two-stage regressions in which the dependent variables in the second stage are generated from estimates of the first-stage regression.

Consider a first-stage estimation in which a vector of variables $\mathbf{Y}$ is regressed on a vector of exogenous variables $\mathbf{X}$:

$$\mathbf{Y} = \mathbf{X}\beta + \delta \tag{1}$$

and a second-stage regression in which the contribution of the kth element of the vector $\mathbf{X}$ is regressed on a vector of exogenous variables $\mathbf{Z}$:

$$\mathbf{X}_k\beta_k = \mathbf{Z}\gamma_k + \varepsilon_k. \tag{2}$$

As $\mathbf{X}_k\beta_k$ cannot be observed directly it has to be replaced by its estimate from the first stage (i.e. $\mathbf{X}_k\hat{\beta}_k$). For the kth element of the vector of first-stage coefficients β we know:

$$\hat{\beta}_k = \beta_k + S_k\eta_k, \tag{3}$$

with $\Omega_k = S_k S_k'$ the covariance matrix of $\hat{\beta}_k$ and η_k i.i.d. $\sim N(0, 1)$.

Multiply equation (3) by $\mathbf{X}_k$ and insert this into equation (2) resulting in:

$$\mathbf{X}_k\hat{\beta}_k = \mathbf{Z}\gamma_k + \varepsilon_k + \mathbf{X}_k S_k \eta_k, \tag{4}$$

which is the empirical equation that Feenstra and Hanson (1997, 1999) present as their second-stage regression equation. The third term on the right-hand side represents, according to the authors, the additional variance introduced in the second-stage regression as a result of the variance of the estimator $\hat{\beta}_k$. The authors conclude that, to estimate the variance of ε_k in the second stage, the variance of the observed residuals has to be corrected by subtracting (an estimate of) this source of extra variance:

$$\hat{\sigma}_\varepsilon^2 = \mathrm{trace}\left(\frac{\mathbf{u}_k\mathbf{u}_k' - \mathbf{X}_k\Omega_k\mathbf{X}_k'}{n}\right), \tag{5}$$

where $\mathbf{u}_k$ is the vector of estimated residuals from equation (4).

Unfortunately, this procedure sometimes results in negative computed variances, namely when $\mathbf{X}_k\Omega_k\mathbf{X}_k'$ exceeds $\mathbf{u}_k\mathbf{u}_k'$.

When equation (4) is actually estimated, the estimator is replaced by the point estimate from the first-stage regression. The difference between this realization and the true parameter β_k corresponds to an unobservable particular realization of η_k. Hence, the second-stage regression is a conditional regression, resulting in conditional estimates of γ_k and ε_k. This implies that, to correct the former for the additional variance induced by using estimated rather than observed values of β_k, one has to derive the unconditional distribution of $\hat{\gamma}_k$, which is the relevant distribution for significance tests. As we will show, this 'unconditioning' of the estimators does not require the subtraction of $\mathbf{X}_k\Omega_k\mathbf{X}_k'$ from $\mathbf{u}_k\mathbf{u}_k'$.

The conditional equation to be estimated in the second stage, takes the form:[1]

$$\mathbf{X}\hat{\beta}\big|_{\hat{\beta}=\tilde{\beta}} = \{\mathbf{Z}\gamma + \varepsilon\}\big|_{\hat{\beta}=\tilde{\beta}}, \tag{6}$$

where $\tilde{\beta}$ represents the estimated value of β from the first stage. The cumbersome notation involved is used to emphasize the important difference between the OLS estimator $\hat{\beta}$ and the actual estimate $\tilde{\beta}$. To simplify notation, however, we will use $\tilde{\beta}$ as shorthand for $\hat{\beta} = \tilde{\beta}$.

