KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and Applied Economics

Orthogonalized regressors and spurious precision

Piet Sercu and Martina Vandebroek

DEPARTMENT OF ACCOUNTANCY, FINANCE AND INSURANCE (AFI)

AFI 0618

Orthogonalized regressors and spurious precision*

Piet Sercu† and Martina Vandebroek‡

First draft: September 2004; this draft: September 2006
Submission to the applied section of the Journal of Banking and Finance

*The authors thanks Christophe Croux and Geert Dhaene for useful suggestions, but accept full responsability for any remaining errors.

†KU Leuven, Leuven School of Business and Economics, Naamsestraat 69, B- 3000 Leuven; Tel: +32 16 326 756; piet.sercu@econ.kuleuven.be.

‡KU Leuven, Leuven School of Business and Economics, Naamsestraat 69, B- 3000 Leuven; Tel: +32 16 326 975; martina.vandebroek@econ.kuleuven.be.

Abstract

The exposure of a stock's return to exchange-rate changes is conventionally estimated by regression. Often, the market return is included as an additional regressor. By first orthogonalizing the market return on the exchange rate one seems to have the best of both worlds: the market factor cannot subsume part of the exposure present in a stock's return, and the SE of the estimate beats both the simple- and the multiple-regression SE's. This last effect is illusory: since the simple and the pseudo-multiple regression always produce the same exposure estimate, given the sample, their precision must be identical too. Technically, the source of the problem is that the uncertainty about the market's exposure estimate is left out of the calculated SE. In published work, the calculated error variances should be corrected upward by 20 to 100 percent.

Keywords: Market Model, currency, exposure.
JEL-codes: .

Orthogonalized regressors and spurious precision

Introduction

A stock's currency exposure is often measured by the slope coeficients of a regression of the stock's return on the percentage changes in the exchange rates. Such a vector of currency exposures, suitably rescaled, tells us what positions should be taken in each currency's forward market to hedge the investment in the stock, at least if the hedger's objective is to minimize variance and no other hedges are used, see Stein (1960) and Johnson (1960). In a more academic application, hedged-stock returns also show up in some versions of the International CAPM, see Sercu (1980) or Adler and Dumas (1983). Thus, demonstrating that exposure exists is a natural first step in testing the relevance of international asset pricing. Since Jorion (1990) it is common practice to add the market return as an additional regressor. Why and how (not) to do so is the issue in this note.

One possible user of this regression information may be the hedger: the multiple regression coefficients provide estimates of the Stein-Johnson hedges ratios if the hedge instruments include not just currency forwards but also an index-futures contract. In academe, however, the objective is not so much to obtain a hedge ratio, but to establish whether exchange-rate exposure exists and is related to the firm's business. Since the market model—the regression of the stock's return on the market return—is *the* workhorse, if anything, in empirical finance, there is a general feeling that any reasonable return-generating process must include at least the market. An additional motivation may have been that the additional regressor can improve the power of the tests: the residual variance shrinks, which, everything else being the same, reduces the variance of the estimator of exposure. On the other hand, there may be some correlation between the market and exchange factors, which then undoes part or all of the gain from the reduced residual variance: the more similar the regressors, the harder it is to sort out their separate contributions. A related problem is that if all stocks have similar exposures to currency factors (possibly up to a firm-specific factor common to all currencies), the market return will exhibit this common exposure pattern too. Since a multiple regression coefficient for the exchange rate measures exposure over and above that already present in the market portfolio, strong similarities in the exchange-rate sensitivities of stocks would kill the chances of finding convincing stock-specific currency effects. To obtain a lower residual noise without giving the market return the chance to subsume the individual currency effects, then, one can first orthogonalise the market return on the exchange rate(s), as in a recent JBF article by Pritamani, Shome and Singa (2004).[1] Our message is that this practice is flawed, in the sense that the drop of the estimator's standard error is illusory and the significance tests unreliable. Sometimes the reverse procedure is adopted, first orthogonalizing the exchange rate on the market.[2] This procedure produces the same exposure coefficient and standard error as the genuine multiple regression; nothing is gained, but nothing is wrong either as long one is not interested in the significance of the market sensitivity.

Section 1 presents the analytical arguments for the case where the market return is the variable being orthogonalized, and Section 2 for the obverse case. In Section 3 we report some Monte-Carlo illustrations; we conclude in Section 4.

