Available online at www.sciencedirect.com
SCIENCE DIRECT®

Finance Research Letters 2 (2005) 131–151

Finance Research Letters

www.elsevier.com/locate/frl

Proxy-quality thresholds: Theory and applications ☆

Timothy Erickson [a], Toni M. Whited [b,*]

[a] *Bureau of Labor Statistics, Postal Square Building, Room 3105, 2 Massachusetts Avenue, NE, Washington, DC 20212-0001, USA*

[b] *University of Wisconsin, Business School, 975 University Avenue, Madison, WI 53706-1323, USA*

Received 3 March 2005; accepted 20 April 2005

Available online 13 June 2005

Abstract

We consider alternative models of a regression containing a proxy for an unobserved regressor. For each model at most two pieces of prior information are necessary to determine the sign of any regressor coefficient: the sign of the partial correlation between the proxy and the unobserved regressor, and a lower bound on the partial or simple correlation between the proxy and the unobserved regressor. We apply our technique to investment and leverage regressions that contain a proxy for the incentive to invest. In both cases proxy quality must be high for the coefficient of interest to be non-zero.

Published by Elsevier Inc.

JEL classification: G31; E22; C29; C49

Keywords: External finance constraints; Investment; Errors-in-variables; Coefficient sign; Prior information

Instrumental variables, or other types of additional identifying information, are often unavailable for consistent estimation of regressions containing proxy variables, which are well known to render OLS estimation inconsistent. Frequently, however, only coefficient *signs* are of interest. In this case although additional prior information or assumptions are

☆ This paper previously circulated under the title "New Results on Signing Coefficients in Regressions Containing a Proxy, with an Application to Financial Constraints and Investment." Computer programs for all calculations are available from the authors upon request.

* Corresponding author.
E-mail addresses: erickson.timothy@bls.gov (T. Erickson), twhited@bus.wisc.edu (T.M. Whited).

1544-6123/$ – see front matter Published by Elsevier Inc.
doi:10.1016/j.frl.2005.04.004

needed to draw inferences, they need not strong enough to identify coefficient *values*. We develop an econometric framework that inputs just enough prior information to identify coefficient signs.

We start with a linear regression with one unobserved regressor and an arbitrary number of perfectly observed regressors. A proxy is available for the unobserved regressor. We consider several assumption sets corresponding to whether the measurement error is correlated with some, any, or all of the other model variables. For each assumption set we show that an index of measurement quality for the unobserved regressor must surpass a threshold in order for a coefficient to retain the sign obtained via OLS. We express this index in terms of either the partial or simple correlation between the unobservable regressor and its proxy. For any coefficient and for either measure of proxy quality, we can compute multiple thresholds, corresponding to our different assumption sets. We suggest reporting all these thresholds, so that readers can use their own prior notions to decide whether the data set is informative about true coefficient signs. These threshold estimates will be particularly interesting if they are either near zero or one. In the first instance it will be hard to accept the hypothesis that the coefficient of interest is zero, and in the second it will be hard to reject this hypothesis. This situation is loosely analogous to that of a t-statistic, which is usually only interesting if it is either very low or very high. Finally, an additional contribution of the paper is the computation of the variances of these threshold bounds. To our knowledge, none of the previous research in errors-in-variables bounds has addressed the issue of threshold variances.

We provide two applications of our technique, both highlighting its computational simplicity and minimal assumption requirements. First, we examine the effects of external finance constraints on investment. Starting with Fazzari et al. (1988), most empirical studies of this issue have examined the sensitivity of investment to cash flow as an indicator of finance constraints. As summarized in Hubbard (1998), these efforts have shown that for groups of firms identified as financially constrained, investment responds strongly to cash flow, even after controlling for a proxy for the incentive to invest. Recently however, several papers have questioned these results, arguing that the usual control for the incentive to invest, Tobin's q, contains substantial measurement error. For example, Erickson and Whited (2000) use measurement-error consistent estimators on investment-q-cash flow regressions, finding that positive cash-flow coefficients produced by OLS are not robust to the use of measurement error remedies, even for financially constrained firms.

Our intent is to determine whether the message in Erickson and Whited (2000) is robust to relaxation of their assumptions. We find this to be the case. Under our less restrictive assumptions the proxy quality thresholds must often be quite high and even near one for the cash-flow coefficient to be positive. Further, the thresholds must also be high to infer a difference in cash-flow coefficients between groups of constrained and unconstrained firms.

Next, we examine the anomalous evidence in Rajan and Zingales (1995) that leverage is decreasing in lagged liquidity. This result counters the intuition from the static trade-off model that higher profits mean more dollars for debt service, more taxable income to shield, and therefore higher target leverage. Because Rajan and Zingales use a proxy for the incentive to invest, our technique is applicable. We find that our threshold must be implausibly high before one can infer a negative coefficient on lagged liquidity.

Our paper is organized as follows. Section 1 describes related literature. Section 2 outlines our econometric model and summarizes our results. Sections 3 and 4 present our investment and leverage applications, and Section 5 concludes. Proofs are in Appendix A.

1. Related econometric literature

We build upon the work of Krasker and Pratt (1986), who show that coefficient signs are indeterminate in a one-mismeasured-regressor errors-in-variables model if the measurement error is correlated with all other model variables. However, they show that coefficient signs are determined by the additional information that the simple correlation between the unobserved regressor and its proxy exceeds a threshold. Our first contribution is to determine coefficient signs by using additional information on quantities other than this simple correlation, though we use this correlation as well. Our second contribution is to examine four alternative sets of assumptions concerning the structure of the measurement error model:

(a) the measurement error (the difference between the proxy and the unobserved regressor) may be correlated with the regression disturbance term and one or more regressors (including the unobserved regressor itself);
(b) the measurement error may be correlated with the disturbance, but is uncorrelated with every regressor;
(c) the measurement error may be correlated with one or more regressors, but is uncorrelated with the disturbance;
(d) the measurement error is uncorrelated with all other variables in the model.

Assumption set (a) is Krasker and Pratt's, while set (d) is the classical errors-in-variables model. Assumption sets (b) and (c) invoke an intermediate number of restrictions on the correlations between the measurement error and other model variables, which, relative to the Krasker–Pratt model, constitute additional information that may assist inference. Like Krasker and Pratt, we also consider additional information in the form of a lower bound on a measure of proxy quality. We use two such measures: the simple correlation between the proxy and the unobserved regressor and the corresponding partial correlation that controls for movements in the perfectly measured regressors. Finally, we consider additional information not previously exploited: knowledge of the sign of a coefficient in the regression of the *proxy* on the unobserved regressor and all perfectly observed regressors. We reproduce their result that the coefficient on the proxy agrees in sign with the coefficient on the unobserved regressor if the simple correlation between the proxy and unobserved regressor exceeds a computable threshold. We state an analogous result in terms of the partial correlation. We do not reproduce the Krasker–Pratt result that the sign of the coefficient on any perfectly observed regressor is determined if the simple correlation between the proxy and the unobserved variable exceeds a (generally different) threshold, since that result is not closed-form. Instead, we give a closed-form threshold based on the additional assumption that the perfectly observed regressor in question has a known sign in the regression of the proxy on all regressors. We give an example showing that our threshold can be substantially lower than that of Krasker–Pratt, indicating the potential value of this type of additional information.

The seminal result in this literature is the well-known reverse-regression bound for the classical errors-in-variables model of Gini (1921) and Frisch (1934). Klepper and Leamer (1984) extend the classical errors-in-variables model and versions of the Gini–Frisch bound to the case of multiple mismeasured regressors, and Leamer (1987) derives bounds in the context of systems of equations. Erickson (1993) provides results from bounding the correlation between the measurement error and equation error in a linear regression with one mismeasured regressor. Klepper (1988) and Bollinger (1996) calculate bounds in the context of dichotomous regressors. Kroch (1988) explores the trade-off between model restrictions and proxy quality in determining the width of an interval that necessarily contains the coefficient of interest. Although we also explore this trade-off, our work can be distinguished from his in that Kroch (like Krasker and Pratt) does not give closed form expressions for inference about the coefficients on perfectly observed regressors.