The conditional ordinary least squares (OLS) estimator of γ is given by:

$$\hat{\gamma}\big|_{\tilde{\beta}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}\hat{\beta}\big|_{\tilde{\beta}}. \tag{7}$$

Our objective is to obtain an expression for the unconditional variance of $\hat{\gamma}$, which is defined as the sum of the variance of the conditional expectation of $\hat{\gamma}$ and the expected value of its conditional variance. The conditional expected value is simply:

$$E\left[\hat{\gamma}\big|_{\tilde{\beta}}\right] = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'E(\mathbf{X})\hat{\beta}\big|_{\tilde{\beta}} \tag{8}$$

and its variance (over all possible realizations of $\tilde{\beta}$) is therefore given by:

$$\mathrm{var}\left\{E\left[\hat{\gamma}\big|_{\tilde{\beta}}\right]\right\} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'E(\mathbf{X})\Omega E(\mathbf{X}')\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}. \tag{9}$$

The conditional variance of $\hat{\gamma}$ can be written as:

$$\mathrm{var}\left[\hat{\gamma}\big|_{\tilde{\beta}}\right] = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathrm{var}\left[\mathbf{X}\hat{\beta}\big|_{\tilde{\beta}}\right]\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}. \tag{10}$$

Given that $\mathbf{X}\hat{\beta}\big|_{\tilde{\beta}}$ is a vector of n independent realizations of linear combinations of the components of $\mathbf{X}$, its conditional variance can be written as:

$$\mathrm{var}\left[\mathbf{X}\hat{\beta}\big|_{\tilde{\beta}}\right] = \hat{\beta}'\big|_{\tilde{\beta}}\Sigma_{\mathbf{X}}\hat{\beta}\big|_{\tilde{\beta}}, \tag{11}$$

where $\mathbf{\Sigma}_X$ is the covariance matrix of $\mathbf{X}$. As equation (6) implies that the mean of the components of $\mathbf{\Sigma}_X$ are linear combinations of the columns of $\mathbf{Z}$, $\mathbf{\Sigma}_X$ is the covariance matrix of $\mathbf{X}$ after regression on $\mathbf{Z}$. The expected value of this conditional variance (over all possible realizations of $\tilde{\beta}$) is then:

$$E\left\{\mathrm{var}\left[\hat{\gamma}\big|_{\tilde{\beta}}\right]\right\} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'E\left[\hat{\beta}'\big|_{\tilde{\beta}}\Sigma_{\mathbf{X}}\hat{\beta}\big|_{\tilde{\beta}}\right]\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}. \tag{12}$$

[1]Henceforth we drop the subscript k to simplify the notation.

Combining equations (12) and (9) yields the unconditional variance of $\hat{\gamma}$:

$$\mathrm{var}\left[\hat{\gamma}\right] = (\mathbf{Z}'\mathbf{Z})^{-1}\left\{\mathbf{Z}'E\left[\hat{\beta}'\big|_{\tilde{\beta}}\boldsymbol{\Sigma}_{\mathbf{X}}\hat{\beta}\big|_{\tilde{\beta}}\right]\mathbf{Z} + \mathbf{Z}'E(\mathbf{X})\Omega E(\mathbf{X}')\mathbf{Z}\right\}(\mathbf{Z}'\mathbf{Z})^{-1}. \quad (13)$$

This result can also be obtained within a method of moments framework, as shown in the Appendix.[2]

It remains to be shown how the two components of this unconditional variance can be estimated. The central part of the second component, $E(\mathbf{X})\ \Omega\ E(\mathbf{X}')$, can be estimated by $\mathbf{X}\hat{\Omega}\mathbf{X}'$ given that $\hat{\Omega}$, the covariance matrix of $\hat{\beta}$, is an unbiased estimator of Ω obtained from the first stage.