1 Orthogonalizing the control variable on the variable of interest

Consider an economy where returns on assets and percentage changes in exchange rates are joint normal processes. This implies that linear "regression" relationships exist between any individual variable on the one hand, and any subset of the other variables or pre-set linear combinations. There is no single true generating process for any variable; the issue just is what variables are observable for analysis or prediction. For simplicity of exposition we consider just one foreign currency, whose percentage changes over period t are denoted by s_t; generalisation to multiple currencies is simple.[3] In terms of relative exposures,[4] the original Dumas (1978) regression is

$$R_{j,t} = \alpha_{1,j} + \gamma_{1,j} s_t + u_{1,j,t}, \tag{1}$$

[1] Other recent papers are *e.g.* Bodnar and Wong, 2003; Bris and Koskinen, 2002; Entorf, Moebert and Sonderhof, 2006; and Priestley and Odegaard, 2002

[2] See e.g. Glaum, Brunner and Himmel, 2000; Hagelin and Pamborg, 2002; and Jorion, 1991.

[3] Following Jorion, one also often collapses the various exchange-rate changes into a single variable, interpreted as the percentage change in the value of a currency basket—typically a trade-weighted one. The objective is to avoid multicollinearity. As Rees and Unni (2005) point out, this assumes that stocks' exposures to the N exchange rates are all proportional to the set of weights used in the basket, an assumption that is *a priori* tenuous and empirically rejected in their tests. We will, however, assume a single exchange-rate regressor for the sake of simplicity of exposition.

[4] The original Dumas (1978) regression was written in terms of values, so that the slope has the dimension of an amount of forex units. For empirical purposes or in asset pricing theory one works with percentage changes rather than values. So the regression provides a dimensionless relative exposure, an elasticity rather than a partial derivative.

where $R_{j,t}$ denotes the stock's return over period t. The merged version of (1) and the market model is

$$R_{j,t} = \alpha_{2,j} + \gamma_{2,j} s_t + \beta_{2,j} R_{m,t} + u_{2,j,t}. \tag{2}$$

The equation used in some studies, lastly, is a hybrid version,

$$R_{j,t} = \alpha_{3,j} + \gamma_{3,j} s_t + \beta_{3,j} u_{1,m,t} + u_{3,j,t}, \tag{3}$$

where $u_{1,m,t}$ is the error from the market portfolio's version of (1),

$$R_{m,t} = \alpha_{1,m} + \gamma_{1,m} s_t + u_{1,m,t}. \tag{4}$$

In the above, the notation refers to the true parameters and errors rather than their estimates. Actually, the statistical problems discussed in this paper stem from the use of imperfect estimates. However, to see the issues we need to understand the relations between the true regressions first; and the results reviewed below for population moments also hold for sample moments and, therefore, for method-of-moments estimators like OLS. To identify the links between the three regressions, substitute Equation (4) into (2):

$$\begin{aligned} R_{j,t} &= \alpha_{2,j} + \gamma_{2,j} s_t + \beta_{2,j}[\alpha_{1,m} + \gamma_{1,m} s_t + u_{1,m,t}] + u_{2,j,t}, \\ &= \underbrace{[\alpha_{2,j} + \beta_{2,j}\alpha_{1,m}]}_{=\alpha_{3,j}} + \underbrace{[\gamma_{2,j} + \beta_{2,j}\gamma_{1,m}]}_{=\gamma_{3,j}=\gamma_{1,j}} s_t + \underbrace{[\beta_{2,j}]}_{=\beta_{3,j}} u_{1,m,t} + \underbrace{u_{2,j,t}}_{u_{3,j,t}}. \end{aligned} \tag{5}$$

The underbrace comments deserve some comments. First, they claim that the square-bracketed expressions must be the coefficients of the hybrid equation. To prove this, first note that the array of regression coefficients is the unique vector that make the residuals orthogonal on the regressors. Next note that $u_{2,j}$, being orthogonal on s and R_m, is also orthogonal on linear combinations of those two, like s and $u_{1,m}$. Therefore Equation (5) indeed is a *bona fide* regression of R_j on s and $u_{1,m}$. The second comment is a corollary of the first: $\beta_{3,j}$ equals $\beta_{2,j}$—a result closely related to the Frisch-Waugh (1933) Theorem.[5] The underbrace text on the gammas mentions a third familiar result: the gamma in the hybrid regression equals the simple gamma from (1). This is because the additional regressor in (3) relative to (1), $u_{1,m,t}$, is orthogonal on the original regressor, s_t; so adding this regressor to the simple equation will not affect the original currency-exposure coefficient. (The result $\gamma_{3,j} = \gamma_{1,j}$ can of course be shown explicitly by working out the expression $\gamma_{2,j} + \beta_{2,j}\gamma_{1,m}$.)