2. Model and taxonomy of results

2.1. Model

Let (y_i, x_i, z_i) be an observable vector and $(u_i, \varepsilon_i, \chi_i)$ be an *un*observable vector. All variables are scalar except $z_i \equiv (z_{i1}, \ldots, z_{ik})$. We measure all variables as deviations from means.

Assumption 1.

(i) (y_i, x_i, z_i) *is related to* $(u_i, \varepsilon_i, \chi_i)$ *and unknown parameters* $\alpha \equiv (\alpha_1, \ldots, \alpha_k)'$ *and* β *according to*

$$y_i = \chi_i \beta + z_i \alpha + u_i, \tag{1}$$

$$x_i = \chi_i + \varepsilon_i; \tag{2}$$

(ii) $(u_i, \varepsilon_i, \chi_i, z_i)$*,* $i = 1, \ldots, n$*, is an i.i.d. sequence*;

(iii) $E(u_i \chi_i) = E(u_i z_{ij}) = 0$*,* $j = 1, \ldots, k$;

(iv) *the covariance matrix* $\operatorname{var}(u_i, \varepsilon_i, \chi_i, z_i)$ *is positive definite.*

Assumption 1 gives the Krasker and Pratt model. It is our most general model, allowing the measurement error, ε_i to be correlated with any other variable. Rather than work directly with this model, we use an alternative but equivalent representation obtained by replacing ε_i with the residual from its projection on (χ_i, z_i). Specifically, we replace (2) with

$$x_i = \chi_i \delta + z_i \gamma + e_i, \tag{3}$$

where

$$e_i = \varepsilon_i - \chi_i \delta_1 - z_i \gamma, \tag{4}$$

$$\binom{\delta_1}{\gamma} = \left(E\left[(\chi_i, z_i)'(\chi_i, z_i)\right]\right)^{-1} E\left[(\chi_i, z_i)'\varepsilon_i\right], \tag{5}$$

$$\delta = 1 + \delta_1. \tag{6}$$

This construction guarantees that e_i, like u_i, is orthogonal to the regressors (χ_i, z_i). Any correlations between the original error ε_i and the regressors are now captured by the slope parameters δ and γ. The literature on bounds in measurement-error models often supposes individuals can input prior beliefs about unknown correlations; in this paper we sometimes also suppose persons can a priori specify the sign of a particular element of (δ, γ). We feel this is a reasonable supposition, since economists are well practiced at discussing the signs of multiple regression coefficients. For example, we regard the following assumption as highly plausible, since the most likely reason one would employ x_i as a proxy for χ_i is the belief that these two variables are positively correlated if z_i is held constant:

Assumption 2. $\delta > 0$.

To state an additional assumption that greatly simplifies the derivation and statement of results, let (b, a') be the coefficient vector from the projection of y_i on (x_i, z_i).

Assumption 3. *Every element of* (b, a') *is positive.*

Note that, if necessary, negative coefficients can be made positive by multiplying their regressors by -1. Assumption 2 implies that if the proxy x_i is so multiplied, then so is χ_i.

To simplify our derivations and to provide the applied researcher with easily computable bounds, we first "partial out" the perfectly-measured variables. Let

$$d \equiv \left[E\left(z_i' z_i\right)\right]^{-1} E\left[z_i' y_i\right], \tag{7}$$

$$m \equiv \left[E\left(z_i' z_i\right)\right]^{-1} E\left[z_i' x_i\right], \tag{8}$$

$$\mu \equiv \left[E\left(z_i' z_i\right)\right]^{-1} E\left(z_i' \chi_i\right), \tag{9}$$

where $E\left(z_i' z_i\right)$ is invertible by Assumption 1(iv). Note that $y_i - z_i d$, $x_i - z_i m$, and $\chi_i - z_i \mu$ are the residuals from projections of y_i, x_i, and χ_i on z_i. Substituting (3) into (8) gives

$$m \equiv \left[E\left(z_i' z_i\right)\right]^{-1} E\left[z_i'(z_i \gamma + \chi_i \delta + e_i)\right] = \gamma + \mu\delta. \tag{10}$$

Substituting (1) into (7) gives

$$\begin{aligned} d &\equiv \left[E\left(z_i' z_i\right)\right]^{-1} E\left[z_i'(z_i \alpha + \chi_i \beta + u_i)\right] \\ &= \alpha + \mu\beta \qquad (11) \\ &= \alpha + \left(\frac{m - \gamma}{\delta}\right)\beta, \qquad (12) \end{aligned}$$

where the third line is obtained by substituting from (10). Subtracting $z_i d$ from both sides of (1) and then substituting in (11) and rearranging gives

$$y_i - z_i d = (\chi_i - z_i \mu)\beta + u_i. \tag{13}$$

Subtracting $z_i m$ from both sides of (3) and then substituting (10) similarly gives

$$x_i - z_i m = (\chi_i - z_i \mu)\delta + e_i. \tag{14}$$

Equations (13) and (14) are, respectively, (1) and (3) in residuals form.

Equations (13) and (14) imply that the second-order moments satisfy

$$\mathrm{var}(y_i - z_i d) = \mathrm{var}(\chi_i - z_i \mu)\beta^2 + \mathrm{var}(u_i), \tag{15}$$

$$\mathrm{cov}(y_i - z_i d, x_i - z_i m) = \mathrm{var}(\chi_i - z_i \mu)\beta\delta + \mathrm{cov}(u_i, e_i), \tag{16}$$

$$\mathrm{var}(x_i - z_i m) = \mathrm{var}(\chi_i - z_i \mu)\delta^2 + \mathrm{var}(e_i). \tag{17}$$

The three moments on the left-hand side can be consistently estimated using the residuals from sample-based regressions between the observable variables, but the resulting information does not suffice to draw inferences about the six unknown quantities on the right-hand side. The assumptions of this section are so general that no restrictions are implied for the observable data. To conduct inference, one must input additional prior information about proxy quality.

We consider two measures of proxy quality: the correlation between the proxy and the true regressor, $\tau = \mathrm{corr}(x_i, \chi_i)$, and that between the corresponding residuals, $\rho = \mathrm{corr}(x_i - z_i m, \chi_i - z_i \mu)$. We will refer to τ as the simple correlation and ρ as the partial correlation. Is ρ superior or inferior to τ as a vehicle for imputing prior information? The answer depends on the relative ease of assessing prior information, a factor that is likely to vary from application to application. However, individuals who prefer or require the conceptual device of holding all else constant in order to form prior opinions about the relationship between two variables may be more comfortable dealing with the partial correlation.

With respect to this issue, note that (14) implies

$$\mathrm{cov}(x_i - z_i m, \chi_i - z_i \mu) = \mathrm{var}(\chi_i - z_i \mu)\delta, \tag{18}$$

so that ρ is positive. In contrast, τ can be negative, because (3) implies

$$\mathrm{cov}(x_i, \chi_i) = \delta\,\mathrm{var}(\chi_i) + \gamma'\,\mathrm{cov}(z_i, \chi_i), \tag{19}$$

and we have assumed no restrictions on γ' or $\mathrm{cov}(z_i, \chi_i)$. We feel this is a failing of τ, since a confident assessment about the sign of δ is likely what motivates the selection of the proxy in the first place. Further, $\tau = 0$ does not imply $\delta = 0$, nor does $\delta = 0$ imply $\tau = 0$. Fortunately, it is not necessary for our purposes to choose between ρ and τ, as the next two results will let us state our propositions in terms of either correlation. Let $R^2_{x\cdot z}$ denote the population coefficient of determination corresponding to the projection of x_i on z_i, defined by $1 - R^2_{x\cdot z} \equiv \mathrm{var}(x_i - z_i m)/\mathrm{var}(x_i)$.

Lemma 1. *If Assumption 1 holds, then*

$$\rho^2 \geqslant \frac{\tau^2 - R^2_{x\cdot z}}{1 - R^2_{x\cdot z}}. \tag{20}$$

This expression holds as an equality if $E(\varepsilon_i \chi_i) = 0$ and $E(\varepsilon_i z_i) = 0$.

