Investment Cash Flow Sensitivity: Fact or Fiction?*

Şenay Ağca
George Washington University

Abon Mozumdar
Virginia Tech

October 25, 2010

Abstract

We examine whether internal funds matter for investment when measurement errors in q are addressed. We reconcile the evidence presented in studies that suggest significance of cash flow to be a result of measurement errors in q with the literature that finds internal funds to be a determinant of investment. Through a detailed analysis of the studies that tackle measurement errors in q, we show that when data and econometric issues are addressed, cash flow cannot be dismissed as an artifact of measurement errors in q. We also find that an analyst forecast based q measure is not superior to a stock market based q measure. We suggest an alternative methodology and show that careful specification and choice of instruments are imperative for obtaining well-specified and powerful models. Our results have considerable implications for studies that examine the relation between investment and financing in panel data settings.

Keywords: Investment, internal funds, measurement error, financial econometrics.

JEL Classification: G14, G31, G32

Şenay Ağca, School of Business, George Washington University, 2201 G Street NW, Funger Hall 505, Washington, D.C. 20052, Tel: (202) 994-9209, Fax: (202) 994-5014, Email: sagca@gwu.edu; Abon Mozumdar, Pamplin College of Business, Virginia Tech, 7054 Haycock Road, Room 352, Falls Church, VA 22043, Tel.:703-538-8414, Fax: 703-538-8415, Email: abon@vt.edu. We are grateful to Jason Cummins, Kevin Hassett, and Toni Whited for making available their data and programs on their websites and through personal correspondence, N.R.Prabhala for extensive and detailed comments, and to Mike Cliff, Fabio Schiantarelli, Toni Whited, and seminar participants at Virginia Tech and 2008 ISB Summer Research Conference, Hyderabad, India for many comments. The previous version of this paper circulated as "Investment-Cash Flow Sensitivity: Myth or Reality?" Agca acknowledges research support from the School of Business, George Washington University and Mozumdar acknowledges summer research support from the Pamplin College of Business, Virginia Tech.

Investment-Cash Flow Sensitivity: Fact or Fiction?

Should internal cash matter for firm investments? Studies that are based on natural experiments where there is a clear exogenous shock to cash flow without a corresponding change in growth opportunities find that cash flow matters for investment (See Blanchard, Lopez-de-Silanes, and Shleifer (1994), Lamont (1997), and Rauh (2006)). On the other hand, Erickson and Whited (2000, 2002) (EW henceforth) and Cummins, Hassett, and Oliner (2006) (CHO henceforth) analyze the relation between investment and cash flow more directly without reference to any specific conditioning events, and find insignificant cash flow effects, after carefully accounting for the measurement error in q in panel data settings. The apparent contradiction between the two approaches demands an explanation. One possibility is that the disparity reflects behavioral differences in how managers respond to unanticipated windfalls compared to how they respond to other types of cash flow shocks in the panel data settings. Alternatively, the responses to unanticipated exogenous shocks may be mimicked in the (much) larger samples used in panel data settings.

In this paper, we find that when data and econometric issues are addressed, both lines of literature show that cash flow cannot be dismissed as a determinant of investment. We pin down the differences with careful attention to specification and estimation in the panel data settings used in investment-cash flow regressions. Our findings have implications for panel data estimations that are rapidly finding application in empirical corporate finance. We show that correct instruments and specifications are crucial to avoid implementation hazards in typical panel data sets in corporate finance.

One clear advantage that the natural experiments approach enjoys is that of simplicity. For example, Lamont (1997) examines capital expenditures of oil companies following the 1986 oil shock. As this exogenous shock reduces cash flows of oil companies without affecting their investment opportunities, it provides a natural setting for examining how capital expenditures respond to cash flow. Lamont finds that investment cutbacks are more severe in single-segment oil firms than in the oil divisions of multi-segment conglomerates in the face of negative oil price shocks. This finding lends itself to a single unambiguous interpretation – that investments respond to the availability of internal funds. Besides natural experiments, following Fazzari, Hubbard and Petersen (1988), an enormous literature uses direct estimation methods to consider the impact of internal funds on investments. Among recent studies, Almeida and Campello (2007), Brown, Fazzari and Petersen (2009), Fee, Hadlock and Pierce (2009) and Polk and Sapienza (2009) follow this direction and find that internal funds matter for investment. Stein (2003) reviews earlier work in the literature.

To interpret the regressions of investment on cash flow, it is important to control for the growth opportunities of firms. The standard approach is to use an observable measure of Tobin's q as a proxy for marginal q, which should be the only determinant of investment according to q theory. Poterba (1988) shows that measurement errors in q can lead to spurious correlations between investment and cash flow. EW and CHO tackle this issue by implementing alternative methodologies that deal with the measurement errors in q and show insignificant cash flow effects. Their findings suggest that cash flow does not matter for investment when measurement error in q is addressed. Both papers use rather different approaches to handle the measurement error in q.

EW propose a GMM estimator based on higher order moments of the regression variables. Applying these estimators to a balanced panel of manufacturing firms over the 1992-1995 period, they find q to have a much greater impact relative to OLS estimates, and cash flow effects to be insignificant. Recently, Almeida, Campello and Galvao (2010) scrutinize the EW methodology and find it to be prone to provide inefficient and biased estimates. In response, Erickson and Whited (2010) indicate that with an appropriate proxy for q and with proper starting values in estimations, EW methodology provides reasonable finite sample performance. Both papers report significant cash flow effects. Interestingly, the findings in Erickson and Whited (2010) contradict those in Erickson and Whited (2000), who report insignificant investment-cash flow sensitivity on their dataset. Because the datasets used in the two studies are different, it remains unclear whether the disparity is driven by sample differences, perhaps due to changing financial policies of firms in the 21st century, or whether investment is indeed sensitive to cash flow even in the earlier period. Our study shows that cash flow cannot be dismissed as an artifact of measurement errors in q even in the earlier period.

CHO propose an innovative approach that uses a measure of q from analyst forecasts. The basic argument underlying CHO's analysis is that the measurement error in stock market-based measures of q is large and persistent enough to significantly distort the estimation. They posit that earnings levels and growth rates predicted by professional analysts are less noisy, and therefore a measure of q based on earnings forecasts is a better proxy for firm growth opportunities. While it is possible for analyst forecasts to be less noisy, it seems curious that such estimates of earnings growth could be systematically superior to those embodied in stock market prices, which incorporate analyst opinions and a lot more besides. Applying analyst forecast based q measure in a dynamic panel setting, CHO find that cash flow ceases to be a significant determinant of investment, and that analyst-forecast based q measure has better explanatory power for investments than the stock market-based q. We provide a detailed analysis of the CHO study and show that insignificant cash flow effects turn significant when data and methodological issues are addressed. We also propose an alternative specification to address the measurement errors in q that is based on best practices on the lines recommended by the econometrics literature on panel data methods.

With regard to CHO, we first scrutinize the data, which indicates that instead of using contemporaneous cash flow (CF_t) as an explanatory factor for investment (I_t), they use one-period lead values of cash flow (CF_{t+1}). We then look at the instrument set. The choice of instruments used in the GMM estimation (third and fourth lags of normalized investments and cash flow) is an unnecessarily restricted one. The estimator efficiency can be easily increased by including longer lags in the instrument set. We also address the observation of analyst forecasts within the investment interval and the use of discounted earnings rather than discounted cash flows for share valuations. We find that when these issues are taken into consideration, the main CHO results cease to hold, both in their original sample and in our reconstructed one. Specifically, we find that an analyst expectations-based measure of q is not superior to a stock price-based one in explaining investments, and that cash flow is a significant determinant of investment.

Regarding EW, we consider a data discrepancy in their original dataset, consider extended periods and use extended observations. Overall, our findings suggest that cash flow cannot be dismissed as a determinant of investment even after addressing measurement errors in q. An important empirical concern is the issue of *negative* q. We find that the proxy used for q in EW produces negative q values in a number of cases. This is clearly not very meaningful because investment opportunities are bounded below by zero. Yet another issue highlighted by our analysis is the need to pay more thorough attention to fit-related diagnostics in panel data settings. We find that the specification tests fail in a majority of cases, an issue that is largely sidestepped in EW. These issues cast doubt on the practical use of the EW methodology and the related result on disregarding cash flow effects as an artifact of measurement errors in q. Thus, our findings reaffirm the conclusions of Almeida, Campello and Galvao (2010).

We also offer novel remedies to the issues of mismeasured q. Our main point is that the existence of two mismeasured proxies for q in fact offer an alternative route to address measurement error problem, by using lags of one as instruments for the other. If the measurement errors in the two proxies are uncorrelated, or even if they are correlated but the cross-correlation decays faster over time than the autocorrelations, then such an estimator will be less subject to measurement error than one using the same proxy as both the regressor and the instrument. We use three estimators suggested in the literature: the instrumental variables estimator of Anderson and Hsiao (1981), and two versions of the dynamic panel GMM estimator developed by Arellano and Bond (1991, 1998) and employed by Blundell, Bond, Devereux, and Schiantarelli (1992) and by CHO. We find that the Anderson and Hsiao (1981) estimator has relatively little power. However, we obtain well-specified and powerful models using the GMM approach. The GMM evidence confirms our earlier findings: internal funds matter for investment, and the stock price-based measure of q is superior to the one based on analyst forecasts. Interestingly, we find model specification tests to be satisfied when lagged q is included in the set of instruments, in contrast to CHO. We also find that estimated coefficients are similar to each other whether the same mismeasured proxy for q is used as both the regressor and the instrument, or separate proxies are used.

The rest of the paper is organized as follows. Section 1 provides a brief summary of the basic q model of investments and the measurement error problem in estimating it. Section 2 describes the data. Section 3 revisits the CHO evidence using both their original sample and an independently reconstructed one. Similarly, Section 4 revisits the EW evidence using original and reconstructed data. Section 5 develops an alternative methodology that utilizes different measures of q as instruments. Section 6 concludes.