[5] The theorem says that if one first regresses both R_j and R_m on s, and then the residuals $e_{1,j}$ on $e_{1,m}$, then one gets the multiple coefficient $\beta_{2,j}$ and its t-statistic without having to run an explicit multiple regression. It is easily shown that if one needs just the coefficient then one orthogonalization actually suffices, and that including s as an additional regressor next to $e_{1,m}$ is a substitute for first orthogonalising R_j on s.

Table 1: **Exposure estimators and their standard errors**

equation to be estimated	exposure estimator	conventional variance of estimate
$R_{j,t} = \alpha_{1,j} + \gamma_{1,j} s_t + u_{1,j,t}$	$\frac{\widehat{\text{cov}}(R_j\,,s)}{\widehat{\text{var}}(s)}$	$\frac{\widehat{\text{var}}(u_{1,j})}{\sum_t (s_{j,t}-\overline{s})^2} = \frac{\widehat{\text{var}}(u_{2,j})+\beta^2_{2,j}\widehat{\text{var}}(u_{1,m})}{\sum_t (s_{j,t}-\overline{s})^2}$
$R_{j,t} = \alpha_{2,j} + \gamma_{2,j} s_t + \beta_{2,j} R_{m,t} + u_{2,j,t}$	$\frac{\hat{\gamma}_{1,j}-\hat{\beta}_{4,j}\hat{\gamma}_{1,m}}{1-\rho^2_{m,s}}$	$\frac{\widehat{\text{var}}(u_{2,j})}{\sum_t (s_{j,t}-\overline{s})^2(1-\hat{\rho}^2_{m,s})}$
$R_{j,t} = \alpha_{3,j} + \gamma_{3,j} s_t + \beta_{3,j} u_{1,m,t} + u_{3,j,t}$	$\frac{\widehat{\text{cov}}(R_j\,,s)}{\widehat{\text{var}}(s)}$	$\frac{\widehat{\text{var}}(u_{2,j})}{\sum_t (s_{j,t}-\overline{s})^2}$

The issue of the paper is the sense, if any, behind the hybrid regression. Obviously, the purpose is neither to detect whether R_m has any influence over and above s or *vice versa*, nor to identify optimal hedge ratios when both currency and market-index futures are available: for those purposes, the standard multiple regression would have been used. Instead, the rationale must have been to come up with statistically more reliable gamma's without letting the market factor subsume the individual stocks' exposures. The estimators and conventional sampling errors for each gamma, in terms of the parameters and the (unobservable) error terms of the regular multiple-regression, are given in Table 1. In that table, $\widehat{\text{cov}}$ and $\widehat{\text{var}}$ denote sample moments, like $\widehat{\text{cov}(x,y)} = \sum_{t=1}^{N}[(y_t-\overline{y})(x_t-\overline{x})]/(N-1)$, and $\hat{\rho}^2_{m,s}$ denotes the sample squared correlation between the two regressor. The beta referred to in the estimator for the second equation is the familiar market-model beta,

$$R_j = \alpha_{4,j} + \beta_{4,j} R_m + u_{4,j}. \tag{6}$$

From the table it seems that the SE of $\widehat{\text{cov}}(R_j\,,s)/\widehat{\text{var}}(s)$ depends on whether it is estimated via the simple regression or the hybrid one, even though in any conceivable sample the two regressions always generate exactly the same number.

At the risk of treading too familiar a path, let us quickly review some of the theory behind the SE's. Much of basic regression theory starts with non-stochastic regressors. Suppose that $R_{j,t}$ measures yield in the t-th hydroponic test bed, $R_{m,t}$ the amount of nutrients administered, and s_t the amount of light administered. Familiarly, the simple regression coefficient $\gamma_{1,j}$ would fail to measure the partial effect of lighting if, due to a careless design, s and R_m are correlated across test beds. Even its SE would be misleading because it would treat all yields not explained by the amount of lighting as utterly unpredictable noise, while in reality part of it stems from nutrient dosage, R_m. This matters: R_m is fixed by the researcher rather than being uncontrollable white noise; and its effect on R_j (and hence on γ_1) can effectively be taken into account in both the current sample and in any out-of-sample prediction. The multiple regression output does take care of both aspects. If the regressors are correlated, he simple regression coefficient would be relevant only if, for some reason, R_m cannot be observed by the researcher in the current sample or in later predictions.