Corollary 1. *If Assumption 1 holds, and the values $E(\varepsilon_i \chi_i) = 0$ and $E(\varepsilon_i z_i) = 0$ are not ruled out by additional assumptions, then for any positive number c, the smallest number c' such that $\tau^2 > c'$ implies $\rho^2 > c$ is given by $c' = R^2_{x\cdot z} + (1 - R^2_{x\cdot z})c$.*

2.2. Taxonomy

Having set up the model, we now provide the applied researcher with a set of thresholds for both of our measures of proxy quality above which a parameter of interest is ensured to be positive. Before discussing our results, we define several quantities: $r^2_{yx\cdot z} \equiv \operatorname{corr}(y_i - z_i d, x_i - z_i m)$, $a^*_j \equiv d_j - m_j b / r^2_{xy\cdot z}$, $c_I \equiv (1 + (a_j/m_j s)^2)^{-1}$, and $c_A \equiv (1 + (b/s)^2)^{-1}$, where $s = \sqrt{\operatorname{var}(y_i - z_i d)/\operatorname{var}(x_i - z_i m) - b^2}$. We also list two further assumptions.

Assumption 4. $E(\chi_i \varepsilon_i) = E(z_{ij} \varepsilon_i) = 0$, $j = 1, \ldots, k$.

Assumption 5. $E(u_i \varepsilon_i) = 0$.

Table 1 provides a summary of our proxy-quality thresholds. These results refer to a population rather than a sample. Applying the results requires substitution of sample-based estimates of the relevant variables into the threshold formulas. We group the results along three dimensions: the two parameters of interest (β and α_j), the two measures of proxy quality (ρ and τ), and our four assumption sets (a)–(d). For each threshold we list the specific assumptions and conditions on observable moments necessary to derive it. All of the results require assumptions (1) and (3); that is, the basic structure of the econometric model and the innocuous assumption that all population least-squares regression slopes are positive. The rest of the assumptions depend on the zero-correlation restrictions we invoke. We give proofs of the assertions in Table 1 in Appendix A.

The first section of Table 1 delineates the results for assumption set (a). Recall that this assumption set is the model of Krasker and Pratt, in which the measurement error may be correlated with the regression error and the regressors. The condition on the simple correlation τ for $\beta > 0$ is equivalent to the solution Krasker and Pratt (1986) give for their Problem 1; we state it here in our own notation for comparison with our original results below.[1] Our results for the parameter α_j are not equivalent to those in Krasker and Pratt, since we invoke a prior sign restriction on γ_j. The value of this restriction can be measured by the difference between the threshold for τ given here and that given by Krasker and Pratt, who do not require a sign restriction. For the case of just one perfectly

[1] Krasker and Pratt do not obtain their result using our Assumption 2. Instead, they implicitly assume $\tau > 0$. It can be shown that our Assumptions 1–3 imply $\tau > 0$ whenever $\tau^2 > R^2_{x\cdot z}$.

Table 1
Taxonomy of thresholds for proxy quality

Assumptions	Conditions	Parameter	Threshold	
			ρ^2	τ^2
(a)				
(1), (2), (3)		β	$1 - r^2_{yx\cdot z}$	$R^2_{x\cdot z} + (1 - R^2_{x\cdot z})(1 - r^2_{yx\cdot z})$
(1), (3), $\gamma_j > 0$		α_j	$\max\{c_I, c_A\}$	$R^2_{x\cdot z} + (1 - R^2_{x\cdot z}) \max\{c_I, c_A\}$
(b)				
(1), (3), (4)		β	$1 - r^2_{yx\cdot z}$	$R^2_{x\cdot z} + (1 - R^2_{x\cdot z})(1 - r^2_{yx\cdot z})$
(1), (3), (4)		α_j	$m_j^2 s^2 (a_j^2 + m_j^2 s^2)^{-1}$	$R^2_{x\cdot z} + (1 - R^2_{x\cdot z}) m_j^2 s^2 (a_j^2 + m_j^2 s^2)^{-1}$
(c)				
(1), (2), (3), (5)		β	0	0
(1), (3), (5), $\gamma_j \geqslant 0$	$a_j^* > 0$	α_j	0	0
(1), (3), (5), $\gamma_j \geqslant 0$	$a_j^* \leqslant 0$	α_j	$m_j b/d_j$	$R^2_{x\cdot z} + (1 - R^2_{x\cdot z}) m_j b/d_j$
(d)				
(1), (3), (4), (5)		β	0	0
(1), (3), (4), (5)	$a_j^* > 0$	α_j	0	0
(1), (3), (4), (5)	$a_j^* \leqslant 0$	α_j	$m_j b/d_j$	$R^2_{x\cdot z} + (1 - R^2_{x\cdot z}) m_j b/d_j$

measured regressor, z_{i1}, with data satisfying $\operatorname{var}(y_i) = \operatorname{var}(x_i) = \operatorname{var}(z_{i1}) = 1$, they table their threshold for various values of $q \equiv b/a_1$, $r \equiv \operatorname{corr}(x_i, z_{i1})$, and $\upsilon \equiv 1 - R^2_{y\cdot xz}$, where $R^2_{y\cdot xz}$ is the multiple correlation coefficient for the population regression of y_i on (x_i, z_{i1}). For $q = 1$, $r = 0.4$, and $\upsilon = 0.2$, their table gives a threshold equal to 0.743. For the same data values Proposition 2 gives 0.542.[2] Of course, the sign restriction on γ_j is also valuable because it permits a closed-form solution. Finally, note that our conditions on α do not require Assumption 2, because, as shown in Appendix A, the derivation of this threshold does not depend on δ. None of the propositions about α given below require Assumption 2 either. Similarly, none of the propositions about β require any information about γ.

The next section of the table lists results for assumption set (b), which includes the additional prior information that the measurement error is uncorrelated with any of the regressors. The value of this information can be measured by the difference between the Krasker–Pratt lower bound on τ for $\alpha_j > 0$ and that given in this second section of Table 1. Using the same given information as in the example above, our bound equals 0.259. This number is less than that given under assumption set (a), because the restriction $\gamma_j = 0$ implied by Assumption 4 is more informative than the knowledge that γ_j is positive.

Assumption set (c) drops the assumption that the measurement error is uncorrelated with the regressors, but adds the assumption that the measurement error is uncorrelated with the regression error. These assumptions ensure that $\beta > 0$ without any prior information on

[2] We square the values in Krasker and Pratt's Table 1 to make them comparable to our thresholds. Also, the other relationships between the quantities defined by Krasker and Pratt and our own are as follows: First, $r = m_1$ under the unit variance assumption. Second, straightforward algebra gives $\upsilon = s^2(1 - m_1^2)$.

proxy quality. This result can be seen by examining the equation

$$b < \frac{\beta}{\delta} < \frac{b}{r_{xy\cdot z}^2}, \tag{21}$$

derived in Appendix A. Recall that b equals the coefficient from the projection of $y_i - z_i d$ on $x_i - z_i m$, and note that $b\big/r_{xy\cdot z}^2$ equals the reciprocal of the coefficient of the projection of $x_i - z_i m$ on $y_i - z_i d$, with $r_{xy\cdot z}$ defined as the correlation between $y_i - z_i d$ and $x_i - z_i m$. These coefficients are called, respectively, direct and reverse regression estimates, and were shown by Gini (1921) and Frisch (1934) to contain β in the classical errors-in-variables model. Thus (21) is the original errors-in-variables interval bound, except that it bounds β/δ rather than β, because of the correlation between the proxy and the unobserved regressor. We now turn to the threshold for α_j. Our example above illustrates the value of prior information that the measurement error is uncorrelated with the regression disturbance and that $\gamma_j \geqslant 0$. Here, the Krasker–Pratt threshold remains 0.743, while the threshold for our assumption set (c) with the condition $a_j^* \leqslant 0$ is 0.4, and the implicit threshold with the condition $a_j^* > 0$ is zero.