The scalar factor in the first component of equation (13), $E\left[\hat{\beta}'\big|_{\tilde{\beta}}\boldsymbol{\Sigma}_{\mathbf{X}}\hat{\beta}\big|_{\tilde{\beta}}\right]$, is estimated from the estimated residuals of the second stage:

$$\hat{\sigma}_{\varepsilon}^2 = \frac{u'u}{n-v} = \frac{1}{n-v}\left(\mathbf{X}\hat{\beta}\big|_{\tilde{\beta}} - \mathbf{Z}\hat{\gamma}\big|_{\tilde{\beta}}\right)'\left(\mathbf{X}\hat{\beta}\big|_{\tilde{\beta}} - \mathbf{Z}\hat{\gamma}\big|_{\tilde{\beta}}\right) = \hat{\beta}'\big|_{\tilde{\beta}}\hat{\boldsymbol{\Sigma}}_{\mathbf{X}}\hat{\beta}\big|_{\tilde{\beta}}, \quad (14)$$

where v is the number of columns of $\mathbf{Z}$. It follows that

$$E\left[\hat{\sigma}_{\varepsilon}^2\right] = E\left\{E\left[\hat{\sigma}_{\varepsilon}^2\big|_{\tilde{\beta}}\right]\right\} = E\left[\hat{\beta}'\big|_{\tilde{\beta}}\boldsymbol{\Sigma}_{\mathbf{X}}\hat{\beta}\big|_{\tilde{\beta}}\right]. \quad (15)$$

Contrary to the suggestion of equation (4) (i.e. the empirical specification considered by Feenstra and Hanson, 1997), $\hat{\sigma}_{\varepsilon}^2$ is an unbiased estimator of the scalar factor in the first component of equation (13). Hence, the correction of FH97 induces a downward bias in the variance of $\hat{\gamma}$ in finite samples.

The estimate of the unconditional variance of $\hat{\gamma}$ in equation (13) cannot assume negative values, as opposed to FH97. It should be stressed that equation (13) is essentially the same as the equation that Feenstra and Hanson (1997, p. 37) propose to correct the standard errors, but with a different value for the estimated variance of the second-stage residuals, and with $E(\mathbf{X})$ replaced by the observed $\mathbf{X}$.

III. Empirical results

In section II, we have shown that the FH97 procedure to correct the standard errors of the second-stage regression, appears to induce a negative bias. This results in the overestimation of the significance of the reported second stage coefficients. To illustrate the problem, we have replicated two studies in which the FH97 procedure has been applied and for which data were available.

First we used the data, kindly provided by Robert Feenstra, to re-estimate the two-stage mandated wage regression proposed in Feenstra and Hanson (1999). In a first stage they regress the sum of changes in

[2]We would like to thank Adrian Pagan for pointing this out.

domestic value added prices and TFP on a vector of exogenous structural determinants $\mathbf{Z}$:[3]

$$\Delta \ln p^{\mathrm{VA}} + \Delta\mathrm{TFP} = \mathbf{Z}\beta + \delta. \tag{16}$$

The contribution of the kth structural determinant to changes in domestic product prices ($\Delta \ln p^{\mathrm{VA}}$) and productivity ($\Delta$TFP) is estimated from the first-stage regression as $\mathbf{Z}_k\hat{\beta}_k$ which is, in the second stage, regressed on the vector of factor shares $\mathbf{V}$:

$$\mathbf{Z}_k\hat{\beta}_k = \mathbf{V}\Delta\omega_k + \varepsilon_k. \tag{17}$$

The estimated coefficients of the second stage ($\Delta\hat{\omega}_k$) are then, following Leamer (1996), considered as estimates of the changes in the rewards of the respective production factors that can be explained by changes in domestic product prices and productivity 'mandated' by the the kth structural variable.

In Table 1, we report the replicated second-stage coefficients of Feenstra and Hanson (1999, p. 933, Table V), as estimated with equation (17). Three production factors are considered: production workers (proxy for low-skilled workers), non-production workers (proxy for high-skilled workers) and capital. Four structural determinants are considered: foreign outsourcing in the narrow sense (i.e. imported intermediate inputs from the same two-digit industry), the difference between foreign outsourcing in the broad sense (i.e. all non-energy imported intermediate inputs) and foreign outsourcing in the narrow sense, the share of computing machinery in total capital and the difference between the share of high-tech capital and computing machinery in total capital. In round brackets we report the standard errors corrected following FH97. In square brackets we report the standard errors as computed with equation (13), using equation (15), given in section II.