In the case of random regressors, much of the above can be salvaged via an interim step. The interim step is that, conditional on the observed values of the regressors, the SE of the multiple regression coefficients would still work. For instance, in hypothetical Monte-Carlo experiments where only the residuals are re-sampled, it does not really matter that the values of the regressors were chosen, once and for all, via a random generator rather than by the researcher: what does matter is that fixed regressors cannot contribute to variability in the coefficients across Monte-Carlo samples. The unconditional SE is then obtained by taking expectations of the conditional one. Under standard assumptions, a conditional SE on average produces a fairly correct estimate of the unconditional SE. One can be lucky or unlucky with the sample's second moments for the regressors, but on average the computed SE's still work. The multiple coefficients still sort out the effects originating from R_m and s, and the SE's take into account both the benefits of lower residuals and the possible problems stemming from of correlated regressors.

In the case of orthogonalized regressors all of the above still works for the SE conditional on the sample, but the unconditional results would only hold if the orthogonalizing coefficient, $\gamma_{1,m}$, were non-random. To show this, we start from the uncontroversial multiple regression, (2), and write it in matrix form, denoting $\mathbf{B}_2 = (\beta_{2,j}, \gamma_{2,j})'$ and $\mathbf{X} = (\mathbf{R_m}, \mathbf{s})$.[6] In the next line we linearly transform the regressors, postmultiplying $\mathbf{X}$ by a 2×2 full matrix $\mathbf{G}$ and taking into account that there must be an offsetting correction $\mathbf{G}^{-1}$ in the coefficients:

$$\begin{aligned} \mathbf{R}_j &= \mathbf{X}\mathbf{B}_2 + \mathbf{u}_{2,j}, &(7)\\ &= [\mathbf{X}\mathbf{G}][\mathbf{G}^{-1}\mathbf{B}_2] + \mathbf{u}_{2,j}. &(8)\end{aligned}$$

Thus, the regression coefficients $\mathbf{B}_3$ w.r.t. the rehashed regressors $\mathbf{XG}$ are given by

$$\mathbf{B}_3 = \mathbf{G}^{-1}\mathbf{B}_2, \tag{9}$$

[6] We ignore the intercept. Think of demeaned returns, or, in a two-factor InCAPM context, of excess returns.

for any $\mathbf{G}$. We are interested in one specific transformation,

$$\mathbf{G} = \begin{bmatrix} 1 & 0 \\ -\gamma_{1,m} & 1 \end{bmatrix} \Rightarrow \mathbf{G^{-1}} = \begin{bmatrix} 1 & 0 \\ \gamma_{1,m} & 1 \end{bmatrix}. \tag{10}$$

The notation, above, by omitting hats, again refers to population values but also holds for sample counterparts in the case of Method-of-Moment estimators like OLS.

We now consider the variance-covariance matrix of the estimation errors in $\hat{\mathbf{B}}_3$, denoted $\mathbf{V}(\hat{\mathbf{B}}_3)$, and we link this to the variance matrix of the orthodox regression. First consider the SE of the hybrid regression conditional on the regressors $\mathbf{X}$. Below, we start by noting that, for given $\mathbf{X}$, $\mathbf{G}$ is nonrandom and can therefore be taken out of $\mathbf{V}(\hat{\mathbf{G}}^{-1}\hat{\mathbf{B}}_2)$. It then suffices to fill out the variance-covariance matrix of $\hat{\mathbf{B}}_2$—we use $\hat{\sigma}^{-2}_{m|s}$ to denote $1/\widehat{\text{var}}(u_{1,m})$—and simplify:

$$\begin{aligned}
\mathbf{V}(\hat{\mathbf{B}}_3|X) &= \mathbf{V}(\hat{\mathbf{G}}^{-1}\hat{\mathbf{B}}_2|\mathbf{X}), \\
&= \hat{\mathbf{G}}^{\mathbf{-1}}\mathbf{V}(\hat{\mathbf{B}}_2|\mathbf{X})[\hat{\mathbf{G}}^{\mathbf{-1}}]', \qquad (11)\\
&= \begin{bmatrix} 1 & 0 \\ \hat{\gamma}_{1,m} & 1 \end{bmatrix} \frac{1}{N-1} \begin{bmatrix} \hat{\sigma}^{-2}_{m|s} & -\hat{\gamma}_{1,m}\frac{1}{N-1}\hat{\sigma}^{-2}_{m|s} \\ -\hat{\gamma}_{1,m}\hat{\sigma}^{-2}_{m|s} & \widehat{\text{var}}(s)^{-1} - \hat{\gamma}^2_{1,m}\hat{\sigma}^{-2}_{m|s} \end{bmatrix} \begin{bmatrix} 1 & \hat{\gamma}_{1,m} \\ 0 & 1 \end{bmatrix} \widehat{\text{var}}(u_{2,j}), \\
&= \frac{1}{N-1} \begin{bmatrix} \hat{\sigma}^{-2}_{m|s} & -\hat{\gamma}_{1,m}\hat{\sigma}^{-2}_{m|s} \\ 0 & \widehat{\text{var}}(s)^{-1} \end{bmatrix} \begin{bmatrix} 1 & \hat{\gamma}_{1,m} \\ 0 & 1 \end{bmatrix} \widehat{\text{var}}(u_{2,j}), \\
&= \frac{1}{N-1} \begin{bmatrix} \hat{\sigma}^{-2}_{m|s} & 0 \\ 0 & \widehat{\text{var}}(s)^{-1} \end{bmatrix} \widehat{\text{var}}(u_{2,j}), \\
&= \frac{1}{N-1} \begin{bmatrix} \widehat{\text{var}}(s) & 0 \\ 0 & \widehat{\text{var}}(u_{1,m}) \end{bmatrix}^{-1} \widehat{\text{var}}(u_{3,j}). \qquad (12)
\end{aligned}$$

This indeed is the variance-covariance matrix of the estimates of the regression of R_j on s and $\hat{u}_{1,m}$. But the usual next step fails: the above is not an unbiased estimate of the unconditional SE. When $\mathbf{X}$, and therefore $\hat{\mathbf{G}}$, are random, $\hat{\mathbf{G}}$ can no longer be factored out of $\mathbf{V}(\hat{\mathbf{B}}_\mathbf{3})$ as we do in Equation (11). Stated differently, if we still factor out $\hat{\mathbf{G}}$ regardless—which is what we do if we accept the SE in Equation (12)—we ignore the variance of $\hat{\mathbf{G}}$ and its covariance with $\hat{\mathbf{B}}_\mathbf{2}$ via $\hat{\mathbf{X}}$.

Explicitly working out these extra (co)variance terms is tedious, but the final result can be obtained via a simple shortcut. If $\hat{\mathbf{G}}$ is not fixed and not independent of $\hat{\mathbf{B}}_\mathbf{2}$, we can first work out the product $\hat{\mathbf{G}}^{-1}\hat{\mathbf{B}}_2$ *inside* the $\mathbf{V}$ operator. We already know that this produces the simple gamma and the multivariate beta estimates:

$$\begin{aligned}
\mathbf{V}(\hat{\mathbf{B}}_3) &= \mathbf{V}(\hat{\mathbf{G}}^{-1}\hat{\mathbf{B}}_2), \\
&= \mathbf{V}(\hat{\beta}_{1,j}, \hat{\gamma}_{2,j}). \qquad (13)
\end{aligned}$$

This says that, in a regression where the market-return regressor has been orthogonalized on the exchange-rate regressor, the SE for the exposure coefficient is the one from the simple regression. The intuition of course is that the procedure deliberately cuts out the mechanism for which the multiple regression is useful: sorting out the interactions between the regressors. Instead, the additional regressor is first doctored so as to guarantee that its inclusion will never lead to any revision of the simple regression coefficient for the regressor of interest, s. In light of this, it is inevitable that inclusion of this doctored variable cannot really improve the SE.

We now turn to a much briefer discussion of the case where the orthogonalisation is done the other way: the market-correlated component is first taken out of s.