Next we consider assumption set (d), which imposes both Assumptions 4 and 5, so that the measurement error is uncorrelated with any variable other than the proxy itself. We find results identical to those for assumption set (c). The similarity arises because Assumption 4 along with Eq. (5) implies $\gamma = 0$ and $\delta = 1$. Therefore, Assumption 4 imposes the restriction $\gamma_j \geqslant 0$ in assumption set (c); and $\delta = 1$ implies that the interval (21) is now precisely the "errors-in-variables bound" of Gini (1921) and Frisch (1934).

Finally, to calculate the variances of these various thresholds, we use the influence-function approach in Erickson and Whited (2002). Specifically, let θ be the vector of observable moments that is used to compute a given threshold, $g(\theta)$. For example, θ may include b, a, c_I, etc. Let $\psi(\theta)$ be the corresponding vector of influence functions for θ; that is, if $\hat{\theta}$ is a consistent estimate of θ, then the influence function, ψ, is defined as a function that satisfies $\sqrt{n}(\hat{\theta} - \theta) = n^{-1/2}\sum_{i=1}^{n} \psi + o_p(1)$. In this case, the asymptotic distribution of $\hat{\theta}$ is a zero-mean multivariate normal with covariance matrix var(ψ), and the delta-method can be used to obtain the asymptotic distribution of $g(\hat{\theta})$.

3. Investment, *Q*, and cash flow

We now use our results to examine the sensitivity of investment to cash flow in regression of investment on cash flow and Tobin's q—a proxy for the incentive to invest.[3] As explained in Erickson and Whited (2000), this proxy is far from perfect, thus motivating us to examine a sample of non-financial firms from COMPUSTAT covering the years 1990 to 1999. We select the sample by first deleting any firm-year observations with missing data, as well as those for which reported debt due in years one through five is greater than

[3] Our method is inappropriate for an investment-q-cash flow regressions if it is non-linear or if the regression error is correlated with true q. Erickson and Whited (2000) delineate assumptions under which simultaneity and non-linearity are unimportant, and their specification tests uncover no evidence of either problem.

reported total debt, and for which reported depreciation, acquisition, and sales of capital goods cannot account for reported changes in the capital stock. Third, we delete any observation for which the firm experienced a merger accounting for more than 15% of the book value of its assets. Finally, a firm must have at least two consecutive years of data to be in our sample. These procedures produce an unbalanced panel whose cross-sectional width ranges from 358 to 1952 firms.

Following Whited (1992) and Erickson and Whited (2000), we classify a firm as financially constrained if it does not have an S&P bond rating, if it never pays dividends, or if it is in the bottom third of the distribution of both total assets and the capital stock for every year that it is in our sample.[4] Finally, we split the sample depending on whether a firm has positive or negative cash flow. This experiment comes from Allayanis and Mozudmar (2004), who argue that firms with negative cash flow are financially distressed and therefore have no capacity to allow investment to respond to cash flow.

Table 2 presents the OLS direct and reverse regressions of investment on the q proxy, cash flow, the constraint indicator, and the interaction between the constraint indicator and cash flow. The indicator is one if a firm is in a constrained group and zero otherwise.[5] The financial constraints hypothesis predicts that the coefficient on the interaction term will be positive; that is, cash flow sensitivity will be higher for the constrained subsample. As is customary in the bounds literature, in presenting the results of the reverse regressions, we rearrange the regression coefficients so that investment is put back on the left side. These regressions pool the different cross sections and contain firm-level dummies to control for fixed effects that could be correlated with the incentive to invest or with cash flow.

The direct regressions in the top panel confirm some of the canonical results in the literature, but not all. All coefficients on cash flow are significant, and the coefficients on the interactions terms based on dividends and negative cash flow confirm the results in Fazzari et al. (1988) and Allayanis and Mozudmar (2004). Low-dividend firms have high sensitivity and negative-cash-flow firms have low sensitivity. However, in the bond-rating and size models, neither interaction term is significant. As in Kaplan and Zingales (1997) and Erickson and Whited (2000), these results contribute to the evidence that the magnitude and direction of differential cash flow sensitivity is sample dependent.

Before presenting our results on the thresholds, we discuss the economic rationale behind our assumptions. First, assumption sets (a) and (c) require that $\delta > 0$. In our context, this assumption implies that in a set of firms with identical cash flow, those with higher true q's will on average have higher q proxies. We view this assumption as highly plausible. Second, assumption set (a) requires that $\gamma > 0$, and assumption set (c) requires that $\gamma \geqslant 0$. In a group of firms with identical true q's, this assumption implies that those firms with higher cash flow (under (a)) or strictly higher cash flow (under (c)) have, on average, higher values for the q proxy. This idea is plausible to the extent that investors over-react to observable information, such as cash flow: see, for example, Odean (1999).

[4] Dividend payout is determined jointly with investment and is therefore correlated with the regression error. Nonetheless, we use it for comparison with the original Fazzari et al. (1988) work.

[5] Defining the dummy by observation instead of by firm produces similar results. Also, using the dummy to split the sample produces results qualitatively similar to those given below.

Table 2
Investment regressions with firm-level finance-constraint indicators

		Constraint indicator			
		Bond rating	Size	Payout	Negative CF
Direct regression					
Q Proxy	0.015	0.015	0.015	0.015	0.015
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cash Flow	0.102	0.081	0.116	0.097	0.130
	(0.011)	(0.025)	(0.013)	(0.011)	(0.022)
Interaction		0.023	−0.043	0.098	−0.112
		(0.026)	(0.024)	(0.038)	(0.026)
R^2	0.109	0.109	0.110	0.111	0.117
Reverse regression					
Q Proxy	0.272	0.272	0.274	0.274	0.286
	(0.019)	(0.019)	(0.020)	(0.020)	(0.021)
Cash Flow	−1.166	−1.103	−1.299	−1.158	−1.541
	(0.105)	(0.175)	(0.120)	(0.106)	(0.158)
Interaction		−0.072	0.372	−0.370	0.771
		(0.165)	(0.107)	(0.139)	(0.077)
R^2	0.207	0.207	0.209	0.208	0.216

Notes. Calculations are based on a sample of manufacturing firms from the combined annual and full coverage 2000 Standard and Poor's COMPUSTAT industrial files. The sample period is 1990 through 1999. CF stand for cash flow, and "interaction" refers to the interaction of cash flow with a dummy variable indicating the presence of liquidity constraints. The results from a model without an interaction term are in the first column, and the results from each of the variables used to define finance constraints are in the next four columns. The reverse regression is a regression of the proxy for the true incentive to invest on investment and the cash flow variables. The results from the reverse regressions are re-arranged to put investment on the left-hand side. All regressions are run using OLS with fixed year and firm effect. Standard errors are corrected for heteroskedasticity using the procedure in White (1980), and standard errors for the reverse regressions are calculated using the delta-method.

Table 3 contains the estimates of our proxy-quality thresholds. First, we consider assumption set (d)—that of the classical errors-in-variables model. Because the direct and reverse regressions coefficients on cash flow and the interaction terms never agree in sign, we must calculate thresholds. In the regression containing just the q proxy and cash flow, we estimate the simple and partial correlation bounds to be 0.53 and 0.43, respectively. Both bounds are estimated very precisely. In the models containing interaction terms, the thresholds for the interaction terms are somewhat smaller and less precisely estimated, especially in the case of the bond-rating model. We compare these results with those in Erickson and Whited (2000), who not only invoke assumption set (d), but assume that $\beta \neq 0$ and also impose restrictions on the distribution for true q. They estimate the squared simple correlation between marginal and observed q to be forty percent. If we were trying to defend the existence of cash-flow sensitivity using these results, the best we could say is that the data are uninformative.