1. The q Model of Investments and Measurement Error

Consider the simplified version of the standard q investment model presented by Erickson and Whited (2000). The firm chooses an investment policy I_t to maximize the expected discounted value of cash flow

$$V_t = \max_{\{I_{t+s}\}_{s=0}^{\infty}} E\left\{\sum_{s=0}^{\infty} \rho^s \left[\pi_{t+s}(K_{t+s-1}) - C_{t+s}(K_{t+s-1}, I_{t+s})\right]\right\} \quad (1)$$

subject to the intertemporal constraint for capital

$$K_t = (1-\delta)K_{t-1} + I_t \quad (2)$$

where I_t denotes gross investment; K_{t-1}: beginning-of-period capital stock; π_t: profits (gross of investments); C_t: cost of investment; ρ: the single period discount factor; and δ: the single period depreciation rate. The price of capital is chosen to be the numeraire. The cost of investment $C_t = I_t + \psi_t(K_{t-1},I_t)$ includes the price of new capital assets (I_t), as well as capital adjustment costs, $\psi_t(K_{t-1},I_t)$. Substituting (2) repeatedly into (1) and differentiating with respect to I_t yields the first order condition

$$1 + \psi_I(K_{t-1}, I_t^*) = E\left\{\sum_{s=1}^{\infty} \rho^s (1-\delta)^{s-1} \left[\pi_K(K_{t+s-1}) - \psi_K(K_{t+s-1}, I_{t+s})\right]\right\} = q_t$$

(3)

where q_t is the marginal cost of capital, and I_t^* solves (3). Assuming a linearly homogeneous quadratic adjustment cost function $\psi_t(K_{t-1},I_t) = c_0 K_{t-1}(c_1 + I_t / K_{t-1})^2$, substituting into (3), and casting in a regression framework yields

$$\frac{I_t}{K_{t-1}} = a_1 + \beta_1 q_t + \varepsilon_t, \qquad \text{with} \quad \varepsilon_t \sim i.i.d.(0, \sigma_\varepsilon^2)$$

(4)

which implies that investments are determined solely by the shadow price of capital, or marginal q. In particular, considerations of the availability of internal funds should play no role in the process. On the other hand, significant deviations from the perfect market paradigm would result in such considerations playing an important role, i.e., a significant coefficient **B** in the regression

$$\frac{I_t}{K_{t-1}} = a_1 + \beta_1 q_t + \mathbf{z}_t\mathbf{B} + \varepsilon_t \tag{5}$$

where **z** represents some measure(s) of internal funds. In empirical work, marginal q (q_t) is typically approximated by Tobin's average q (Q_t), while the most commonly used measure of internal funds is cash flow. Given the severity of the assumptions required, the strict structural interpretation of the model may not hold exactly, but equation (5) still has a natural interpretation: are investments determined exclusively by the attractiveness of investment opportunities as measured by marginal q (or Tobin's q), or does internal funds matter for investment?

Unfortunately, estimation of (5) is hindered by the problem of measurement error in q. CHO and EW list the several layers of approximation that lie between the marginal q of theory and the various versions of Tobin's q used in practice. Let Q_t be the mismeasured empirical proxy of the true marginal q_t, i.e.,

$$Q_t = a_0 + q_t + u_t, \qquad \text{with} \qquad u_t \sim i.i.d.(0, \sigma_u^2) \tag{6}$$

Substituting (6) into (5), we get

$$\begin{aligned} \frac{I_t}{K_{t-1}} &= (a_1 - \beta_1 a_0) + \beta_1 Q_t + \mathbf{z}_t\mathbf{B} + (\varepsilon_t - \beta_1 u_t) \\ &= a + \beta_1 Q_t + \mathbf{z}_t\mathbf{B} + e_t \end{aligned} \tag{7}$$

Clearly, $Cov(Q_t, e_t) = -\beta_1\sigma_u^2 \neq 0$, so the usual OLS condition of independence between errors and regressors is violated and the OLS estimates of (β_1,**B**) are inconsistent. In particular, the probability limit of the estimate of β_1 equals $\beta_1/(1+\sigma_u^2\Sigma_{11})$ where $\mathbf{\Sigma}$ = plim $\mathbf{X'X}/n$, and $\mathbf{X}=(q, \mathbf{z})$, i.e., the estimate is biased toward 0. For the estimate of the i-th element of **B**, the probability limit is $\mathbf{B}_i - (\beta_1\sigma_u^2\Sigma_{i+1,1}) / (1+\sigma_u^2\Sigma_{11})$, which may be either greater or smaller than $\mathbf{B}_i$. (See Greene (2003), p. 83-86 for details.) The CHO approach to overcoming this problem is to use a measure of Q for which the measurement error u is small, while that of EW uses the information in higher order moments to explicitly isolate its impact.

2. Data

Our sample is obtained from the universe of U.S. public firms over the 1982-2003 period. All accounting data and year-end stock prices are collected from COMPUSTAT. Details on constructing regression variables from COMPUSTAT data are provided in Appendix A1. Data on analyst expectations about future earnings and growth rates are obtained from I/B/E/S. Stock betas based on the Scholes and Williams (1977) approach are obtained from CRSP. Since I/B/E/S data are available for only a small subset of the CRSP-COMPUSTAT coverage, the sample size reduces for tests that require the use of I/B/E/S data. The basic sample is modified in various ways for implementing the different tests in the analysis. Details of samples used in the different tests are reported in the tables and in their discussions in the text.

3. Analyst Forecast Based Measure of q

The basic argument underlying CHO's analysis is that the measurement error u in traditional equity market-based measures of Tobin's q (Q_E) is large and persistent enough to significantly distort investment-cash flow sensitivity estimation. They posit that earnings levels and growth rates predicted by professional analysts are less noisy, and therefore a measure of q based on earnings forecasts is a better proxy for firm growth opportunities. They construct such a measure $\hat{Q}$ from analyst forecasts as reported by I/B/E/S and find that it has better explanatory power for investments than the stock market-based Q_E. Importantly, when $\hat{Q}$ is used to control for investment opportunities, cash flow ceases to be a significant determinant of investment.

Examining CHO methodology and implementation, we find that CHO evidence is affected by data discrepancy and restrictive choice of instruments. When these issues as well as some other minor problems are taken into account, the evidence is contrary to those reported in CHO: Investment is sensitive to cash flow and analyst forecast based q measure is not superior to stock market based q measure. We explore these issues in the next subsections.

3.1 A Data Discrepancy

CHO's analysis and conclusions are affected by an apparent discrepancy in their dataset, which shows up while attempting to replicate their Figure 2. The three panels in their figure plot annual average percent changes in two variables at a time: investments and $\hat{Q}$ in the top panel, investments and Q_E in the middle one, and investments and cash flow in the bottom. They also report the adjusted R^2 values from regressing the investments series on the $\hat{Q}$, Q_E, and cash flow series as 0.71, 0.03, and 0.40, respectively.

It is easy to replicate the top two panels of their figure. For the bottom panel, however, plotting cash flows based on the data reported in their Gauss dataset fulld.dat produces a series that is time-shifted back by one period relative to their plot. Naturally, when we time-shift the cash-flow data forward by one period, we obtain their plot. Our plots based on the original and the time-shifted CHO data are shown in Figure 1. We also find that the adjusted R^2 from regressing the investment series on the cash flow series is -0.068 when we use the (unshifted) CHO data, but is 0.4030, as in CHO, when we use the shifted data.

The problem is more than one of mere reporting, since almost all of CHO's tables can be replicated only by using the unshifted data. (The sole exception is the set of results for the unrated sample of their Table 3, which can be replicated only by using the shifted data.) We are thus forced to conclude that either their figure or most of their tables is/are incorrect. Since CHO do not include firm identifiers in their dataset, we are unable to directly check which one of the cash flow series – unshifted or shifted – is the correct one. However, the high (low) adjusted R^2 from regressing investment on the shifted (unshifted) cash flow series, and the high correlation between investment and cash flow found by earlier studies, together indicate that the shifted series, and therefore their reported figure, are correct. This implies that what has been used by CHO as contemporaneous cash flow CF_t in most of their regression analysis is really CF_{t+1}. While this is only an indirect inference at this point, it is supported by results from our independently constructed dataset, as reported in Section 3.4.

3.2 Choice of Instruments for GMM Estimation

The dynamic panel data GMM estimator used by CHO is based on earlier work by Arellano and Bond (1991, 1998) and Blundell, Bond, Devereux, and Schiantarelli (1992) (hereafter BBDS). The error term e in (7) can be decomposed into a firm-specific $FIRM_i$, a time-specific $YEAR_t$, and an idiosyncratic disturbance w_{it} to yield

$$\frac{I_{it}}{K_{it-1}} = a + \beta_1 q_{it} + \beta_2 \frac{CF_{it}}{K_{it-1}} + FIRM_i + YEAR_t + w_{it} \tag{8}$$

where the additional subscript i indicates the firm, and scaled cash flow (CF/K) is used as the internal liquidity measure **z**. First-differencing (8) eliminates the firm-specific component $FIRM_i$, and replaces the original set of time-specific $YEAR_t$'s with their first-differences

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = \beta_1 \Delta q_{it} + \beta_2 \Delta\left(\frac{CF_{it}}{K_{it-1}}\right) + \Delta YEAR_t + w^*_{it} \tag{9}$$

While endogenous variables lagged by two periods or more are legitimate instruments if the w_{it} are serially uncorrelated, CHO use as instruments normalized investments and cash flows with three or more lags, since they find that specification tests generally reject models with *t-2* dated instruments. Their final choice of instruments consists of *t-3* and *t-4* dated investments and cash flows.

This is an unnecessarily restricted choice of instruments. In similar panel data settings, it is customary to increase estimator efficiency by exploiting large numbers of moment conditions, resulting in models that are vastly over-identified. In particular, this entails using the maximum possible lags of eligible endogenous variables as instruments. (See, e.g., Holtz-Eakin, Newey, and Rosen (1988), Arellano and Bond (1991), BBDS, and Greene (2003).) Earlier studies such as BBDS were forced to restrict the numbers of remote instruments due to computing limitations, but currently available computing power makes such constraints irrelevant for panels with moderate time horizons, as is the case here. At the same time, however, there is the risk of introducing an over-fitting bias by using too many instruments (Arellano and Bond (1998), Wooldridge (2002, p.305), Alvarez and Arellano (2003)). It is therefore necessary to ensure that efficiency gains are not pursued at the cost of estimator unbiasedness.

The number of available observations is also reduced by CHO's requiring a complete set of instruments for all observations. Again, this is not necessary. The earlier observations for each firm which are excluded from CHO's sample can in fact legitimately be included, albeit with fewer moment restrictions than the other observations.

3.3 Tests of Common Factor Restrictions

BBDS and CHO find that the assumption of the errors w_{it} in (8) being serially uncorrelated is generally rejected by the m_2 test of second order serial correlation (Arellano and Bond (1991)). They find that the data are more consistent with AR(1) disturbances which implies a dynamic model with non-linear common factor restrictions on the estimated coefficients. Consider model (8), but with AR(1) w_{it}

$$w_{it} = \rho\, w_{it-1} + \zeta_{it} \qquad (10)$$

Substituting for w_{it-1} from (8), collecting terms, and taking first differences, we have

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = \rho\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + \beta_1\Delta q_{it} - \rho\beta_1\Delta q_{it-1} + \beta_2\Delta\left(\frac{CF_{it}}{K_{it-1}}\right) - \rho\beta_2\Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it} \qquad (11)$$

The coefficients from the corresponding unrestricted model

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = b_0\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + b_1\Delta q_{it} + b_2\Delta q_{it-1} + b_3\Delta\left(\frac{CF_{it}}{K_{it-1}}\right) + b_4\Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it} \qquad (12)$$

can be used to generate the restricted coefficients as in BBDS, and tested for the common factor restrictions $b_2 = -b_0 b_1$ and $b_4 = -b_0 b_3$ by means of a χ^2 test. BBDS approach consists of minimizing the normed distance between a non-linear transformation of the unrestricted estimates and a linear transformation of the restricted ones, the minimized function being asymptotically $\chi^2(2)$ distributed under the null.