2 Orthogonalising the Variable of Interest on the Control Variable

If s is orthogonalised on R_m using $s = \alpha_{4,s} + \beta_{4,s} R_m + u_{4,s}$, then its coefficient is the same as it would have been in the straightforward multiple regression, and so is its SE. The first claim is analogous to our earlier result that $\beta_{2,j} = \beta_{3,j}$. The second claim follows from the inverse of the covariance matrix in the multiple regression,

$$\mathrm{E}(\mathbf{X}'\mathbf{X})^{-1} \quad = \quad \frac{1}{N-1}\begin{bmatrix} \hat{\sigma}^{-2}_{s|m} & -\hat{\beta}_{4,s}\hat{\sigma}^{-2}_{s|m} \\ -\hat{\beta}_{4,s}\hat{\sigma}^{-2}_{s|m} & \widehat{\mathrm{var}}(R_m)^{-1} - \hat{\beta}^2_{4,s}\hat{\sigma}^{-2}_{s|m} \end{bmatrix}. \tag{14}$$

The first element is exactly the same as the first element in the regression with s orthogonalized on R_m. To close the argument, note that the residual variances of the multiple and this hybrid regression are the same. Thus, the first element of $\mathbf{V}(\mathbf{X})^{-1}\mathrm{var}(e)$ is unaffected by the orthogonalization.[7]

3 Monte Carlo simulations

The Monte-Carlo simulations in Tabel 3 illustrate all of the above. In each set of simulations we generate 10,000 times series of 50 records $\{R_j, R_m, s\}$ each, as follows. The independent random variables are s, $u_{1,m}$ and $u_{2,j}$. From these we construct $R_m = \gamma_{1,m}s + u_{1,m}$ and $R_j = \gamma_{2,j}s + \beta_{2,j}R_m + u_{2,j}$. Next we estimate $R_m = \hat{\alpha}_{1,m} + \hat{\gamma}_{1,m}\, s + e_{1,m}$ and retrieve $e_{1,m}$, the estimates of $u_{1,m}$. We then run the three regressions discussed in the text, plus a variant of the hybrid, labeled regression 5, where in the orthogonalisation step we use the true market exposure $\gamma_{1,m}$ rather than the sample's estimate. We produce three sets of simulations. In

[7] The Frisch-Waugh requirement that one also orthogonalize R_j on R_m is necessary only if the second-pass regression is also a simple one; here R_m is included into the regression.

Table 2: **Monte-Carlo simulation results**

Equation	true γ_j	$\overline{\hat{\gamma}_j}$	$\overline{\text{var}}_{\text{OLS}}(\hat{\gamma}_j)$	$\text{var}(\hat{\gamma}_j)$	var ratio
Simulation 1					
$R_{j,t} = \alpha_{2,j} + \gamma_{2,j} s_t + \beta_{2,j} R_{m,t} + u_{2,j,t}$	1.00	0.996	.1706	.1667	0.98
$R_{j,t} = \alpha_{1,j} + \gamma_{1,j} s_t + u_{1,j,t}$	2.25	2.249	.0864	.0855	0.99
$R_{j,t} = \alpha_{3,j} + \gamma_{3,j} s_t + \beta_{3,j} e_{1,m,t} + u_{3,j,t}$	2.25	2.249	.0651	.0855	1.31
$R_{j,t} = \alpha_{5,j} + \gamma_{5,j} s_t + \beta_{5,j} u_{1,m,t} + u_{5,j,t}$	2.25	2.249	.0665	.0668	1.00
Simulation 2					
$R_{j,t} = \alpha_{2,j} + \gamma_{2,j} s_t + \beta_{2,j} R_{m,t} + u_{2,j,t}$	1.00	0.999	.0122	.0121	1.01
$R_{j,t} = \alpha_{1,j} + \gamma_{1,j} s_t + u_{1,j,t}$	2.25	2.249	.2762	.2744	0.99
$R_{j,t} = \alpha_{3,j} + \gamma_{3,j} s_t + \beta_{3,j} e_{1,m,t} + u_{3,j,t}$	2.25	2.249	.0106	.2744	25.89
$R_{j,t} = \alpha_{5,j} + \gamma_{5,j} s_t + \beta_{5,j} u_{1,m,t} + u_{5,j,t}$	2.25	2.250	.0109	.0109	1.00
Simulation 3					
$R_{j,t} = \alpha_{2,j} + \gamma_{2,j} s_t + \beta_{2,j} R_{m,t} + u_{2,j,t}$	1.00	0.989	.9856	.9629	0.98
$R_{j,t} = \alpha_{1,j} + \gamma_{1,j} s_t + u_{1,j,t}$	2.25	2.249	.1382	.1377	1.00
$R_{j,t} = \alpha_{3,j} + \gamma_{3,j} s_t + \beta_{3,j} e_{1,m,t} + u_{3,j,t}$	2.25	2.249	.1328	.1378	1.04
$R_{j,t} = \alpha_{5,j} + \gamma_{5,j} s_t + \beta_{5,j} u_{1,m,t} + u_{5,j,t}$	2.25	2.249	.1357	.1363	1.00