Next, we allow the measurement error to be correlated with cash flow and/or "true q," as in assumption set (c). This correlation may be important because the replacement value of the capital stock deflates both the left- and right-hand side variables of the regression,

Table 3
Proxy quality thresholds for investment regressions with firm-level finance-constraint indicators

	Constraint indicator				
	Pooled model	Bond rating	Size	Payout	Negative CF
Cash flow coefficient					
Partial correlation thresholds					
(a)	0.943	0.943	0.944	0.944	0.949
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
(b)	0.904	0.927	0.898	0.909	0.900
	(0.019)	(0.041)	(0.021)	(0.019)	(0.031)
(c)	0.430	0.468	0.420	0.435	0.411
	(0.042)	(0.096)	(0.043)	(0.044)	(0.057)
(d)	0.430	0.468	0.420	0.435	0.411
	(0.042)	(0.096)	(0.043)	(0.044)	(0.057)
Simple correlation thresholds					
(a)	0.952	0.952	0.953	0.953	0.958
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
(b)	0.919	0.939	0.915	0.923	0.917
	(0.016)	(0.035)	(0.018)	(0.016)	(0.026)
(c)	0.520	0.552	0.514	0.526	0.514
	(0.034)	(0.081)	(0.035)	(0.035)	(0.045)
(d)	0.520	0.552	0.514	0.526	0.514
	(0.034)	(0.081)	(0.035)	(0.035)	(0.045)
Interaction coefficient					
Partial correlation thresholds					
(a)		0.943	0.944	0.944	0.949
		(0.008)	(0.008)	(0.008)	(0.008)
(b)		0.508	0.846	0.575	0.772
		(0.568)	(0.146)	(0.190)	(0.081)
(c)		0.200	0.363	0.221	0.300
		(0.340)	(0.119)	(0.086)	(0.058)
(d)		0.200	0.363	0.221	0.300
		(0.340)	(0.119)	(0.086)	(0.058)
Simple correlation thresholds					
(a)		0.952	0.953	0.953	0.958
		(0.007)	(0.007)	(0.007)	(0.007)
(b)		0.586	0.871	0.644	0.811
		(0.478)	(0.122)	(0.160)	(0.067)
(c)		0.327	0.466	0.346	0.421
		(0.286)	(0.100)	(0.071)	(0.044)
(d)		0.327	0.466	0.346	0.421
		(0.286)	(0.100)	(0.071)	(0.044)

Notes. Calculations are based on a sample of manufacturing firms from the combined annual and full coverage 2000 Standard and Poor's COMPUSTAT industrial files. The sample period is 1990 through 1999. The figures presented are the lower bounds for either the partial or simple correlation between the true incentive to invest and its proxy. These bounds are necessary conditions for the coefficient on either cash flow or the interaction term to retain the sign presented in Table 2, under the assumption sets (a)–(d). In (a) the measurement error may be correlated with the regression disturbance term and one or more regressors (including the unobserved regressor itself); in (b) the measurement error may be correlated with the disturbance, but is uncorrelated with every regressor; in (c) the measurement error may be correlated with one or more regressors, but is uncorrelated with the disturbance; and in (d) the measurement error is uncorrelated with all other variables. CF stand for cash flow, and "interaction coefficient" refers to the coefficient on the interaction of cash flow with a dummy variable indicating the presence of liquidity constraints. The results from a model without an interaction term are in the first column, and the results from each of the variables used to define finance constraints are in the next four columns. Standard errors, calculated using the delta method, are in parentheses under the threshold estimates.

and because this replacement value may be measured with error. In either case, because of the similarity of the results for assumption sets (c) and (d), we find identical results.

Under assumption set (b), we allow correlated regression and measurement errors, but impose a zero correlation between the measurement error and cash flow. This problem arises because investment-q regressions are derived from investment adjustment cost functions in which the regression error has a negative effect on marginal adjustment costs. (See Hayashi and Inoue (1991).) Firms with high u_i's have low marginal installation costs and tend to adopt new technologies unknown to the market. Measurement error in the observed q's of these firms is large, since the market tends to misvalue their capital. In this case we find that the simple correlation bound for a positive cash-flow coefficient ranges from 0.915 to 0.939, and that the partial correlation thresholds range from 0.586 to 0.871. Both sets of bounds are noticeably higher than those for assumption set (d) and, except in the case of the bond-rating model, are once again estimated quite precisely. Finally, under assumption set (a) we find that both thresholds must be near one to ensure that the cash-flow and interaction-term coefficients retain their sign.

Comparing this final result with those for the classical errors-in-variables model highlights the trade-off between the strength of model restrictions and the size of correlation thresholds in determining coefficient signs. We also see that relaxing the distributional and independence assumptions used by Erickson and Whited (2000) only reinforces their results. Finally this reinforcement helps interpret the evidence in Blanchard et al. (1994), Lamont (1997), and Rauh (2005) that firm investment responds to cash windfalls. Although these papers clearly show that external finance is more costly than internal, the evidence of robustness in this paper reemphasizes that these "natural experiments" shed little light on the link between finance constraints and garden-variety investment-cash flow sensitivities.

4. Leverage and liquidity

We next examine the leverage regression in Rajan and Zingales (1995) using the data in Hennessy and Whited (2005).[6] Rajan and Zingales regress the ratio of net debt to book assets on (1) the market-to-book ratio, (2) the ratio of book fixed assets to total book assets, (3) the log of sales, and (4) the lagged ratio of earnings before interest and taxes to book assets.[7] As noted in the introduction, Rajan and Zingales find a negative coefficient on the fourth regressor. Table 4 shows that we can replicate this result. However, we also find that the reverse regression coefficient is positive, implying that under assumption sets (c) and (d), the true coefficient value is unbounded. We therefore turn to the second and third panels of Table 4, which present our thresholds. Here, we find first that, as in the previous application, the threshold values increase as we move from assumption sets (d) to (a). Of more interest are the high thresholds for the liquidity coefficient, which imply that the measurement quality of the proxy for true q must be very high in order to infer a negative coefficient value.

[6] See the latter for a complete description of the data and variables.

[7] Rajan and Zingales also deflate their variables by the market value of assets, with similar results.

Table 4
Leverage regressions: Estimates and proxy quality thresholds

		Q proxy	Tangibility	Log sales	EBIT	R^2
Direct regression						
		−0.070	0.268	0.026	−0.138	0.216
		(0.003)	(0.012)	(0.001)	(0.023)	
Reverse regression						
		−0.738	−0.326	0.021	2.182	0.247
		(0.027)	(0.039)	(0.005)	(0.133)	
Partial correlation thresholds						
	(a)	0.906	0.906	0.906	0.967	
		(0.006)	(0.009)	(0.009)	(0.002)	
	(b)	0.906	0.339	0.005	0.967	
		(0.006)	(0.021)	(0.001)	(0.002)	
	(c)	0.000	0.188	0.000	0.637	
			(0.014)		(0.273)	
	(d)	0.000	0.188	0.000	0.637	
			(0.014)		(0.273)	
Simple correlation thresholds						
	(a)	0.922	0.922	0.922	0.973	
		(0.005)	(0.008)	(0.008)	(0.002)	
	(b)	0.922	0.450	0.172	0.973	
		(0.005)	(0.021)	(0.014)	(0.002)	
	(c)	0.000	0.325	0.000	0.698	
			(0.013)		(0.231)	
	(d)	0.000	0.325	0.000	0.698	
			(0.013)		(0.231)	

Notes. Calculations are based on a sample of non-financial firms from the annual 2002 COMPUSTAT industrial files. The sample period is 1993 to 2001. Leverage is measured as the ratio of total long term debt less cash to the book value of assets. The Q Proxy is the market-to-book ratio; tangibility is the ratio of the book value of fixed assets to the book value of total assets; and EBIT is the ratio of earnings before interest and taxes to the book value of total assets. The direct and reverse regressions in the first panel are estimated by OLS. The thresholds presented are the lower bounds for either the partial or simple correlation between the true incentive to invest and its proxy. These bounds are necessary conditions for the coefficient on either cash flow or the interaction term to retain the sign presented in Table 2, under the assumption sets (a)–(d). In (a) the measurement error may be correlated with the regression disturbance term and one or more regressors (including the unobserved regressor itself); in (b) the measurement error may be correlated with the disturbance, but is uncorrelated with every regressor; in (c) the measurement error may be correlated with one or more regressors, but is uncorrelated with the disturbance; and in (d) the measurement error is uncorrelated with all other variables. Standard errors are in parentheses below all parameter estimates.