Comparing the two procedures reveals that FH97 overestimates the significance of the second stage coefficient estimates, although overall the differences between both procedures seem relatively small. This is explained by the fact that the second-stage component of the total (unconditional) variance of $\Delta\hat{\omega}$ is outweighed by its first-stage component, i.e. the dominance of the second term over the first term in the formula of Feenstra and Hanson (1997, p. 37). However, our computed standard error does differ substantially from the one of Feenstra and Hanson (1999) for the coefficient of the non-production workers in the regression with the high-tech share as dependent variable.

[3]Feenstra and Hanson argue that a specification in which domestic product price changes are regressed on $\mathbf{Z}$ and ΔTFP, with the coefficient of the latter providing an estimate of the pass-through of productivity changes into domestic product prices, would suffer from an endogeneity bias given the link, by definition, between product price changes and TFP. They therefore prefer a specification in which the sum of changes in domestic product prices and TFP is considered as the dependent variable in the first-stage regression.

TABLE 1

Estimated factor price changes (1979–1990) (using ex post rental prices for computer share and high-tech share)

	Dependent variable			
Independent variable	*Outsourcing (narrow)*	*Outsourcing (difference)*	*Computer share*	*High-tech share (difference)*
Change in value-added prices plus TFP explained by				
Production labour share	−0.010	0.020	−0.005	0.026
FH97 corrected standard errors	(0.009)	(0.014)	(0.012)	(0.025)
Our corrected standard errors	[0.010]	[0.015]	[0.014]	[0.025]
Non-production labour share	0.099	0.063	0.248	0.007
FH97 corrected standard errors	(0.049)	(0.039)	(0.100)	(0.004)
Our corrected standard errors	[0.050]	[0.041]	[0.100]	[0.011]
Capital share	0.002	−0.001	0.001	0.004
FH97 corrected standard errors	(0.003)	(0.003)	(0.004)	(0.004)
Our corrected standard errors	[0.004]	[0.004]	[0.005]	[0.005]

Notes: This table reports the results from a re-estimation of the second stage regression, i.e. equation (17), by Feenstra and Hanson (1999, Table V, p. 933). The dependent variables are constructed by using the coefficient estimates from the first-stage regression, i.e. equation (16) (our coefficient estimates are the same as in Feenstra and Hanson, 1999).

Source: Data from the website of Robert C. Feenstra at University of California, Davis.

In Table 2, we report the replications of the two-stage estimation of the sector bias of SBTC by Haskel and Slaughter (2002) for the United States in the 1970s and the 1980s. In a first stage, Haskel and Slaughter regress the level change in the total wage bill share of the non-production workers (proxy for high-skilled workers) in sector k on changes in the relative wages of non-production workers (i.e. skill premium) and changes in the capital–output ratio:

$$\Delta\omega_k = a_0 + a_1 \Delta \log\left(\frac{w_s}{w_u}\right)_k + a_2 \Delta \log\left(\frac{K}{Y}\right)_k + \varepsilon_k. \tag{18}$$

The variation in $\Delta\omega_k$ that cannot be explained by the skill premium or the capital–output ratio is considered as an estimate of SBTC. In the second stage, this estimate is regressed on the start-of-decade skill intensity (ratio of high-skilled to low-skilled workers):

$$\widehat{\mathrm{SBTC}}_k = \alpha + \beta\left(\frac{S}{U}\right)_k + u_k. \tag{19}$$

Haskel and Slaughter consider $\hat{\beta}$ as an estimate of the sector bias of SBTC, i.e. a positive coefficient would indicate that SBTC is concentrated in high-skill-intensive industries. The authors consider four alternative estimates of SBTC as the dependent variable in their second stage regression equation (19). Using data from the NBER Productivity database,

TABLE 2

Estimated sector bias of skill-biased technical change in the US (1970–1990)