Key In each simulation we generate 10,000 samples, each of 50 records $\{R_j, R_m, s\}$, as follows. The independent random variables are s, $u_{1,m}$ and $u_{2,j}$. From these we construct $R_m = \gamma_{1,m} s + u_{1,m}$ and $R_j = \gamma_{2,j} s + \beta_{2,j} R_m + u_{2,j}$. Lastly we estimate $R_m = \hat{\alpha}_{1,m} + \hat{\gamma}_{1,m} R_m + e_{1,m}$ and retrieve $e_{1,m}$,. We then run the three regressions discussed in the text, plus a variant of the hybrid, regression 4, where we use the true market exposure rather than the sample's estimate. The assumed *p.a.* parameter values, along with some implied numbers, are as follows:

	Assumed values						Implied parameters							
							volatilities		market model		Eq (1) for m		Eq (1) for j	
	σ_s	$\sigma_{u_{1m}}$	$\sigma_{u_{2j}}$	γ_{1m}	γ_{2j}	β_{2j}	σ_j	σ_m	β_{4j}	ρ^2_{jm}	γ_{1m}	ρ^2_{ms}	γ_{1j}	ρ^2_{js}
S1	.20	.20	.35	1.25	1.00	1.00	.41	.21	1.06	.27	0.25	.06	0.50	.06
S2	.20	.50	.10	1.25	1.00	1.00	.56	.68	1.16	.91	1.25	.20	2.25	.44
S3	.20	.10	.50	1.25	1.00	1.00	.68	.27	1.70	.45	1.25	.86	2.25	.44

For each equation we show the mean of the 10,000 $\hat{\gamma}_{.,j}$s, the average of the 10,000 error variances predicted by the regression program, and the cross-sectional variance of the 10,000 estimated gamma's. The ratio of the last two is then shown under the heading "var ratio".

the first, we choose a realistic set of parameters producing a moderate bias in the estimated variance, while the other two are characterized by a much stronger or much weaker bias, respectively. The assumed *per annum* parameter values, along with some implied numbers, are shown in the Key to the Table. In simulation S1, the numbers are calibrated to what one gets with monthly data: *per annum* volatilities 20 and 40 for market and stock, respectively; a market model that explains about one-quarter of the return variability; and a weak exposure effect. In simulation S2, the market factor and the gamma's are overemphasized relative to the ideosyncratic variance, resulting in quite high ρ^2s for the simple regressions and quite precise estimates. In S3 the numbers are swapped, producing a high-power exposure regression with little genuine role for the market and, as a result, more imprecise estimates. For each equation we show the mean of the 10,000 $\hat{\gamma}_{.,j}$s, the average of the 10,000 error variances predicted by the regression program, and then the cross-sectional variance of the 10,000 estimated gamma's, a reliable estimate for the true unconditional variance of the estimate. The last column then shows the ratio of this true variance of the exposure estimate over the average variance produced by OLS; any non-unit value of course signals that the regression output cannot be trusted.

In the two orthodox regressions, the multivariate and the simple, the OLS-computed variances match the true variability across samples quite well. In the regression with the doctored data, the third one, the regression program claims to come up with a SE that is even better than the multivariate while preserving the simple-regression estimate, but this SE underestimates the true one. The theoretical ratio of actual estimation variance over calculated variance for the third regression, which from Table 1 equals $1 + \beta_{2,j}^2 \text{var}(u_{1,m})/\text{var}(u_{2,j})$, equals 1.32 in S1, 26 in S2, and 1.02 in S3 when calculated from the (known) population parameters; the other variance ratios should all be equal to unity. This is close to what we see in the average variances.

In S1, variance ratios tend to be somewhat below their predicted values. To see to what extent this reflect a systematic effect rather than randomness, we add simulations S2 and S3 where estimates are much more precise or less precise, respectively, than in S1 and where the potential of randomness in the variance ratios is accordingly higher or lower. We find no traces of bias in the tight case, S2, and larger but unsystematic deviations in scenario S3, suggesting that it all comes down to randomness.