5. Conclusion

We provide a new econometric method for making inferences in the presence of a mismeasured regressor. We give a menu of different prior information sets that identify the *sign* of a coefficient on a mismeasured or perfectly measured regressor. The information sets are arguably weaker than those necessary to identify the precise values of the coefficients.

First, we apply this technique to regressions of investment on a proxy for true q and cash flow. We show that the partial and simple correlations between observed and true q must be unrealistically large before one can legitimately infer a positive cash-flow coefficient in the regression containing *true* q. Second, we apply this technique to a regression of leverage on a proxy for true q, lagged liquidity, and other controls, finding a similar result with regard to the negative OLS coefficient on lagged liquidity. Finally, the methodology itself is of broader interest: the use of proxies, including Tobin's q, is widespread in empirical corporate finance.

Acknowledgments

We thank Evangelos Falaris and Drew Roper for helpful comments and Peng Sun for research assistance.

Appendix A

A.1. Assumption set (a)

Proposition 1. *Suppose Assumptions 1–3 hold. If* $\rho^2 > 1 - r^2_{yx\cdot z}$ *or* $\tau^2 > R^2_{x\cdot z} + (1-R^2_{x\cdot z})(1 - r^2_{yx\cdot z})$, *then* $\beta > 0$.

Proof. to make use of information on proxy quality, first use (18) to write $\rho = \delta\sqrt{\operatorname{var}(\chi_i - z_i\mu)/\operatorname{var}(x_i - z_i m)}$. Using this expression to eliminate $\operatorname{var}(\chi_i - z_i\mu)$ from (15)–(17) yields

$$\operatorname{var}(y_i - z_i d) = \rho^2 \operatorname{var}(x_i - z_i m)\left(\frac{\beta}{\delta}\right)^2 + \operatorname{var}(u_i), \tag{22}$$

$$\operatorname{cov}(y_i - z_i d, x_i - z_i m) = \rho^2 \operatorname{var}(x_i - z_i m)\left(\frac{\beta}{\delta}\right) + \operatorname{cov}(u_i, e_i), \tag{23}$$

$$\operatorname{var}(x_i - z_i m) = \rho^2 \operatorname{var}(x_i - z_i m) + \operatorname{var}(e_i). \tag{24}$$

Equation (24) translates the requirement $\operatorname{var}(e_i) > 0$ into the restriction $\rho^2 < 1$. Substituting (22)–(24) into the requirement $\operatorname{var}(u_i)\operatorname{var}(e_i) - \operatorname{cov}^2(u_i, e_i) > 0$, using the definition $b = \operatorname{cov}(y_i - z_i d, x_i - z_i m)/\operatorname{var}(x_i - z_i m)$, and rearranging yields

$$b - s\sqrt{\frac{1-\rho^2}{\rho^2}} < \frac{\beta}{\delta} < b + s\sqrt{\frac{1-\rho^2}{\rho^2}}. \tag{25}$$

Since $\delta > 0$, inference about the sign of β simplifies to inference about the sign of β/δ, the quantity constrained to the interval (25). A lower bound on ρ^2 answers the question "what is the smallest number c such that $\rho^2 > c$ implies β/δ and b have the same sign?" The lower endpoint of (25) is increasing in ρ^2, while the upper endpoint is decreasing in ρ^2. Therefore, since $b > 0$, c equals that value of ρ^2 that sets the lower endpoint of (25) equal to zero. The solution is $c = 1 - r^2_{xy\cdot z}$. Invoking Corollary 1 completes the proof. □

Proposition 2. *Suppose Assumptions 1 and 3 hold. If* $\gamma_j > 0$, *and* $\rho^2 > \max\{c_I, c_A\}$ *or* $\tau^2 > R^2_{x\cdot z} + (1-R^2_{x\cdot z})\max\{c_I, c_A\}$, *then* $\alpha_j > 0$.

Proof. First we note the following lemma, which is proved at the end of this section.

Lemma 2. $\sup_{\gamma_j>0} c(\gamma_j) = \max\{c_I, c_A\}$, *where* $c(\gamma_j) = \frac{(m_j-\gamma_j)^2 s^2}{(m_j-\gamma_j)^2 s^2 + (a_j+\gamma_j b)^2}$, $c_I \equiv (1 + (a_j/m_j s)^2)^{-1}$, *and* $c_A \equiv (1 + (b/s)^2)^{-1}$.

Next, solving (12) for α gives, element-wise,

$$\alpha_j = d_j - (m_j - \gamma_j)\frac{\beta}{\delta}. \tag{26}$$

(26) implies we can multiply each term of (25) by $-(m_j - \gamma_j)$ and add d_j to obtain

$$\begin{aligned} & d_j - (m_j - \gamma_j)b - |m_j - \gamma_j| s\sqrt{\frac{1-\rho^2}{\rho^2}} \\ & \quad < \alpha_j < d_j - (m_j - \gamma_j)b + |m_j - \gamma_j| s\sqrt{\frac{1-\rho^2}{\rho^2}}. \end{aligned} \tag{27}$$

We simplify (27) by noting that the standard partialling result $a = d - mb$ implies

$$a_j \equiv d_j - m_j b, \tag{28}$$

and therefore

$$d_j - (m_j - \gamma_j)b = a_j + \gamma_j b. \tag{29}$$

Inequality (27) becomes

$$a_j + \gamma_j b - |m_j - \gamma_j| s\sqrt{\frac{1-\rho^2}{\rho^2}} < \alpha_j < a_j + \gamma_j b + |m_j - \gamma_j| s\sqrt{\frac{1-\rho^2}{\rho^2}}. \tag{30}$$

Suppose that γ_j is fixed and not equal to m_j, and that $a_j + \gamma_j b > 0$. Then there exists a number $c(\gamma_j)$ such that $\rho^2 = c(\gamma_j)$ sets the lower endpoint of (30) equal to zero and $\rho^2 > c(\gamma_j)$ implies $\alpha_j > 0$. Straightforward algebra shows that this number is

$$c(\gamma_j) = \frac{(m_j - \gamma_j)^2 s^2}{(m_j - \gamma_j)^2 s^2 + (a_j + \gamma_j b)^2}. \tag{31}$$

Suppose that it is known only that $\gamma_j > 0$. Then $\alpha_j > 0$ if ρ^2 exceeds the supremum of $c(\gamma_j)$ over positive γ_j. Then, Lemma 1 together with Corollary 1 imply Proposition 2. □

A.2. Assumption set (b)

Proposition 3. *Suppose Assumptions 1, 3, and 4 hold. If* $\rho^2 > m_j^2 s^2 (a_j^2 + m_j^2 s^2)^{-1}$ *or* $\tau^2 > R_{x\cdot z}^2 + (1 - R_{x\cdot z}^2) m_j^2 s^2 (a_j^2 + m_j^2 s^2)^{-1}$, *then* $\alpha_j > 0$.

Proof. By Eq. (5), Assumption 4 implies $\gamma = 0$ and $\delta = 1$. An immediate consequence is that Assumption 4 can replace Assumption 2 in the sufficient conditions of Proposition 1 without changing its conclusion. Inference about β is otherwise unchanged. To determine the consequences for inference about α, set $\gamma_j = 0$ in (31) to obtain $c(0) = m_j^2 s^2 (a_j^2 + m_j^2 s^2)^{-1}$ as the smallest number c such that $\rho^2 > c$ implies $\alpha_j > 0$. Corollary 1 determines a corresponding threshold for the simple correlation. □

A.3. Assumption set (c)

Proposition 4. *If Assumptions 1–3, and 5 hold, then* $\beta > 0$.