3.4 The CHO Evidence Revisited

We begin by examining the original CHO dataset but with the apparent data discrepancy mentioned in Section 3.1 corrected by time-shifting the cash flow data forward by one year. Table 1 Panel A reports results for the traditional OLS estimation in first-differences, as specified in (9). Cash flow is found to be positive and significant in explaining investments, when growth opportunities are controlled for by using any of the three proxies – Q_E, $\hat{Q}$, or *LTG* (analysts's long term growth forecast from I/B/E/S). The bottom row of Panel A reports results for a combined model in which both Q_E and $\hat{Q}$ are used as explanatory variables; both are found to be positive and significant.

Panel B reports GMM results from estimating model (9) that are directly comparable to Table 2 in CHO.[1] In sharp contrast to CHO, the cash flow coefficient is positive and significant in all cases. In the bottom row, Q_E retains its significance when both it and $\hat{Q}$ are used as controls for growth opportunities. The evidence thus indicates that CHO's fundamental results of the insignificance of cash flow and the inferiority of Q_E relative to

---

[1] As in CHO, we follow Arellano and Bond's (1991) simulation-based recommendation to use first-step GMM estimates of parameters and the m_2 statistic, and second-step estimates of the *J*-statistic. Standard errors are robust to heteroscedasticity and account for first-order serial correlation in the (first-differenced) errors (see Arellano and Bond (1998) p. 7).

analyst expectations-based measures are possibly driven by the apparent use of CF_{t+1} in place of CF_t. In contrast to CHO, the Sargan (J) test of over-identifying conditions (Sargan (1958), Hansen (1982)) as well as Arellano and Bond's (1991) m_2 test of second order serial correlation reject consistently, indicating that AR(1) disturbances and the extended dynamic structure of (11) or (12) may be more appropriate.

Panels C and D extend the sample and the set of instruments as described in Section 3.2. Panel C includes observations with third and fourth lags of investment and cash flow as instruments when both lags are available, and third lags only when fourth lags are not. Panel D utilizes more moment conditions by using all available lags of length greater than two of investments and cash flow as instruments. Results are very similar to those of Panel B – cash flow is positive and significant in all cases, while the specification tests continue to reject. Standard errors in Panel D are significantly smaller indicating efficiency gains from using the much larger instrument set. Parameter estimates are somewhat smaller relative to Panels B and C, suggesting the possibility of some over-fitting bias, but it is not large enough to change inferences qualitatively. In any case, the consistent rejection of the specification tests clearly shows that these estimates cannot be considered reliable.

Table 2 reports results from estimating the restricted coefficients of (11) from the unrestricted model (12) using the approach of BBDS. Panel A is directly comparable to Table 3 of CHO. The results here provide some support to CHO's findings: cash flow is positive but insignificant at conventional levels when $\hat{Q}$ is used as a control, while $\hat{Q}$ is positive and significant. Q_E is insignificant in the combined model. Note, however, that the estimated cash flow coefficients are two orders of magnitude greater than those reported by CHO, indicating that the data discrepancy continues to play an important role. The m_2 test generally fails to reject indicating that the AR(1) error structure is appropriate. The Sargan (J) test fails to reject the $\hat{Q}$-based models, although in the combined model, the p-value is only marginally greater than conventional rejection levels. The BBDS test of common factor restrictions fails to reject. Note that in the combined model, there are three common factor restrictions (one each for Q_E, $\hat{Q}$, and cash flow) and the test statistics are distributed $\chi^2(3)$.

Panels B and C, however, suggest that when more information in the form of additional observations and/or moment conditions are brought to bear on the estimation, the evidence contradicts CHO's findings. When observations with less than the full set of instruments are included (Panel B), cash flow becomes weakly significant, although $\hat{Q}$ continues to dominate Q_E in the combined model. When all possible instruments are used (Panel C), cash flow becomes strongly significant and Q_E retains its significance in the combined model. While the m_2 test continues to fail to reject, the J-test as well as the common factor restrictions test reject comprehensively, indicating that the data are incapable of satisfying the large number of conditions imposed at this level.

Panel D presents results using the full set of available instruments as in Panel C, but with the unshifted CHO data. In contrast to Panel C but consistent with CHO, the estimated investment-cash flow sensitivity is small and insignificant, which suggests that CHO's finding of an insignificant cash flow effect is not only due to the low power of the limited instrument set that they use, but also to the data discrepancy. Both Q_E and $\hat{Q}$ are significant in the combined model, which further contradicts CHO's conclusion of $\hat{Q}$ being a superior measure of growth opportunities than Q_E. As in Panel C, however, all specification tests except m_2 reject comprehensively, which casts some doubt on the validity of the model with these data.

Recall from Section 3.1 that our inference about a time-shift in CHO's cash flow series is only an indirect one, based on the apparent discrepancy between their diagram and tables. Therefore, we next repeat the above analysis using an independently constructed dataset. In order to maintain comparability with the original CHO data, we adopt their observation and valuation approaches, and the sample period is the same 1982-99. Results are reported in Panels A, B, and C of Table 3; Panel D extends the sample to 2003. To conserve space, we report GMM results only for model (11), as the simpler model (9) with uncorrelated disturbances is comprehensively rejected for all combinations of explanatory variables (Q_E, $\hat{Q}$, *LTG*, and cash flow) and instruments[2]. We include OLS results for (9) in Panel A for comparison with the earlier literature and with Table 1: not surprisingly, the cash flow coefficient is positive and significant in all cases. In the combined model, Q_E is found to be positive and significant, while $\hat{Q}$ is insignificant, although it is significant when used as the sole control for growth opportunities.

---

[2] Results not reported and are available on request.

GMM results are presented in Panels B, C, and D. Again to conserve space, we report results only for the cases with $\hat{Q}$ and Q_E as the controls for growth opportunities; results with *LTG* are qualitatively similar and are not reported. Panels B and C correspond to Panels A and C in Table 2: Panel B uses instruments as in CHO, while Panel C uses all available lags of length greater than two of investments and cash flow as instruments. In Panel B, all estimated coefficients on Q_E, $\hat{Q}$, and cash flow are statistically insignificant and standard errors large, indicating that the instruments lack power. All specification tests fail to reject, which further confirms the inference of low power due to the limited instrument set.

When the instrument set is expanded to include all available lags of length greater than two (Panel C), cash flow becomes significant, as in Panels B and C in Table 2. Both Q_E and $\hat{Q}$ are positive and significant as sole controls for growth opportunities, but Q_E drives out $\hat{Q}$ in the combined model, which is contrary to CHO's finding. The specification tests fail to reject in all cases, indicating that the models are well-specified.

Panel D incorporates more recent data by extending the sample period to 1982-2003. Results are very similar to Panel C, with cash flow positive and significant, and Q_E and $\hat{Q}$ also positive and significant when used alone. As in Panel C, however, Q_E dominates $\hat{Q}$ in the combined model. All specification tests are again satisfied. Overall, the cash flow evidence from this table is consistent with that from Table 2 – estimated investment-cash flow sensitivity is positive and significant when efficient estimators are used. The evidence on the relative superiority of Q_E and $\hat{Q}$ is now in favor of Q_E .

3.5 Timing of Analyst Forecasts and the Valuation of Equity

We address two additional issues regarding CHO methodology and then examine the evidence. Specifically, we look at the timing of the analyst forecasts and the equity valuation formula used in CHO.

Regarding analyst forecasts, CHO use the earliest forecast in a year to determine $\hat{Q}$. This approach results in a median gap of 10 months between the date of investment and the date of one-year ahead analyst forecast observations. As a result, $\hat{Q}$ is partially codetermined with investment. We overcome this problem by increasing the gap to at least 12 months for one-year ahead forecasts and 24 months for next-to-next year forecasts.[3]

Next, we address a problem in the calculation of the value of equity. CHO uses below formula for equity valuation:

$$\hat{V}_t = E_t\left[\pi_t + \rho_t\pi_{t+1} + \rho_t{}^2(1+LTG_t)\overline{\pi}_t + \rho_t{}^3(1+LTG_t)^2\,\overline{\pi}_t + \rho_t{}^4(1+LTG_t)^3\,\overline{\pi}_t + \rho_t{}^5\,\frac{(1+LTG_t)^3\,\overline{\pi}_t}{\overline{r}-\overline{g}}\right] \tag{13}$$

[3] If data are not available for a one-year ahead forecast with a gap of 12 months, one with a gap of 13 months is chosen, and so on, to a maximum gap of 18 months. If data are not available for a next-to-next year forecast with a gap of 24 months, we search for a similar forecast with a gap of 25, 26, or 27 months, failing which we use the earliest next-year forecast with a gap of 23 to 19 months. By this modified algorithm, the mean (median) gap for one-year-ahead forecasts is 12.16 (12) months, while it is 21.91 (22) months for two-years-ahead forecasts This represents a major improvement over the CHO algorithm for the one-year-ahead forecasts and a marginal one for the two-years-ahead ones.

where $E_t[\cdot]$ represents analyst expectations at time-t , π_t and π_{t+1} are earnings at the end of the first and second years, $\overline{\pi}_t$ is the average of π_t and π_{t+1} , LTG_t is the long term growth rate, ρ_t is the single period discount factor, and $\bar{r}-\bar{g}$ the difference between the long run average returns on equity and average nominal GDP growth rate (assumed to be 9%). The earnings are projected for the end of the year, so the first EPS π_t should be discounted for one year, the second EPS π_{t+1} for two years, and so on. Moreover, the last term in the equation that refers to the present value of the terminal value of the stock should be replaced by $\rho_t^4 \frac{(1+LTG)^4 \overline{\pi}_t}{\bar{r}-\bar{g}}$. Given the role that this term typically plays in discounted cash flow valuations, inaccuracy in the formula may lead to undervaluation of the stock.

In Table 4, we use the independently reconstructed dataset as in Table 3 but address the issues related to the timing of analyst forecasts and the equity valuation formula as discussed above. Following standard practice, utility and financial firms are excluded from the sample, and outliers are excluded on a year-by-year basis rather than for the sample as a whole.

Panel A presents OLS results, which are similar to OLS results in Tables 1 and 3 and in earlier papers. As in Table 3, Q_E is positive and significant in the combined model, while $\hat{Q}$ is insignificant. Cash flow is positive and significant in all cases. Panels B, C, D, and E present GMM results. Panel B reports results from using the CHO instrument set. The specification tests reveal significant problems with the models: the *J*-test strongly rejects in all three cases while the m_2 test does in two. The estimated coefficient on lagged investment is significantly negative in two out of three cases, which is difficult to interpret meaningfully. Cash flow is insignificant in two cases, and positive and significant in one; Q_E is positive and significant while $\hat{Q}$ is insignificant. The main impression from this panel is the general rejection of the specification tests, which suggests that the choice of instruments is inappropriate.