	Dependent variable		
Independent variable	*SBTC1*	*SBTC2*	*SBTC3*
1970s			
Sector bias of SBTC	−0.03	−0.02	−0.01
FH97 corrected standard errors	(0.005)	(0.006)	(0.006)
Our corrected standard errors	[0.008]	[0.008]	[0.008]
1980s			
Sector bias of SBTC	0.05	0.04	0.02
FH97 corrected standard errors	(0.006)	(0.006)	(0.009)
Our corrected standard errors	[0.007]	[0.008]	[0.011]

Notes: This table reports the results from a re-estimation of the second-stage regression, i.e. equation (19) by Haskel and Slaughter (2002, first three columns Table 2, p. 1774). The dependent variables are three alternative measures for skill-biased technical change and are generated from the first-stage regression of equation (18). Our coefficient estimates are the same as in Haskel and Slaughter (2002), except for the coefficient of the sector bias in the SBTC1 specification for the 1980s, for which we find 0.05 whereas Haskel and Slaughter (2002) report an estimate of 0.06.

Source: NBER Productivity Database.

we could replicate the estimations for three alternative measures: SBTC1 (estimated intercept plus residual of the first-stage regression: $a_0 + \varepsilon_k$), SBTC2 (residual of the first-stage regression from which the skill premium has been omitted) and SBTC3 (residual of the first-stage regression from which the skill premium has been omitted but in which sector dummies have been inserted). As the dependent variable of the second-stage regression is generated from estimates of the first-stage regression, Haskel and Slaughter applied the FH97 procedure to correct the standard errors of the second-stage estimates.

Our estimates for the coefficients are the same as those by Haskel and Slaughter (2002) except for the SBTC1 regression for the 1980s for which we find a coefficient of 0.05 compared with 0.06 reported by Haskel and Slaughter (2002).

We compare the standard errors reported in Haskel and Slaughter (2002) (in round brackets) to the standard errors computed with our procedure (reported in square brackets). Our corrected standard errors are all substantially higher than the reported standard errors using the FH97 procedure. Qualitatively however, the results of Haskel and Slaughter are not affected, although they are somewhat less robust: the sector bias of skill-biased technical change turns out to be significant for the first and second specification (SBTC1 and SBTC2) but not for the third (SBTC3), at least not at the 5% significance level. The fact that the estimated coefficients remain significant is due to the combination of two effects: the low value of the uncorrected standard errors and the sample size (which reduces the impact of the bias).[4]

IV. Conclusions

In this paper, we proposed a procedure to correct the standard errors of the second-stage estimates of a two-stage regression in which the dependent variable in the second stage is generated from the estimates of the first-stage estimation. These two-stage regressions have apparently become increasingly popular in recent empirical studies. The second-stage regression results in conditional estimates of the residuals that do not account for the additional variance that is induced by using estimated rather than true dependent variables. To test the significance of the second-stage coefficients, unconditional estimates of the standard errors accounting for this additional variance, should be computed.

Feenstra and Hanson (1997) proposed a procedure to correct the standard errors in the second stage of a two-stage mandated wage regression, used to estimate the impact of international trade competition and SBTC on wage inequality. The proposed correction procedure has since been used in other two-stage estimations (e.g. Haskel and Slaughter, 2001, 2002, 2003). However the proposed procedure does not warrant positive variances. Apart from the fact that this implies that in some cases 'corrected' standard errors cannot be computed and hence no inference on the statistical significance of the second stage estimates can be made, this indicates a negative bias in the proposed correction.

We have shown that this negative bias is due to the fact that Feenstra and Hanson (1997) subtract the variance of the first-stage residuals from the variance of the conditional second-stage residuals. We proposed an expression of the unconditional variance of the second-stage estimates that is not subject to this bias. The correction that we propose can alternatively be derived with the method of moments estimator of the asymptotic covariance matrix as proposed by Hansen (1982) and applied to two-stage regressions by Newey (1984) and Pagan (1986).