Proof. Assumption 5, (4), and Assumption 1(iii) imply $E(u_i e_i) = 0$. The requirement that $\mathrm{var}(u_i, e_i)$ be positive definite implies that both $\mathrm{var}(e_i)$ and $\mathrm{var}(u_i)$ are positive. As already noted, (24) translates $\mathrm{var}(e_i) > 0$ into the restriction $\rho^2 < 1$. To make use of $\mathrm{var}(u_i) > 0$, first set $\mathrm{cov}(u_i, e_i) = 0$ in (23), and rearrange the result as

$$\frac{\beta}{\delta} = \frac{\mathrm{cov}(y_i - z_i d, x_i - z_i m)}{\mathrm{var}(x_i - z_i m)\rho^2}. \tag{32}$$

Next, substitute (32) into (22) to obtain $\mathrm{var}(u_i) = \mathrm{var}(y_i - z_i d) - \mathrm{cov}^2(y_i - z_i d, x_i - z_i m)/[\mathrm{var}(x_i - z_i m)\rho^2]$. Restricting the right-hand side of this equality to be strictly positive defines an inequality that is easily manipulated to give $\rho^2 > r^2_{xy\cdot z}$. Thus,

$$r^2_{xy\cdot z} < \rho^2 < 1. \tag{33}$$

Equation (32) maps this interval for ρ^2 into the following restrictions on β/δ: $b < \beta/\delta < b/r^2_{xy\cdot z}$. This immediately implies the result. □

Proposition 5. *Suppose Assumptions 1, 3, and 5 hold, and* $a_j^* > 0$. *Then* $\gamma_j \geqslant 0$ *implies* $\alpha_j > 0$.

Proof. Substituting (32) into (26) gives

$$\alpha_j = d_j - (m_j - \gamma_j)\frac{b}{\rho^2}. \tag{34}$$

The result follows because this expression maps (33) into an interval containing α_j, with one endpoint equal to (29) and the other endpoint equal to

$$d_j - (m_j - \gamma_j)\frac{b}{r^2_{xy\cdot z}} = a_j^* + \frac{\gamma_j b}{r^2_{xy\cdot z}}. \qquad \square \tag{35}$$

Proposition 6. *Suppose Assumptions 1, 3, and 5 hold, and* $a_j^* < 0$. *If* $\gamma_j \geqslant 0$ *and either* $\rho^2 > m_j b/d_j$ *or* $\tau^2 > R^2_{x\cdot z} + (1 - R^2_{x\cdot z}) m_j b/d_j$, *then* $\alpha_j > 0$.

Proof. Write (34) as

$$\alpha_j = \left(d_j - m_j b/\rho^2\right) + \gamma_j b/\rho^2. \tag{36}$$

Refer to (28) to see that the bracketed term on right side of (36) will be positive if ρ^2 is sufficiently close to unity. The smallest number c such that $\rho^2 > c$ implies $d_j - m_j b/\rho^2 >$

0 is the value of ρ^2 that sets $d_j - m_j b/\rho^2 = 0$, which is $c = m_j b/d_j$. It follows that if $\gamma_j > 0$ and $\rho^2 > m_j b/d_j$ then $\alpha_j > 0$. Corollary 1 extends this result to the simple correlation. □

A.4. Assumption set (d)

A formal statement about the sign of β is given by replacing Assumption 2 with Assumption 4 in the sufficient conditions for Proposition 4. The results for α are given by adding Assumption 4 and deleting $\gamma_j \geqslant 0$ from the sufficient conditions of Propositions 5 and 6.

A.5. Proof of Lemma 1

Let $R^2_{x\cdot\chi z} \equiv 1 - E(e_i^2)/E(x_i^2)$, $R^2_{x\cdot z} \equiv 1 - E((x_i - z_i m)^2)/E(x_i^2)$, and $R^2_* \equiv 1 - E(e_i^2)/(E(x_i - \chi_i\phi)^2)$, where $\phi \equiv [E(\chi_i'\chi_i)]^{-1}E(\chi_i' x_i)$ is the coefficient from the projection of x_i on χ_i. Because the population R^2 of a simple linear regression equals the square of the simple correlation between the dependent and independent variables we have $\tau^2 \equiv 1 - E((x_i - \chi_i\phi)^2)/E(x_i^2)$ and, by (14), $\rho^2 = 1 - E(e_i^2)/E((x_i - z_i m)^2)$. Note that

$$1 - R^2_{x\cdot\chi z} = \left(1-\rho^2\right)\left(1 - R^2_{x\cdot z}\right) \tag{37}$$

$$= \left(1 - R^2_*\right)\left(1-\tau^2\right). \tag{38}$$

Equating the right-hand sides of (37) and (38) and then solving for ρ^2 yields

$$\rho^2 = \frac{1 - R^2_{x\cdot z} - (1 - R^2_*)(1-\tau^2)}{1 - R^2_{x\cdot z}}. \tag{39}$$

Inequality (20) is established by noting that (39) is increasing in R^2_* and that $R^2_* \geqslant 0$.

Next note from (3)–(6) that $E(\varepsilon_i \chi_i) = 0$ and $E(\varepsilon_i z_i) = 0$ imply $x_i = \chi_i + e_i$, which in turn implies $\phi = 1$. It follows that $x_i - \chi_i\phi = e_i$ and therefore $R^2_* = 0$, implying that the right-hand side of (39) equals $(\tau^2 - R^2_{x\cdot z})/(1 - R^2_{x\cdot z})$. □

A.6. Proof of Lemma 2

We begin by establishing some properties of $c(\gamma_j)$.

Property (a). If $m_j = -a_j/b$ then $c(\gamma_j) = (1 + (b/s)^2)^{-1}$ for all γ_j.

Property (b). If $m_j \neq -a_j/b$ then $c(\gamma_j)$ has a unique maximum equal to 1 at $\gamma_j = -a_j/b$, a unique minimum equal to zero at $\gamma_j = m_j$, and an asymptote equal to $(1 + (b/s)^2)^{-1}$ as $\gamma_j \to \pm\infty$.

Property (c). If $m_j > -a_j/b$ then $c(\gamma_j)$ is strictly increasing on $(-\infty, -a_j/b)$, strictly decreasing on $(-a_j/b, m_j)$, and strictly increasing on (m_j, ∞).

Property (d). If $m_j < -a_j/b$ then $c(\gamma_j)$ is strictly decreasing on $(-\infty, m_j)$, strictly increasing on $(m_j, -a_j/b)$, and strictly decreasing on $(-a_j/b, \infty)$.

To establish Property (a), rewrite (31) as

$$c(\gamma_j) = \frac{(m_j - \gamma_j)^2 s^2}{(m_j - \gamma_j)^2 s^2 + (a_j/b + \gamma_j)^2 b^2} \tag{40}$$

$$= \frac{1}{1 + \left(\frac{a_j/b+\gamma_j}{m_j-\gamma_j}\right)^2 \frac{b^2}{s^2}}, \tag{41}$$

and then note that it reduces to $(1 + b^2/s^2)^{-1}$ if $m_j = -a_j/b$. The remaining properties are established under the assumption $m_j \neq -a_j/b$, which, it should be noted, ensures that the denominator of (40) is positive, and hence $c(\gamma_j)$ is well defined, for all $\gamma_j \in (-\infty, \infty)$. The unique maximum of Property (b) is established by inspecting (41) to see that $c(\gamma_j) \leqslant 1$ for all $\gamma_j \in (-\infty, \infty)$, and that $c(\gamma_j) = 1$ if and only if $\gamma_j = -a_j/b$. The unique minimum follows by noting from (40) that $c(\gamma_j)$ cannot be negative, and that $c(\gamma_j) = 0$ if and only if $\gamma_j = m_j$. To derive the asymptote, use (28) to eliminate a_j from (41) and obtain

$$c(\gamma_j) = \frac{1}{1 + \left(\frac{d_j/b}{m_j-\gamma_j} - 1\right)^2 \frac{b^2}{s^2}}, \tag{42}$$

and then note that $(d_j/b)/(m_j - \gamma_j)$ tends to zero as $\gamma_j \to \pm\infty$. To establish Property (c), note that the derivative of (42) with respect to γ_j is:

$$\frac{\partial c}{\partial \gamma_j} = \frac{-2bd_j\left(\frac{d_j/b}{(m_j-\gamma_j)} - 1\right)}{\left(1 + \left(\frac{d_j/b}{(m_j-\gamma_j)} - 1\right)^2 \frac{b^2}{s^2}\right)^2 (m_j - \gamma_j)^2 s^2}. \tag{43}$$