Panel C examines the case in which negative cash flow observations are excluded to eliminate the impact of financial distress as in Allayannis and Mozumdar (2004) and Bhagat, Moyen, and Suh (2005). The specification tests now do much better: the m_2 test fails to reject in all three cases while the *J*-test rejects in one. The common factors restriction is satisfied in all cases as well. Cash flow, in addition to Q_E, is now found to be consistently positive and significant, while $\hat{Q}$ is significantly negative in one specification and insignificant in the other. Panel D reports results for the sample including negative cash flow observations but with the enhanced instrument set. Cash flow is again found to be positive and significant, as in Panel C, but the magnitude of the estimated sensitivity is lower. Q_E is consistently positive and significant. $\hat{Q}$ is positive and significant when used as the sole measure of growth opportunities but is driven out by Q_E in the combined model. With the exception of the *J*-test for the $\hat{Q}$-based model, all specification tests fail to reject.

Finally, Panel E presents results for the extended sample (1984-2003), positive cash flows, and enhanced instruments case. The findings are qualitatively very similar to those in Panels C and D: cash flow and Q_E are consistently positive and significant, while $\hat{Q}$ is positive and significant when used by itself but insignificant in the combined model. All specification tests fail to reject, with the exception of the common factor test which rejects the $\hat{Q}$-based model.

Overall, the evidence suggests that when an efficient instrument set is used, analyst forecast based measure $\hat{Q}$ cannot displace cash flow and the traditional stock price-based Q_E as the determinants of investment.

4. Estimation Based on Higher Order Moments

4.1 The Erickson and Whited (2000) Estimator

As with CHO, EW too indicate that biases induced by measurement error (u in equation 6) in the observed proxy (Q_E) for q may be substantial and may be responsible for the estimated coefficients on Q_E being low and those on cash flow being high, as reported in earlier papers. They propose a class of measurement error-consistent GMM estimators that utilize the information in the high (third and higher) order moments of the data to explicitly separate out the impact of such error. The first step in their approach is to obtain the residuals from projecting I/K, Q_E, and q on **z** (suppressing the time subscript)

$$(\gamma,\omega,\eta)=(\frac{I}{K},Q_E,q)-z\mu(\frac{I}{K},Q_E,q) \tag{14}$$

where $\mathbf{z}\mu(x)$ projects x on **z,** i.e., $\mu(x) = (E(\mathbf{z}'\mathbf{z}))^{-1}E(\mathbf{z}'x)$. The moments of ω and γ can be estimated from their sample counterparts and can in turn be used to estimate (β_1 ,$\boldsymbol{\alpha}$) where $\boldsymbol{\alpha} = (a,\mathbf{B})$. (Note that $\boldsymbol{\alpha} = \mu(I/K) - \beta_1\mu(Q)$.) EW use three second order moments,

$$E(\gamma^2) = \beta_1^2 E(\eta^2) + E(\varepsilon^2) \tag{15}$$

$$E(\omega\gamma) = \beta_1 E(\eta^2) \tag{16}$$

$$E(\omega^2) = E(\eta^2) + E(u^2) \tag{17}$$

two third order moments,

$$E(\gamma^2\omega) = \beta_1^2 E(\eta^3) \tag{18}$$

$$E(\gamma\omega^2) = \beta_1 E(\eta^3) \tag{19}$$

and three fourth order moments

$$E(\gamma^3\omega) = \beta_1^3 E(\eta^4) + 3\beta_1 E(\eta^2)E(\varepsilon^2) \tag{20}$$

$$E(\omega^2\gamma^2) = \beta_1^2 E(\eta^4) + E(\eta^2)\left[E(\varepsilon^2) + E(u^2)\right] + E(\varepsilon^2)E(u^2) \tag{21}$$

$$E(\omega^3\gamma) = \beta_1\left[E(\eta^4) + 3E(\eta^2)E(u^2)\right] . \tag{22}$$

Equations (15-19) constitute an exactly identified system of five equations in five unknowns (β_1, $E(\eta^2)$, $E(\varepsilon^2)$, $E(u^2)$, $E(\eta^3)$) which yields their GMM3 estimator. The full set of equations (15-22) represents an over-identified system of eight equations in six unknowns ($E(\eta^4)$ is the sixth unknown) which yields their GMM4 estimator. They also use a system including fifth order moments to generate a GMM5 estimator. GMM4 and GMM5 are based on over-identified models that are implemented with the inverse of the influence function-adjusted asymptotic covariance matrix (Newey and McFadden (1994)) as the weighting matrix of the quadratic form objective function.

In addition to the usual *J*-test for over-identifying restrictions for the GMM4 and GMM5 estimators (Sargan (1958), Hansen (1982)), the EW approach also requires a joint test for the two following conditions to be satisfied for the model to be identified:

$$\beta_1 \neq 0 \tag{23}$$

$$E(\eta^3) \neq 0. \tag{24}$$

Note that this requires the identification test to reject and the over-identification test to fail to reject.

Applying their estimators to a balanced panel of 737 U.S. manufacturing firms over 1992-1995, EW find that for each of the four years, the estimated coefficient on q is an order of magnitude greater than the OLS estimate and is *always highly significant* statistically. They also find that the estimated coefficient on cash flow is *always statistically insignificant*. The identification and over-identification tests are *all* uniformly satisfied. They conclude that 'most of the stylized facts produced by investment-cash flow regressions are artifacts of measurement error' and that 'cash flow does not matter'. Recently Almeida, Campello and Galvao (2010) scrutinize EW methodology and show that it is prone to yield biased and inefficient estimates. Erickson and Whited (2010), on the other hand, discuss that when appropriate starting values and a proper proxy for q is used in estimations, EW methodology has reasonable performance. Both of these papers show positive investment-cash flow sensitivities. It is curious that Erickson and Whited (2010) yield support for positive investment-cash flow sensitivities whereas the earlier study of EW does not. Can the difference be due to the changing financial structure of corporations from mid 1990s to 2000s or is there an alternative explanation? Below, we show that even in the earlier periods cash flow cannot be disregarded as an artifact of measurement errors in q.

Our analysis suggests that the original EW findings are not sufficiently general to justify their categorical conclusions. We address these issues as follows. First, we revisit the EW sample and methodology. Here we work on the EW's original data after addressing a data discrepancy and removing negative q values which are not interpretable as investment opportunities cannot be negative. After addressing these data issues, we find that for the 1992-1995 period that is considered in EW, investment-cash flow sensitivity is positive and significant in 1995 and q is not significant in 1994 with GMM3 estimations. For periods before 1992 as well as for GMM5 estimation in 1995, identification tests fail in all cases except in 1983. Next, we consider a subsample of original EW sample where the data consists of those observations where all variables in the EW data are matched with COMPUSTAT. For this sample, we consider both the proxy for q used by EW as well as a more conventional proxy for q. Again, the results show a large number of identification test failures. These results are in line with Almeida, Campello and Galvo (2010) that suggest the poor performance of EW methodology cannot be attributed to the choice of the proxy for q. In cases where identification and overidentification tests are satisfied, results suggest that cash flow cannot be dismissed as a determinant of investment.

Next, we extend the sample to all manufacturing firms that have at least five years of continuous data in COMPUSTAT. We use EW's assumptions for this extended data as well as standard definitions of variables used in the literature. In these extended samples, identification tests are rejected in most cases but overidentifying restrictions are generally not satisfied. The results show strong evidence contradictory to the original EW conclusions regarding investment-cash flow sensitivity

Overall, these findings cast serious doubt on the original EW findings that positive investment-cash flow sensitivity is driven mainly by measurement errors in q and that EW is a powerful methodology to be employed for correcting measurement errors in q. We discuss these findings in detail in the following subsections.

4.2 The EW Evidence Revisited

Table 5, Panel A presents results from applying the EW estimator to the original data set used in Erickson and Whited (2000).[4] As the Erickson and Whited (2000) dataset covers years earlier than 1992, we report the results for the 1982-1995 sample period. For the 1992-1995 period, the results are consistent with those reported by EW – q is significant, cash flow is insignificant, the identification test rejects, and the *J*-tests fail to reject in each of the four years.[5]

Regarding proxy for q, we observe that the EW proxy for q provides negative values in a number of cases. Negative q values are not interpretable as investment opportunities cannot be negative. These negative values are driven by the numerator of the proxy for q used in EW when inventory is greater than the market value of equity plus book value of debt.[6]

---

[4] We are grateful to Toni Whited for kindly sharing her data and making available on her web page GAUSS programs for implementing the EW tests and estimators.

[5] The set of perfectly measured regressors includes an intercept term and cash flow. When we also include a debt ratings dummy and an interaction term between the dummy and cash flow, as in EW, we are able to exactly replicate their reported results. All our reported results are qualitatively very similar to those obtained with the extended regressor set. We have chosen to report results for the smaller regressor set to conserve space and because our primary objective is to examine the existence of investment-cash flow sensitivity *per se*, and not to distinguish between financially constrained and unconstrained firms.

[6] For example, for the firm with CUSIP 464893 in 1993, the proxy for q used in EW is -0.44. The numerator of this proxy for q is equal to the market value of debt (2.26) plus market value of equity (0.95)

Next, we construct a sample by accessing COMPUSTAT data for the list of CUSIP numbers and all variables in Erickson and Whited (2000). For 595 firms, we obtain an exact match with EW with respect to CUSIP numbers and all variables except finished goods inventories (INVFG in COMPUSTAT and FGINV in EW) and income before extraordinary items (IB in COMPUSTAT and INC in EW).[7] We find that income and finished goods inventories are incorrectly sorted in reverse order in terms of years. Two examples are provided in Appendix 2. In EW, cash flow is defined as income before extraordinary items plus depreciation scaled by the replacement value of capital stock. Thus the reverse sorting of income before extraordinary items is an important issue that needs to be addressed.

We first consider all 737 firms in the original EW data and change the sorting of income and finished goods inventories.[8] We omit observations with negative q variables during our analyses as they are not interpretable. The results are in Table 5, Panel B. For the period 1992-1995, all identification tests are satisfied. Overidentification tests are satisfied except for GMM5 estimations in 1995. Investment-cash flow sensitivity is positive and significant in 1995 and q is not significant in 1994 with GMM3 estimation. For periods before 1992, identification tests fail in all years except 1983. These results suggest that the original EW findings are not sufficiently general to justify their

---

minus inventory (4.06). Thus an inventory value greater than the sum of the market values of debt and equity results in negative q values.

[7] There are some variations in capital expenditure numbers as well but generally the numbers are close to each other in both samples.

[8] Since the subsample of 595 matched firms were reverse sorted in income and finished goods inventories, we consider reverse sorting for all 737 firms in the original data.

categorical conclusions using their dataset but with corrections. These findings also indicate that the low power of EW methodology that is discussed in Almeida, Campello and Galvao (2010) cannot be attributed to the choice of q as advocated in Erickson and Whited (2010).

Next we consider a subsample of EW dataset for 595 firms where we find an exact match with all variables with the COMPUSTAT. We use original EW assumptions on this sample for the 1982-1995 period. Again, we omit observations with negative q values in the estimations. The results are given in Table 6, Panel A. Identification tests fail in 8 out of 14 years. In the remaining 6 years, there are only 2 years where all GMM results show insignificant investment-cash flow sensitivity. We also find negative coefficients on the q variable. In short, these results do not provide support for good performance of EW methodology or the insignificance of investment-cash flow sensitivity after addressing measurement errors in q.