We replicated the estimation of the two-stage mandated wage regression by Feenstra and Hanson (1999) and the two-stage estimation of the sector bias of SBTC by Haskel and Slaughter (2002) to compare the results of the correction proposed by Feenstra and Hanson (1997) to the correction we propose. The difference between the standard errors corrected with the two distinct procedures is small for the estimation by Feenstra and Hanson (1999) but is relatively high for the estimation results of Haskel and Slaughter (2002),

[4]This can be verified by observing that $\mathbf{X}_k\hat{\Omega}_k\mathbf{X}_k'$ is divided by n in equation (5).

although the main conclusions still hold given the relatively high significance of the estimates. One coefficient estimate of Haskel and Slaughter (2002), however, is significant at the usual 5% error level when correcting the standard errors following Feenstra and Hanson (1997) but only at the 10% level when following our correction procedure.

The negative bias of the Feenstra and Hanson (1997) correction is higher the smaller the sample size is. The size of the samples used by Feenstra and Hanson (1999) and Haskel and Slaughter (2002) is rather high given their use of the NBER Productivity Database covering some 450 industries. The difference between the two correction procedures and thus the overestimation of the significance of the estimated second stage effects in two-stage regressions with generated dependent variables is likely to be more severe in samples of more frequently occurring (i.e. smaller) size.

Final Manuscript Received: October 2004

References

Baldwin, R. E. and Hilton, R. S. (1984). 'A technique for indicating comparative costs and predicting changes in trade ratios', *The Review of Economics and Statistics*, Vol. 64, pp. 105–110.

Bird, E. J. (2001). 'Does the welfare state induce risk-taking?', *Journal of Public Economics*, Vol. 80, pp. 357–383.

Feenstra, R. C. and Hanson, G. H. (1997). *Productivity Measurement and the Impact of Trade and Technology on Wages: Estimates for the US, 1972–1990*. NBER Working Paper No. 6052.

Feenstra, R. C. and Hanson, G. H. (1999). 'The impact of outsourcing and high-technology capital on wages: estimates for the United States, 1979–1990', *Quarterly Journal of Economics*, Vol. 114, pp. 907–940.

Gaston, N. and Trefler, D. (1995). 'Union wage sensitivity to trade and protection: theory and evidence', *Journal of International Economics*, Vol. 39, pp. 1–25.

Hansen, L. P. (1982). 'Large sample properties of generalized method of moments estimators', *Econometrica*, Vol. 50, pp. 1029–1054.

Haskel, J. and Slaughter, M. J. (2001). 'Trade, technology and UK wage inequality', *The Economic Journal*, Vol. 111, pp. 163–187.

Haskel, J. and Slaughter, M. J. (2002). 'Does the sector bias of skill-biased technical change explain changing skill premia?', *European Economic Review*, Vol. 46, pp. 1757–1783.

Haskel, J. and Slaughter, M. J. (2003). 'Have falling tariffs and transportation costs raised US wage inequality?', *Review of International Economics*, Vol. 11, pp. 630–650.

Leamer, E. E. (1996). *In Search of Stolper–Samuelson Effects on US Wages*, NBER Working Paper No. 5427 (published in 1998 in S. M. Collins (ed.), *Imports, Export, and the American Worker*, Brookings Institution Press, Washington, DC).

Newey, W. K. (1984). 'A method of moments interpretation of sequential estimators', *Economics Letters*, Vol. 14, pp. 201–206.

Pagan, A. (1986). 'Two-stage and related estimators and their applications', *Review of Economic Studies*, Vol. 53, pp. 517–538.

Sinn, H.-W. (1995). ‘A theory of the welfare state’, *Scandinavian Journal of Economics*, Vol. 83, pp. 263–283.

Appendix

In this appendix, we show how the ‘unconditional’ variance of the second-stage coefficient estimates can alternatively be derived using a method of moments approach.

The properties of the method of moments estimator of the asymptotic covariance matrix have been studied by Hansen (1982). Newey (1984) and Pagan (1986) have applied Hansen’s results in two-stage regression models.