Since the denominator of (43) is positive, the sign of the derivative is the same as the sign of the numerator. Using (28) to eliminate d_j, the numerator can be rewritten as

$$2b^2\left(\frac{a_j}{b} + m_j\right)\left(\frac{-a_j/b - \gamma_j}{m_j - \gamma_j}\right).$$

The restriction $m_j > -a_j/b$ implies $2b^2(\frac{a_j}{b} + m_j) > 0$, and also that $\left(\frac{-a_j/b-\gamma_j}{m_j-\gamma_j}\right)$ is positive if $\gamma_j \in (-\infty, -a_j/b)$, negative if $\gamma_j \in (-a_j/b, m_j)$, and positive if $\gamma_j \in (m_j, \infty)$. The derivative (43) is therefore positive if $\gamma_j \in (-\infty, -a_j/b)$, negative if $\gamma_j \in (-a_j/b, m_j)$, and positive if $\gamma_j \in (m_j, \infty)$. To establish Property (d), note that $m_j < -a_j/b$ implies $2b^2(\frac{a_j}{b} + m_j) < 0$, and that $\left(\frac{-a_j/b-\gamma_j}{m_j-\gamma_j}\right)$ is positive if $\gamma_j \in (-\infty, m_j)$, negative if $\gamma_j \in (m_j, -a_j/b)$, and positive if $\gamma_j \in (-a_j/b, \infty)$; therefore, (43) is negative if $\gamma_j \in (-\infty, m_j)$, positive if $\gamma_j \in (m_j, -a_j/b)$, and negative if $\gamma_j \in (-a_j/b, \infty)$.

Given these four properties, it is clear that the supremum of $c(\gamma_j)$ over the interval $[0, \infty)$ equals $\max\{c_I, c_A\}$, where $c_I \equiv c(0) = (1 + (a_j/m_j s)^2)^{-1}$ is the intercept of $c(\gamma_j)$, and $c_A = (1 + (b/s)^2)^{-1}$ is the asymptote. Figure 1 illustrates the possible shapes of $c(\gamma_j)$ under our assumptions $a_j > 0$ and $b > 0$. The first panel assumes the hypothesis of Property (d); the remaining panels assume that of Property (c). □

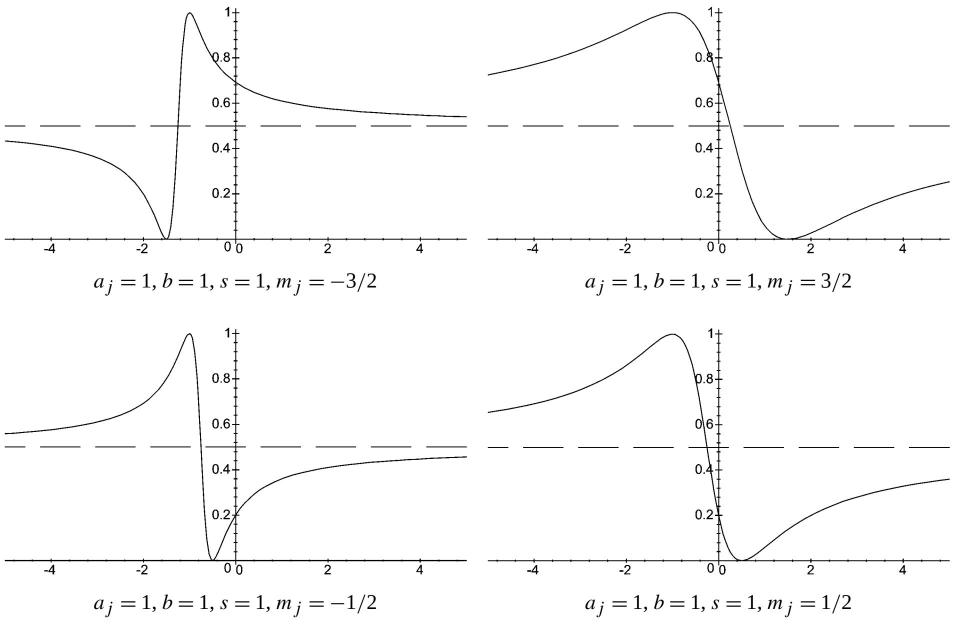

Fig. 1. This figure illustrates the possible shapes for 0(7) in Lemma 2. $c(\gamma)$ is the function that defines, under assumption set (a) the partial correlation threshold necessary for the coefficients on the perfectly measured regressors to retain their sign obtained via OLS. Under this assumption set the measurement error may be correlated with the regression disturbance term and one or more regressors (including the unobserved regressor itself) b is the coefficient from the least squares projection of the left hand side variable on the proxy and a vector of perfectly observed regressors. a_j is the jth element of vector of coefficients on the perfectly measured regressors. m_j is the jth element of the coefficient vector obtained from a least squares projection of the proxy on the perfectly measured regressors. s is a function, given in Section 2, of the error variances in these two projections and b.

References

Allayanis, Y., Mozudmar, A., 2004. The investment-cash flow sensitivity puzzle: Can negative cash flow observations explain it? Journal of Banking and Finance 28, 901–930.

Blanchard, O.J., Lopez-de-Silanes, F., Shleifer, A., 1994. What do firms do with cash windfalls? Journal of Financial Economics 36, 337–360.

Bollinger, C., 1996. Bounding mean regressions when a binary regressor is mismeasured. Journal of Econometrics 73, 387–399.

Erickson, T., 1993. Restricting regression slopes in the errors-in-variables model by bounding the error correlation. Econometrica 61, 959–969.

Erickson, T., Whited, T.M., 2000. Measurement error and the relationship between investment and q. Journal of Political Economy 108, 1027–1057.

Erickson, T., Whited, T.M., 2002. Two-step GMM estimation of the errors-in-variables model using high-order moments. Econometric Theory 18, 776–799.

Fazzari, S., Hubbard, R.G., Petersen, B.C., 1988. Financing constraints and corporate investment. Brookings Papers on Economic Activity 1, 141–195.

Frisch, R., 1934. Statistical Confluence Analysis by Means of Complete Regression Systems. University Institute of Economics, Oslo, Norway.

Gini, C., 1921. Sull’interpolazione di una retta quando i valori della variabile indipendente sono affetti da errori accidentali. Metroeconomica 1, 63–82.

Hayashi, F., Inoue, T., 1991. The firm growth – q relationship with multiple capital goods: Theory and evidence from panel data on Japanese firms. Econometrica 59, 731–754.

Hennessy, C.A., Whited, T.M., 2005. Debt dynamics. Journal of Finance 60, 1129–1165.

Hubbard, R.G., 1998. Capital market imperfections and investment. Journal of Economic Literature 36, 193–238.

Kaplan, S., Zingales, L., 1997. Do financing constraints explain why investment is correlated with cash flow? Quarterly Journal of Economics 112, 169–216.

Klepper, S., Leamer, E.E., 1984. Consistent sets of estimates for regressions with errors in all variables. Econometrica 52, 163–183.

Klepper, S., 1988. Bounding the effects of measurement error in regressions involving dichotomous variables. Journal of Econometrics 37, 343–359.

Krasker, W.S., Pratt, J.W., 1986. Bounding the effects of proxy variables on regression coefficients. Econometrica 54, 641–655.

Kroch, E., 1988. Bounds on specification error arising from data problems. Journal of Econometrics 37, 171–192.

Leamer, E.E., 1987. Errors in variables in linear systems. Econometrica 55, 893–909.

Lamont, O., 1997. Cash flow and investment: Evidence from internal capital markets. Journal of Finance 52, 83–109.

Odean, T., 1999. Do investors trade too much? American Economic Review 89, 1279–1298.

Rajan, R.G., Zingales, L., 1995. What do we know about capital structure? Some evidence from international data. Journal of Finance 50, 1421–1460.

Rauh, J., Investment and financing constraints: Evidence from the funding of corporate pension plans, Journal of Finance, in press.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.

Whited, T.M., 1992. Debt, liquidity constraints, and corporate investment: Evidence from panel data. Journal of Finance 47, 1425–1460.