Variables in the EW dataset are non-standard in that they are constructed using algorithms that are highly assumption-, data-, and computation-intensive, as described in Whited (1992). It has been customary in the literature to use simpler variable definitions (see, e.g., Kaplan and Zingales (1997), Cleary (1999), Love (2003), and Almeida and Campello (2007)). We use variable definitions that are standard in the literature for this subsample by accessing COMPUSTAT for constructing these variables as explained in Appendix A1 and apply EW estimators with these definitions. To reduce the impact of outliers, we winsorize variables in the top and bottom percentiles of investment, q, and cash flow. Since we do not require the additional variables used in the original EW data that is not from COMPUSTAT such as investment tax credit rate, aggregate measure of the mix of structures and equipment, maturity distribution of debt, deflator for nonresidential investment, Moody's medium-grade dividend yield, etc., we extend the sample period beyond 1995 until 2003.

Results are presented in Table 6, Panel B. In the 1982-1995 period, which is comparable to the sample run with original EW assumptions (Table 6, Panel A), the results are qualitatively similar. Only 5 years out of 14 years until 1995 pass the identification tests, and only 2 out of 5 years do not provide any support for significant and positive investment-cash flow sensitivity. When we examine the sample period 1982-2003, 11 out of 22 years pass the identification tests and 8 out of these 11 years provide support for significant and positive investment-cash flow sensitivity.

Overall, the results with both the original EW assumptions as well as more conventional variable definitions fail to support EW conclusion that EW methodology has good finite sample performance and the significance of cash flow is entirely an artifact of measurement error in *q*.

4.3 Extending the Sample Size

We next examine the generality of findings by extending the sample to include all manufacturing firms (primary SIC codes 2000-3999) with at least five years of continuous data on COMPUSTAT. Results are in Table 7. Panel A employs original EW assumptions for the 1982-1995 period and omit observations with negative *q* values. Panel B uses standard variable definitions given in Appendix A1 for the 1982-2003 period. Investment, cash flow and q are winsorized at the top and bottom percentiles. Thus Table 7 is constructed in a similar way as in Table 6 but includes more observations per year.

By increasing the sample size, identification tests reject in every case in Panel A and in all cases but one in Panel B in Table 7. However, the performance of the *J*-tests of over-identification is mixed. In Panel A, of the 14 years in the sample, the GMM4 estimator's overidentifying restrictions are rejected in 8 years at the 10% significance level; and the GMM5 estimator's in 11 years. In Panel B, performance of *J*-tests of over-identification is better for the GMM4 estimator (8 out of 22 years are rejected) but not for the GMM5 estimator (15 out of 22 years are rejected). The combination of EW's identification and over-identification conditions has been found to be difficult to satisfy in other studies as well. See, e.g., Almeida and Campello (2007) and Polk and Sapienza (2009), who report having to exclude large sections of their data to satisfy these conditions.

In the models for which the specification tests are satisfied, the estimated regression coefficients provide strong evidence contradictory to the original EW conclusions regarding investment-cash flow sensitivity. [9,10]

[9] Paradoxically, it is the OLS estimator that appears to provide some support for the EW findings. In Table 7, Panel B, the coefficient on cash flow estimated by OLS is negative although not significant for 5 years and negative and significant for 2 years. However, we find that this result can be explained by the impact of negative cash flow observations as reported earlier by Allayannis and Mozumdar (2004) and Bhagat, Moyen, and Suh (2005). When negative cash flow observations are excluded from the sample, the OLS estimated coefficient on cash flow is consistently positive and significant, and the GMM evidence shows positive and significant coefficient in 16 out of 25 times. (Results in Appendix A3.)

Overall, these findings suggest that EW methodology is not easily applicable to real data as it is hard to satisfy both identification and overidentification tests, and the finding of positive cash flow coefficient in investment equations cannot be disregarded as an artifact of measurement errors in q.

5. Using Alternative Measures of q as Instruments in GMM Estimation

In this section, we examine the traditional instrumental variables approach for addressing the problem of measurement error in q. The availability of two noisy proxies for the firm's growth opportunities – one based on stock market prices and another on analyst expectations – allows using lags of one as instruments for the other as a solution to the problem of measurement error. We also revisit CHO's finding of specification tests rejecting models in which lagged values of Q_E (or $\hat{Q}$) are used as instruments with their own current values as regressors, and check if a similar result holds in our sample.

Note that there exists an important distinction between the GMM panel data estimator of Arellano and Bond (1991, 1998) and the simpler instrumental variables estimator of, e.g., Anderson and Hsiao (1981). In the latter approach, given an *a priori* theoretical restriction of zero correlation between the error at time t and the instrument at time t-s, the model has a *single* moment restriction $\Delta e_{it}\, z_{it-s} = 0$, $\forall\, t \in (s+1, s+2,\ldots,T)$, where Δe is

[10]According to Fazzari, Hubbard and Petersen (1988), investment should be positively correlated with internal funds if a company is not able to obtain external funding. Theoretical analyses by Gomes (2001), Alti (2003), and Moyen (2004) suggest that financing constraints may be neither necessary nor sufficient to induce the empirically observed sensitivity of investments to cash flow. These studies mainly focus on average q not being able to capture growth opportunities. To the extent that CHO and EW address measurement errors in q, this issue should not be a concern.

the regression error (in first-differenced form), z the instrument, T the number of years of data, and i indexes the firm. On the other hand, in the Arellano and Bond (1991, 1998) approach, the same theoretical restriction is used to generate T-s distinct moment conditions, i.e., a separate moment condition for each t:

$$\begin{pmatrix} \Delta e_{is+1} z_{i1} \\ \Delta e_{is+2} z_{i2} \\ \vdots \\ \Delta e_{iT} z_{iT-s} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

While the Anderson and Hsiao (1981) estimator is more intuitive, being based on a direct correspondence between regressors and instruments, it is inefficient relative to the GMM estimator. Arellano and Bond (1991) also develop a mixed model with disaggregated moment conditions for the lagged dependent variable(s) and time-aggregated ones for the other regressor(s), i.e., the moment conditions assume the form

$$\begin{pmatrix} y_{i1} & 0 & \cdots & 0 \\ 0 & y_{i1} & \cdots & 0 \\ 0 & y_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & y_{i1} \\ 0 & 0 & \cdots & y_{i2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & y_{iT} \\ \Delta x_{i3} & \Delta x_{i4} & \cdots & \Delta x_{iT} \end{pmatrix} \begin{pmatrix} \Delta e_{i3} \\ \Delta e_{i4} \\ \vdots \\ \Delta e_{iT} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

where the set of instrumental variables consists of the endogenous lagged dependent variables(s) y, and the other (possibly exogenous or predetermined) variables x.

As in Section 3, we find that the m_2 test for the lack of second order serial correlation is consistently rejected for model (9)[11], suggesting AR(1) errors as in (10) and the dynamic specification with common factor restrictions as in (12). Results from estimating (12) are reported in Table 8. Within each Panel A – D, the first row reports results for the case in which current and once-lagged first-differences of Q_E are used as regressor variables, and further lags of (in some panels, first-differenced) Q_E as instruments; the second row for Q_E as regressors and $\hat{Q}$ as instruments; the third row for $\hat{Q}$ as regressors and $\hat{Q}$ as instruments; and the fourth row for $\hat{Q}$ as regressors and Q_E as instruments. Panel A presents Anderson and Hsiao (1981) estimates using third and fourth lags of changes in investment, cash flow, and q as instruments. The estimator is clearly inefficient – in each of the four cases considered, estimated standard errors are large, resulting in few parameter estimates being significant. The dynamic feedback coefficient on lagged investment is also estimated at implausible negative values. None of the specification tests reject, which is natural given the imprecision of the estimator. Overall, the evidence from this panel suggests that the Anderson and Hsiao (1981) estimator is not adequately powerful to be useful here.

Panel B presents results for the mixed model of Arellano and Bond (1991) described above. Third and fourth lags of investment (in levels) are now used as instruments to generate two distinct moment conditions for each year in the sample. Instruments and moment conditions for the other variables – cash flow and q – are as in Panel A. While the standard errors are smaller and parameter estimates appear reasonable (Q_E and cash

[11] Results are available on request.

flow coefficients are positive and significant, while those on $\hat{Q}$ are insignificant),the *J*-test rejects consistently, i.e., the over-identifying conditions are not satisfied. At the same time, however, the m_2 and common factor tests fail to reject, suggesting that the model is probably not grossly misspecified.

Panel C presents results for the panel data GMM estimator of Arellano and Bond (1991, 1998), using a choice of instruments similar to that of CHO: third and fourth lags of all explanatory variables – investment, cash flow, and q – are used in levels to generate a total of six separate moment conditions for each year in the sample. Unlike CHO who report specification test rejections when lags of q are used as instruments, we find that specification tests are largely satisfied, with the exception of the *J*-test in two cases and the m_2 test in one. As in Panel B, estimated coefficients on cash flow and Q_E are consistently positive and significant, while those on $\hat{Q}$ are insignificant.

Panel D extends the instrument set to include all available lags (third and higher) of investment, cash flow, and *q*, as in Panels D and E of Table 4, and as described in Arellano and Bond (1991). Despite the vastly over-identified nature of the model, the m_2 test fails to reject in every case and the *J*-test rejects in one. Estimated standard errors are often approximately 30-40% smaller relative to Panel C, indicating substantial efficiency gains due to the augmented instruments set. The cash flow coefficient estimates are consistently smaller than those in Panel C, suggesting a possible over-fitting bias arising from using too many instruments (Arellano and Bond (1998), Wooldridge (2002, Alvarez and Arellano (2003)). At the same time, however, estimated coefficients on cash flow and

Q_E continue to be positive and significant, and those on $\hat{Q}$ to be insignificant, as in earlier panels. It thus appears that the possible too-many-instruments bias is not an overwhelming concern here. The common factors restriction is rejected in three out of four cases, especially for the $\hat{Q}$-based models, suggesting that the data are probably close to their limit in satisfying such a large number of moment conditions within such a parsimonious model.

Finally, Panel E considers the combined model in which both Q_E and $\hat{Q}$ are included as regressors as well as instruments. The top row in the panel is for the case in which the instruments set is restricted to the third and fourth lags of the relevant variables, and the bottom row for the case of all available lags (greater than two) being included. As in the earlier panels, estimated coefficients on Q_E and cash flow are always positive and significant. The coefficient on $\hat{Q}$ is insignificant. The *J*-tests and m_2 tests fail to reject in both cases; the common factors restriction test is satisfied in the former case but rejects in the latter, in line with results in Panels C and D.