The results for the fourth regression, where the orthogonalisation uses the true $\gamma_{1,m}$ rather than the sample's estimate, clearly illustrate that the problem originates from the variability of the market's gamma across samples: if there had been no such variability, then the SE with the orthogonalized regression would have been right on target. Unfortunately, given the weakness of the link between stock returns and exchange-rate changes, in reality the variability of the market's gamma estimate is high.

4 Conclusion

By orthogonalizing the market return on the exchange rate one seems to have the best of both worlds: the market factor cannot subsume part of the forex exposure present in a stock's return, and the SE of the estimate beats both the simple- and the multiple-regression SE's.

This last effect is illusory: since in any particular sample the simple and the pseudo-multiple regression coefficients are always equal to one another, their precision must be identical too. Technically, the source of the problem is that the uncertainty about the market's exposure estimate is left out of the calculated SE.

How large the effect is in a real-world situation depends on the sample, but the order of magnitude is easily calculated. Volatilities are about 0.20 *p.a.* for the market, and 0.30 to 0.40 for stocks (Hull, 1993). Low-volatility stocks tend to be large and low-beta, and *vice versa*: Fama (1976) reports an average $\beta = 0.61$ and $\rho^2 = 0.20$ for the 30 largest stocks, and $\beta = 1.00$ and $\rho^2 = 0.27$ for average stocks. With low ρ^2's for the Dumas regression, the ratio of true to reported variance, $1 + \beta_{2,j}^2 \sigma_{1,m}^2 / \sigma_{2,j}^2$, is about 1.25 for Fama's large stocks, 1.32 for Fama's average stocks, and between 1.56 and 2.25 for stocks with beta 1.5, depending on volatility (0.30 or 0.40). So the divergence can easily be as large as the difference between e.g. Dickey-Fuller and regular critical values and should not be ignored, especially as the estimator has no other benefits.

References

ADLER, M. and B. DUMAS, 1983, International Portfolio Choice and Corporation Finance: a. Synthesis, *The Journal of Finance*, 38, 925-984

BODNAR, G.M. and M. H. F. WONG, 2003. Estimating Exchange Rate Exposures: Issues in Model Structure, *Financial Management* 32(1), 35-63.

BRIS A.; KOSKINEN Y. and PONS-SANZ V.P., 2002. Corporate Financial Policies and Performance Around Currency Crises, Yale ICF Working Paper No. 00-61

DUMAS, B., 1978. The theory of the trading firm revisited, *Journal of Finance* 33(3): 1019-1029.

ENTORF, H., J. MOEBERT and K. SONDERHOF, 2006. The Foreign Exchange Exposure of Nations, Darmstadt UT, Working paper.

FRISCH, R. and F. WAUGH, 1933, Partial time regressions as compared with individual trends, *Econometrica*, 45, 939-953.

GLAUM, M., M. BRUNNER and H. HIMMEL, 2000. the DAX and the Dollar: the economic exchange rate exposure of German corporations, *Journal of International Business Studies*, 31(4), 715-724

HAGELIN N. and PRAMBORG B., 2002. Hedging Foreign Exchange Exposure: Risk Reduction from Transaction and Translation Hedging. *Journal of international financial management and accounting*, vol.15, pp.1-20.

JAYASINGHE P. and B. PRAMBORG, 2002. Hedging Foreign Exchange Exposure: Risk Reduction from Transaction and Translation Hedging, *Journal of International Financial Management and Accounting*, 15, 1-20

JOHNSON, L.L., 1960: The theory of hedging and speculation in commodity futures, *Review of Economic Studies* 27, 139-151

JORION, P., The Exchange Rate Exposure of US Multinationals, *Journal of Business* (July 1990), 331-345.

PRIESTLEY R. and ODEGAARD B.A.(2002). New Evidence on Exchange Rate Exposure, Norwegian school of management working paper

PRITAMANI M.D., SHOME D.K. and SINGAL, 2004. Foreign exchange exposure of exporting and importing firms, *Journal of Banking and Finance*, vol.28(7), pp.1697-1710.

REES, W. and S. C. UNNI, Exchange Rate Exposure among European Firms: Evidence from France, Germany and the UK. *Accounting and Finance*, 45(3) November 2005, pp. 479-497

SERCU, P., A generalisation of the international asset pricing model, 1980, *Journal de l'Association Française de Finance*, 1(1), pp. 91 - 135.

STEIN, J.L. 1961, The simultaneous determination of spot and futures prices, *American Economic Review* 51, 1012-1025