The two equations (stages) to be estimated are:

$$\mathbf{Y} = \mathbf{X}\beta + \delta \tag{A1}$$

$$\mathbf{X}\hat{\beta} = \mathbf{Z}\gamma + \varepsilon \tag{A2}$$

that is, the estimated β from the first stage ($\hat{\beta}$) is used to generate the dependent variables of the second-stage regression.[5]

The moment conditions:

$$E[g(\mathbf{X}, \beta] = 0 \tag{A3}$$

$$E[h(\mathbf{Z}, \gamma, \hat{\beta})] = 0 \tag{A4}$$

for this problem are:

$$E[\mathbf{X}'(\mathbf{Y} - \mathbf{X}\beta)] = 0 \tag{A5}$$

$$E[\mathbf{Z}'(\mathbf{X}\hat{\beta} - \mathbf{Z}\gamma)] = 0. \tag{A6}$$

Define

$$G_\beta = E\left[\frac{\partial g(\mathbf{X}, \beta)}{\partial \beta}\right] \tag{A7}$$

$$H_\beta = E\left[\frac{\partial h(\mathbf{Z}, \gamma, \hat{\beta})}{\partial \beta}\right] \tag{A8}$$

[5]It should be noted that the generated regressand problem is not strictly identical to the generated regressor problems considered by Newey and Pagan since the generated variables from the first-step estimates enter the second-step equation as the dependent rather than as the independent variables. This distinction does not affect the results obtained.

$$H_\gamma = E\left[\frac{\partial h(\mathbf{Z}, \gamma, \hat{\beta})}{\partial \gamma}\right] \tag{A9}$$

$$V_{gg} = E[g(\mathbf{X}, \beta)g(\mathbf{X}, \beta)'] \tag{A10}$$

$$V_{gh} = V_{hg} = E[g(\mathbf{X}, \beta)h(\mathbf{Z}, \gamma, \hat{\beta})'] \tag{A11}$$

$$V_{hh} = E[h(\mathbf{Z}, \gamma, \hat{\beta})h(\mathbf{Z}, \gamma, \hat{\beta})']. \tag{A12}$$

A consistent estimator of the asymptotic covariance matrix of γ is $F^{-1}VF^{-1\prime}$ (Hansen, 1982), where:

$$F = \begin{bmatrix} G_\beta & 0 \\ H_\beta & H\gamma \end{bmatrix}$$

and

$$V = \begin{bmatrix} V_{gg} & V_{gh} \\ V_{hg} & V_{hh} \end{bmatrix}.$$

The square matrices F and V have been partitioned as in Newey (1984). Applying these results gives:

$$F = \begin{bmatrix} -X'X & 0 \\ Z'X & -Z'Z \end{bmatrix}$$

$$V = \begin{bmatrix} \sigma_\delta^2 X'X & 0 \\ 0 & \sigma_\varepsilon^2 Z'Z \end{bmatrix}$$

with $V_{gh} = V_{hg} = 0$ under the assumption of independence between δ and ε.

The corrected covariance matrix for the second step coefficients is given by (Newey, 1984):

$$\Omega_\gamma = H_\gamma^{-1} V_{hh} H_\gamma^{-1\prime} + H_\gamma^{-1} H_\beta \left[G_\beta^{-1} V_{gg} G_\beta^{-1\prime}\right] H_\beta' H_\gamma^{-1\prime}. \tag{A13}$$

Inserting the appropriate sub-matrices of F and V yields:

$$\Omega_\gamma = (\mathbf{Z}'\mathbf{Z})^{-1}\left[\mathbf{Z}'\hat{\sigma}_\varepsilon^2\mathbf{Z} + \mathbf{Z}'\mathbf{X}\hat{\Omega}_\beta\mathbf{X}'\mathbf{Z}\right](\mathbf{Z}'\mathbf{Z})^{-1} \tag{A14}$$

which is exactly equation (13).