We draw several inferences from the evidence of Table 8. First, we continue to find consistently positive and significant estimates of investment-cash flow sensitivity and thus this coefficient cannot be ruled out as an artifact of measurement error in *q*.[12] Second, the estimated coefficient on Q_E being positive and significant and that on $\hat{Q}$ being mostly insignificant contradict CHO's conclusion that analyst expectations-based

[12] The positive cash flow sensitivity is despite assuming cash flow to be endogenous. This is in contrast to Blundell, Bond, Devereux, and Schiantarelli (1992) who find current cash flow to be a significant determinant of investment only if it is assumed to be exogenous.

measures of q are superior to stock price-based ones in explaining firm investments. Third, the fact that whether one uses lags of Q_E or $\hat{Q}$ as instruments makes little difference to estimated parameters and standard errors implies that there is no real benefit to instrumenting Q_E with lags of $\hat{Q}$. This is a fortunate finding given that stock price data are available for all publicly traded firms but analyst expectations for only a few. For studies using international data in particular, constructing $\hat{Q}$ is impossible for most countries, while Q_E can be constructed from stock market data. Finally, while there is some evidence of an over-fitting bias due to too many instruments, it is not severe enough in this sample to alter results qualitatively.

6. Conclusion

A significant branch of the literature in finance and economics analyzes the sensitivity of investment to internal funds. While studies that examine natural shocks to cash flows indicate significant sensitivity of investment to internal funds, the findings from panel data methods used by CHO and EW indicate that there is no sensitivity when measurement errors in q are addressed. We reexamine the apparent disconnect between these findings.

We find that cash flow cannot be dismissed as a determinant of investment. There is little support for the view that such sensitivities are an artifact of measurement errors in q. We also propose and implement a methodology that uses lagged values of stock market based and analyst forecast based proxies for q as instruments, which yields models that are at once powerful and well-specified. The estimated coefficients strongly suggest that a stock price-based measure of q is superior to an analyst forecast-based one. Estimated investment-cash flow sensitivities, while being smaller than their OLS counterparts, are consistently positive and significant.

References

Allayannis, George, and Abon Mozumdar, The impact of negative cash flow and influential observations on investment-cash flow sensitivity estimates, 2004, *Journal of Banking and Finance*, 28, 901-930.

Almeida, Heitor, and Murillo Campello, 2007, Financial constraints, asset tangibility and corporate investment, *Review of Financial Studies* 20, 1429-1460.

Almeida, Heitor, Murillo Campello and Antonio F. Galvao, 2010, Measurement errors in investment equations, forthcoming, *Review of Financial Studies* 23, 3279-3328.

Alti, Aydogan, 2003, How sensitive is investment to cash flow when financing is frictionless? *Journal of Finance* 58, 707–722.

Alvarez, Javier, and Manuel Arellano, 2003, The time-series and cross-sectional asymptotics of dynamic panel data estimators, *Econometrica* 71, 1121-1159.

Anderson, T.W., and C. Hsiao, 1981, Estimation of dynamic models with error components, *Journal of the American Statistical Association* 76, 598-606.

Arellano, Manuel, and Stephen R. Bond, 1991, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, 58, 277-297.

Arellano, Manuel, and Stephen R. Bond, 1998, Dynamic panel data estimation using DPD98 for Gauss: A guide for users, Oxford University and CEMFI, Spain.

Bhagat, Sanjai, Nathalie Moyen, and Inchul Suh, Investment and internal funds of distressed firms, 2005, *Journal of Corporate Finance* 11, 449-472.

Blanchard, Oliver J., Florencio Lopez-de-Silanes, and Andrei Shleifer, 1994, What do firms do with cash windfalls? *Journal of Financial Economics* 36.

Blundell, Richard, Stephen Bond, Michael Devereux, and Fabio Schiantarelli, 1992, Investment and Tobin's q, *Journal of Econometrics* 51, 233-257.

Brown, James R., Steven M. Fazzari and Bruce C. Petersen, 2009, Financing Innovation and Growth: Cash Flow, External Equity, and the 1990s R&D Boom, Journal of Finance, 64, 151-185.

Cleary, Sean, 1999, The relationship between firm investment and financial status, *Journal of Finance* 54, 673-692.

Cummins, Jason, Kevin Hassett, and Stephen Oliner, 2006, Investment behavior, observable expectations, and internal funds, *American Economic Review* 96, 796-810.

Erickson, Timothy, and Toni M. Whited, 2000, Measurement error and the relationship between investment and *q*, *Journal of Political Economy*, 108, 1027-1057.

Erickson, Timothy, and Toni M. Whited, 2002, Two-Step GMM estimation of the errors-in-variables model using high-order moments, *Econometric Theory* 18, 776-799.

Erickson, Timothy, and Toni M. Whited, 2006, On the accuracy of different measures of q, *Financial Management* 35, 5-33.

Erickson, Timothy, and Toni M. Whited, 2010, Treating measurement error in Tobin's q, Working Paper, University of Rochester.

Fazzari, Steven M., R. Glenn Hubbard, and Bruce C. Petersen, Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, 1988, 141-195.

Fee, Edward C., Charles Hadlock, Joshua Pierce, 2009, Investment, Financing Constraints, and Internal Capital Markets: Evidence from the Advertising Expenditures of Multinational Firms, Review of Financial Studies 22, 2361-2392.

Gomes, Joao, 2001, Financing Investment, *American Economic Review* 91, 1263-1285.

Greene, William H., 2003, Econometric Analysis, 5th edition, Prentice Hall, New Jersey.

Hansen, Lars P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029-1054.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen, 1988, Estimating vector autoregressions with panel data, *Econometrica* 56, 1371-1395.

Hubbard, R. Glenn, 1998, Capital-market imperfections and investment. *Journal Of Economic Literature,* 36, 193-225.

Kaplan, Steve, and Luigi Zingales, 1997, Do investment-cash flow sensitivities provide useful measures of financing constraints? *The Quarterly Journal of Economics* 112, 169-215.

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

Love, Inessa, 2003, Financial development and financing constraints: International evidence from the structural investment model, *Review of Financial Studies* 16, 765-791.

Moyen, Nathalie, 2004, Investment-cash flow sensitivities: Constrained versus unconstrained firms. *Journal of Finance*, 59, 2061-2092.

Newey, Whitney, and McFadden, D., 1994, Large sample estimation and hypothesis testing, *Handbook of Econometrics* 4 (Engle R., McFadden D., eds.), Chapter 36.

Polk, Christopher K., and Paola Sapienza, 2009, The stock market and corporate investment: a test of catering theory, *Review of Financial Studies* 22, 187-217.

Poterba, James M., 1988, Comment on 'Financing constraints and corporate investment', Brookings Paper on Economic Activity 1, 200–204.

Rauh, Joshua D., 2006, Investment and financing constraints: evidence from the funding of corporate pension plans, *Journal of Finance* 61, 33-71.

Sargan, John D., 1958, The estimation of economic relationships using instrumental variables, *Econometrica* 26, 393-415.

Scholes, Myron, and Joseph Williams, 1977, Estimating betas from nonsynchronous data, *Journal of Financial Economics* 5, 309-327.

Stein, Jeremy 2003. Agency, information, and corporate investment. In G. Constantinides, M. Harris, and R. Stulz (eds.), *Handbook of the Economics of Finance*, Elsevier, North Holland.

Wooldridge, J.M., 2002, Econometric analysis of cross-section and panel data, MIT Press.

Whited, Toni M., Debt, liquidity constraints, and corporate investment: Evidence from panel data, *Journal of Finance* 47, 1425-1460.

TABLE 1
Static Model Estimation Using Shifted CHO Data

This table reports results from estimating

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = \beta_1 \Delta q_{it} + \beta_2 \Delta\left(\frac{CF_{it}}{K_{it-1}}\right) + \Delta YEAR_t + w^*_{it}$$

using the original dataset of Cummins, Hassett, and Oliner (2006), with the cash flow data shifted forward by one year. Q_E, $\hat{Q}$, or *LTG* are used as proxies for *q*; the last row in each panel includes both Q_E and $\hat{Q}$ as explanatory variables. Q_E represents the stock price-based proxy for *q*, $\hat{Q}$ represents the analyst forecast-based proxy for *q*, and *LTG* is the long term growth rate predicted by analysts as reported by I/B/E/S. The subscripts *i* and *t* indicate firm and year, respectively, in the data panel. Panel A reports OLS results. Panels B, C, and D report GMM results. Panel B uses 3rd and 4th lags of normalized investment and cash flow as instruments. Panel C uses 3rd and 4th lags as instruments when both lags are available, and only 3rd lags when 4th lags are not available. Panel D uses all available lags of length greater than 2 of normalized investment and cash flow as instruments. Robust standard errors are reported in parentheses. [a], [b], and [c] denote significance at the 1%, 5%, and 10% levels, respectively. Panels B, C, and D also report *p*-values for *J*-tests for over-identifying conditions and m_2 tests for second order serial correlation as in Arellano and Bond (1991).

Ordinary Least Squares Estimation						
	Q_E	*LTG*	$\hat{Q}$	*CF/K*		
Panel A						
Sample period: 1984-1999 Number of firms: 1066 Number of observations: 9299	0.014[a] (0.001)			0.148[a] (0.015)		
		0.060[c] (0.036)		0.175[a] (0.017)		
			0.028[a] (0.003)	0.137[a] (0.016)		
	0.010[a] (0.001)		0.017[a] (0.003)	0.132[a] (0.015)		
Panel Data Generalized Method of Moments Estimation						
	Q_E	*LTG*	$\hat{Q}$	*CF/K*	*J*	m_2
Panel B						
Sample period: 1986-1999 Number of firms: 1066 Number of observations: 7167 Instruments used: $(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i = 3,4	0.040[a] (0.008)			0.459[a] (0.119)	0.005	0.008
		1.791[a] (0.250)		0.225[a] (0.072)	0.024	0.000
			0.111[a] (0.011)	0.280[a] (0.064)	0.042	0.003
	0.016[c] (0.009)		0.090[a] (0.023)	0.308[b] (0.131	0.042	0.009
Panel C						
Sample period: 1984-1999 Number of firms: 1066 Number of observations: 9299 Instruments used: $(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i =3, min(t-1,4)	0.033[a] (0.008)			0.393[a] (0.127)	0.003	0.001
		1.694[a] (0.363)		0.301[a] (0.113)	0.039	0.000
			0.100[a] (0.019)	0.276[b] (0.129)	0.044	0.000
	0.004 (0.009)		0.095[a] (0.021)	0.280[b] (0.129)	0.032	0.000
Panel D						
Sample period: 1984-1999 Number of firms: 1066 Number of observations: 9299 Instruments used: $(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i =3,4,….,t-1	0.024[a] (0.004)			0.209[a] (0.056)	0.001	0.000
		0.987[a] (0.148)		0.246[a] (0.052)	0.019	0.000
			0.052[a] (0.008)	0.164[a] (0.054)	0.012	0.000
	0.009[c] (0.005)		0.038[a] (0.010)	0.164[a] (0.054)	0.006	0.000

TABLE 2
Dynamic Model Estimation Using Shifted/Unshifted CHO Data

This table reports GMM results from estimating ρ, β_1, and β_2 in the restricted model

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = \rho\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + \beta_1\Delta q_{it} - \rho\beta_1\Delta q_{it-1} + \beta_2\Delta\left(\frac{CF_{it}}{K_{it-1}}\right) - \rho\beta_2\Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it}$$

from the unrestricted model

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = b_0\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + b_1\Delta q_{it} + b_2\Delta q_{it-1} + b_3\Delta\left(\frac{CF_{it}}{K_{it-1}}\right) + b_4\Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it}$$

as in Blundell, Bond, Devereaux, and Schiantarelli (1992), using the original dataset of Cummins, Hassett, and Oliner (2006). The cash flow data are shifted forward by one year in Panels A, B, and C, and unshifted in Panel D. Q_E and/or $\hat{Q}$ are used as proxies for q; the last row in each panel includes both Q_E and $\hat{Q}$ as explanatory variables. Q_E represents the stock price-based proxy for q, while $\hat{Q}$ represents the analyst forecast-based proxy for q. The subscripts i and t indicate firm and year, respectively, in the data panel. Panel A uses 3rd and 4th lags of normalized investment and cash flow as instruments. Panel B uses 3rd and 4th lags as instruments when both lags are available, and only 3rd lags when 4th lags are not available. Panels C and D use all available lags of length greater than 2 of normalized investment and cash flow as instruments. Robust standard errors are reported in parentheses. [a], [b], and [c] denote significance at the 1%, 5%, and 10% levels, respectively. Also reported are p-values for J-tests for over-identifying conditions, m_2 tests for second order serial correlation as in Arellano and Bond (1991), and the BBDS comfac test of common factor restrictions.

Panel Data Generalized Method of Moments Estimation							
	Q_E	$\hat{Q}$	*CF/K*	*I/K*	*J*	m_2	*Comfac*
Panel A							
Sample period: 1986-1999	0.027[a]		0.239[b]	0.240[a]	0.041	0.471	0.309
Number of firms: 1066	(0.007)		(0.103)	(0.027)			
Number of observations: 7167		0.085[a]	0.098	0.074[a]	0.152	0.175	0.898
Instruments used:		(0.020)	(0.104)	(0.012)			
$(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i = 3,4	0.012	0.070[a]	0.142	0.048[a]	0.102	0.222	0.994
	(0.008)	(0.022)	(0.107)	(0.008)			
Panel B							
Sample period: 1985-1999	0.020[a]		0.286[a]	0.287[a]	0.046	0.932	0.236
Number of firms: 1066	(0.007)		(0.114)	(0.031)			
Number of observations: 8233		0.083[a]	0.196[c]	0.124[a]	0.152	0.428	0.761
Instruments used:		(0.021)	(0.115)	(0.019)			
$(I/K)_{t-i}$ and $(CF/K)_{t-i}$,	0.004	0.078[a]	0.208[c]	0.104[a]	0.104	0.449	0.946
i =3, min(t-1,4)	(0.008)	(0.024)	(0.115)	(0.016)			
Panel C							
Sample period: 1985-1999	0.014[a]		0.151[a]	0.373[a]	0.024	0.424	0.016
Number of firms: 1066	(0.004)		(0.049)	(0.033)			
Number of observations: 8233		0.035[a]	0.116[a]	0.366[a]	0.035	0.347	0.021
Instruments used:		(0.008)	(0.047)	(0.037)			
$(I/K)_{t-i}$ and $(CF/K)_{t-i}$,	0.008[c]	0.024[a]	0.118[a]	0.363[a]	0.031	0.350	0.051
i =3,4,….,t-1	(0.005)	(0.010)	(0.048)	(0.035)			
Panel D							
Sample period: 1984-1999	0.018[a]		0.004	0.410[a]	0.018	0.665	0.001
Number of firms: 1066	(0.004)		(0.043)	(0.034)			
Number of observations: 9299		0.040[a]	-0.011	0.381[a]	0.020	0.808	0.003
Instruments used:		(0.009)	(0.043)	(0.039)			
$(I/K)_{t-i}$ and $(CF/K)_{t-i}$,	0.011[b]	0.025[b]	-0.014	0.375[a]	0.020	0.738	0.011
i =3,4,….,t-1	(0.004)	(0.010)	(0.043)	(0.038)			

TABLE 3
Dynamic Model Estimation Using Replicated CHO Data

This table reports results using an independently reconstructed dataset seeking to replicate that of Cummins, Hassett, and Oliner (2006) (CHO). Panels A, B, and C use the same sample period as CHO, i.e., 1982-99; Panel D extends the sample period to 1982-2003. Q_E and/or $\hat{Q}$ are used as proxies for q; the last row in each panel includes both Q_E and $\hat{Q}$ as explanatory variables. Q_E represents the stock price-based proxy for q, while $\hat{Q}$ represents the analyst forecast-based proxy for q. The subscripts i and t indicate firm and year, respectively, in the data panel. Panel A reports OLS results from estimating

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = \beta_1 \Delta q_{it} + \beta_2 \Delta\left(\frac{CF_{it}}{K_{it-1}}\right) + \Delta YEAR_t + w^*_{it}$$

Panels B, C, and D report GMM results from estimating ρ, β_1, and β_2 in the restricted model

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = \rho\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + \beta_1 \Delta q_{it} - \rho\beta_1\Delta q_{it-1} + \beta_2 \Delta\left(\frac{CF_{it}}{K_{it-1}}\right) - \rho\beta_2\Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it}$$

from the unrestricted model

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = b_0\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + b_1 \Delta q_{it} + b_2\Delta q_{it-1} + b_3 \Delta\left(\frac{CF_{it}}{K_{it-1}}\right) + b_4\Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it}$$

as in Blundell, Bond, Devereaux, and Schiantarelli (1992). Panel B uses 3[rd] and 4[th] lags of normalized investment and cash flow as instruments. Panels C and D use all available lags of length greater than 2 of normalized investment and cash flow as instruments. Robust standard errors are reported in parentheses. [a], [b], and [c] denote significance at the 1%, 5%, and 10% levels, respectively. Also reported are p-values for J-tests for over-identifying conditions, m_2 tests for second order serial correlation as in Arellano and Bond (1991), and the BBDS comfac test of common factor restrictions.

Ordinary Least Squares Estimation							
	Q_E	*LTG*	$\hat{Q}$	*CF/K*			
Panel A							
Sample period: 1983-1999 Number of firms: 1631 Number of observation:14195	0.058[a] (0.006)			0.188[a] (0.013)			
		0.068 (0.078)		0.202[a] (0.013)			
			0.038[a] (0.009)	0.201[a] (0.013)			
	0.056[a] (0.007)		0.007 (0.010)	0.188[a] (0.013)			
Panel Data Generalized Method of Moments Estimation							
	Q_E	$\hat{Q}$	*CF/K*	*I/K*	*J*	m_2	*Comfac*
Panel B							
Sample period: 1986-1999 Number of firms: 1631	-0.036 (0.058)		0.067 (0.061)	0.245[a] (0.039)	0.301	0.645	0.680
Number of observations:9302 Instruments used:		0.002 (0.089)	0.068 (0.062)	0.286[a] (0.059)	0.271	0.393	0.576
$(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i = 3,4	-0.058 (0.068)	0.016 (0.104)	0.063 (0.068)	0.274[a] (0.049)	0.484	0.327	0.722
Panel C							
Sample period: 1984-1999 Number of firms: 1631	0.082[a] (0.023)		0.134[a] (0.038)	0.205[a] (0.045)	0.142	0.929	0.469
Number of observation:12564 Instruments used:		0.107[a] (0.040)	0.144[a] (0.038)	0.260[a] (0.042)	0.157	0.502	0.228
$(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i =3,4,….,t-1	0.067[c] (0.025)	0.045 (0.044)	0.136[a] (0.037)	0.218[a] (0.036)	0.264	0.858	0.536
Panel D							
Sample period: 1984-2003 Number of firms: 2132	0.077[a] (0.016)		0.100[a] (0.029)	0.286[a] (0.041)	0.525	0.577	0.322
Number of observation:17118 Instruments used:		0.091[a] (0.027)	0.110[a] (0.027)	0.302[a] (0.0043)	0.230	0.263	0.170
$(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i =3,4,….,t-1	0.067[a] (0.017)	0.043 (0.027)	0.101[a] (0.026)	0.272[a] (0.045)	0.440	0.560	0.618

TABLE 4

Dynamic Panel Estimation Using Modified CHO Data

This table reports results using an independently reconstructed dataset seeking to improve upon that of Cummins, Hassett, and Oliner (2006) (CHO). Panels A, B, and C use the same sample period as CHO, i.e., 1982-99; Panel D extends the sample period to 1982-2003. Q_E and/or $\hat{Q}$ are used as proxies for q; the last row in each panel includes both Q_E and $\hat{Q}$ as explanatory variables. Q_E represents the stock price-based proxy for q, while $\hat{Q}$ represents the analyst forecast-based proxy for q. The subscripts i and t indicate firm and year, respectively, in the data panel. Panel A reports OLS results from estimating

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = \beta_1 \Delta q_{it} + \beta_2 \Delta\left(\frac{CF_{it}}{K_{it-1}}\right) + \Delta YEAR_t + w^*_{it}$$

Panels B, C, D, and E report GMM results from estimating ρ, β_1, and β_2 in the restricted model

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = \rho\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + \beta_1 \Delta q_{it} - \rho\beta_1 \Delta q_{it-1} + \beta_2 \Delta\left(\frac{CF_{it}}{K_{it-1}}\right) - \rho\beta_2 \Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it}$$

from the unrestricted model

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = b_0\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + b_1 \Delta q_{it} + b_2 \Delta q_{it-1} + b_3 \Delta\left(\frac{CF_{it}}{K_{it-1}}\right) + b_4 \Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it}$$

as in Blundell, Bond, Devereaux, and Schiantarelli (1992). Panels B and C use 3rd and 4th lags of normalized investment and cash flow as instruments. Panels D and E use all available lags of length greater than 2 of normalized investment and cash flow as instruments. Panels C and E exclude negative cash flow observations. Robust standard errors are reported in parentheses. [a], [b], and [c] denote significance at the 1%, 5%, and 10% levels, respectively. Also reported are *p*-values for *J*-tests for over-identifying conditions, m_2 tests for second order serial correlation as in Arellano and Bond (1991), and the BBDS comfac test of common factor restrictions.

Ordinary Least Squares Estimation							
	Q_E	*LTG*	$\hat{Q}$	*CF/K*			
Panel A							
Sample period: 1983-1999 Number of firms: 1272 Number of observation:10983	0.060[a] (0.007)			0.119[a] (0.015)			
		0.187[c] (0.071)		0.131[a] (0.015)			
			0.025[a] (0.006)	0.132[a] (0.015)			
	0.058[a] (0.007)		0.006 (0.007)	0.120[a] (0.015)			
Panel Data Generalized Method of Moments Estimation							
	Q_E	$\hat{Q}$	*CF/K*	*I/K*	*J*	m_2	*Comfac*
Panel B							
Sample period: 1986-1999 Number of firms: 1272 Number of observations:7153 Instruments used: $(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i = 3,4	0.100[b] (0.045)		0.029 (0.052)	-0.017[a] (0.004)	0.003	0.033	0.9996
		0.090 (0.068)	0.104[c] (0.053)	0.149[a] (0.045)	0.014	0.200	0.782
	0.114[b] (0.047)	-0.028 (0.080)	0.039 (0.051)	-0.012[a] (0.003)	0.006	0.050	1.000
Panel C							
Sample period: 1986-1999 Number of firms: 1147 Number of observation:6219 Instruments used: $(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i =3,4	0.076[b] (0.038)		0.207[a] (0.067)	0.138[b] (0.061)	0.062	0.612	0.853
		-0.045 (0.0057)	0.243[a] (0.073)	0.202[b] (0.102)	0.137	0.732	0.964
	0.136[a] (0.054)	-0.146[b] (0.073)	0.235[a] (0.072)	0.197[a] (0.060)	0.392	0.819	0.930
Panel D							
Sample period: 1984-1999 Number of firms: 1272 Number of observation:9697 Instruments used: $(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i =3,4,….,t-1	0.111[a] (0.018)		0.051[b] (0.027)	0.181[a] (0.041)	0.305	0.342	0.455
		0.064[b] (0.026)	0.081[a] (0.027)	0.219[a] (0.033)	0.053	0.107	0.259
	0.103[a] (0.019)	0.026 (0.027)	0.047[c] (0.028)	0.178[a] (0.040)	0.394	0.273	0.660
Panel E							
Sample period: 1984-2003 Number of firms: 1548 Number of observation:11816 Instruments used: $(I/K)_{t-i}$ and $(CF/K)_{t-i}$, i =3,4,….,t-1	0.075[a] (0.014)		0.069[a] (0.022)	0.172[a] (0.0037)	0.179	0.183	0.400
		0.047[b] (0.018)	0.089[a] (0.024)	0.230[a] (0.026)	0.127	0.122	0.081
	0.070[a] (0.014)	0.022 (0.020)	0.067[a] (0.022)	0.167[a] (0.039)	0.169	0.151	0.650

TABLE 5
Higher Order Moments Estimation Using Erickson and Whited (2000) Data

This table presents OLS and GMM results from estimating

$$\frac{I_t}{K_{t-1}} = a + \beta_1 q_t + \beta_2 \frac{CF_t}{K_{t-1}} + e_t$$

with Q_E used as a mismeasured proxy for *q* over the sample period 1982-95. Panel A gives the results for the original EW dataset of 737 firms used in Erickson and Whited (2000). Panel B gives the results when the EW dataset of 737 firms is corrected for reverse-sorting in income and finished goods inventories and negative q values are omitted. Number of observations in each year is reported under the years in italics. GMM estimates are obtained using the measurement error-consistent higher-order moments estimators of Erickson and Whited (2000). The set of perfectly measured regressors includes a constant and normalized cash flow. Robust standard errors for OLS and Newey-McFadden (1994) influence function adjusted standard errors for GMM are reported in parentheses. ^{a}, ^{b}, and ^{c} denote significance at the 1%, 5%, and 10% levels, respectively. Also reported are test statistics for identification tests and *J*-tests of over-identifying restrictions along with their *p*-values. Instances in which both identifying and over-identifying conditions are satisfied have results reported in grey cells.

Panel A

	ID Test	J-Test		QE				CF			
		GMM4	GMM5	OLS	GMM3	GMM4	GMM5	OLS	GMM3	GMM4	GMM5
1982	2.118	0.534	2.918	0.191^{a}	0.201	0.201	0.201	0.831^{a}	0.852	0.853	0.853
602	(0.347)	(0.766)	(0.713)	(0.007)	(0.002)	(0)	(0)	(0.037)	(0.015)	(0.012)	(0.012)
1983	3.898	2.341	3.417	0.017^{c}	0.051	0.062	0.061	-0.009	0.009	0.014	0.014
611	(0.142)	(0.31)	(0.636)	(0.009)	(0.036)	(0.012)	(0.007)	(0.018)	(0.051)	(0.042)	(0.042)
1984	5.584	3.782	4.088	0.018^{a}	0.033^{c}	0.059^{a}	0.07^{a}	-0.002	-0.049	-0.127^{b}	-0.159^{a}
615	(0.061)	(0.151)	(0.537)	(0.006)	(0.017)	(0.022)	(0.013)	(0.026)	(0.056)	(0.054)	(0.053)
1985	4.735	0.641	4.035	0.015^{b}	0.026	0.031^{a}	0.1^{b}	0.272^{b}	0.266^{b}	0.263^{b}	0.221
625	(0.094)	(0.726)	(0.544)	(0.006)	(0.026)	(0.012)	(0.046)	(0.123)	(0.124)	(0.119)	(0.139)
1986	2.617	1.441	6.661	-0.003	1.692	0.054	0.17	0.563^{a}	-0.163	0.538	0.489
643	(0.27)	(0.486)	(0.247)	(0.009)	(19.938)	(0.031)	(0.077)	(0.055)	(8.566)	(0.07)	(0.107)
1987	5.537	29.828	197.076	0.013^{a}	-0.039	0	0	0.21^{a}	0.214	0.211	0.211
669	(0.063)	(0)	(0)	(0.004)	(0.083)	(0.029)	(0.289)	(0.022)	(0.019)	(0.022)	(0.03)
1988	1.645	1.208	3.673	-0.006	0.345	0.001	-0.237	0.884^{a}	-0.007	0.867	1.469
685	(0.439)	(0.547)	(0.597)	(0.008)	(0.966)	(0.132)	(0.136)	(0.323)	(2.059)	(0.423)	(0.438)
1989	4.738	1.654	7.59	0.014^{a}	0.027^{a}	0.032^{a}	0.019^{a}	0.222^{a}	0.165^{b}	0.142^{c}	0.2^{a}
699	(0.094)	(0.437)	(0.18)	(0.004)	(0.007)	(0.01)	(0.006)	(0.058)	(0.064)	(0.074)	(0.047)
1990	4.835	3.86	6.418	0.011c	0.103	0.117^{b}	0.081^{b}	0.459^{a}	0.379^{a}	0.368^{a}	0.398^{a}
717	(0.089)	(0.145)	(0.268)	(0.006)	(0.071)	(0.05)	(0.04)	(0.017)	(0.112)	(0.122)	(0.094)
1991	4.976	1.017	2.482	0.013^{a}	0.062	0.114^{c}	0.128^{b}	0.487^{a}	0.454^{a}	0.418^{a}	0.409^{a}
737	(0.083)	(0.601)	(0.779)	(0.004)	(0.072)	(0.063)	(0.054)	(0.05)	(0.106)	(0.11)	(0.108)
1992	11.174	1.353	2.995	0.015^{a}	0.05^{b}	0.026^{a}	0.027^{a}	0.168^{a}	-0.082	0.089	0.077
737	(0.004)	(0.508)	(0.701)	(0.003)	(0.022)	(0.006)	(0.005)	(0.064)	(0.188)	(0.077)	(0.081)
1993	8.383	1.656	7.123	0.014^{a}	0.042^{a}	0.035^{a}	0.033^{a}	0.101^{a}	-0.055	-0.016	-0.008
737	(0.015)	(0.437)	(0.212)	(0.003)	(0.007)	(0.007)	(0.008)	(0.037)	(0.065)	(0.058)	(0.052)
1994	10.907	0.817	6.53	0.015^{a}	0.069	0.044^{a}	0.023^{a}	0.146^{a}	-0.34	-0.115	0.072
737	(0.004)	(0.665)	(0.258)	(0.003)	(0.047)	(0.008)	(0.004)	(0.053)	(0.444)	(0.093)	(0.06)
1995	8.009	2.546	5.505	0.019^{a}	0.048^{a}	0.038^{b}	0.063^{a}	0.138^{c}	-0.155	-0.054	-0.31^{b}
737	(0.018)	(0.28)	(0.357)	(0.004)	(0.013)	(0.015)	-0.012	(0.079)	(0.147)	(0.167)	(0.151)

Table 5, Panel B

	ID Test	J-Test		QE				CF			
		GMM4	GMM5	OLS	GMM3	GMM4	GMM5	OLS	GMM3	GMM4	GMM5
1982	1.496	3.467	10.065	0.074[a]	0.08	0.08	0.08	1.505[a]	1.508	1.508	1.508
598	(0.473)	(0.177)	(0.073)	(0.006)	(0)	(0)	(0)	(0.042)	(0.04)	(0.04)	(0.04)
1983	4.737	0.703	5.757	0.017[a]	0.098[a]	0.109[a]	0.113[a]	0.006	-0.695	-0.783	-0.822
607	(0.094)	(0.704)	(0.331)	(0.012)	(0.026)	(0.024)	(0.008)	(0.155)	(0.519)	(0.624)	(0.457)
1984	1.994	0.688	6.207	0.017[a]	0.066	0.062	0.042	0.019	-0.533	-0.485	-0.259
611	(0.369)	(0.709)	(0.287)	(0.008)	(0.022)	(0.019)	(0.008)	(0.083)	(0.336)	(0.292)	(0.141)
1985	1.188	1.343	6.88	-0.008	-0.04	-0.086	-0.039	0.577[a]	0.72	0.927	0.715
620	(0.552)	(0.511)	(0.23)	(0.009)	(0.046)	(0.016)	(0.023)	(0.152)	(0.214)	(0.145)	(0.148)
1986	0.055	7.007	6.185	-0.025	0.07	0	0.084	0.95[a]	0.806	0.912	0.785
639	(0.973)	(0.03)	(0.289)	(0.017)	(2.022)	(0.063)	(0.035)	(0.204)	(2.819)	(0.245)	(0.26)
1987	3.863	2.59	8.174	0.017[a]	0.007	0.025	0.024	0.215[a]	0.216	0.215	0.215
665	(0.145)	(0.274)	(0.147)	(0.005)	(0.046)	(0.004)	(0.003)	(0.017)	(0.015)	(0.017)	(0.017)
1988	2.279	0.601	14.217	-0.01	0.562	0.231	-0.032	0.827[a]	-0.45	0.29	0.877
680	(0.32)	(0.741)	(0.014)	(0.011)	(1.331)	(0.042)	(0.016)	(0.247)	(2.533)	(0.338)	(0.225)
1989	2.574	0.806	3.495	0.01[a]	0.023	0.022	0.059	0.207[a]	0.156	0.159	0.011
697	(0.276)	(0.668)	(0.624)	(0.004)	(0.01)	(0.007)	(0.025)	(0.043)	(0.068)	(0.055)	(0.124)
1990	3.916	2.368	10.606	0.01	0.082	0.059	0.15	0.423[a]	0.362	0.382	0.304
713	(0.141)	(0.306)	(0.06)	(0.007)	(0.029)	(0.012)	(0.009)	(0.032)	(0.093)	(0.062)	(0.132)
1991	2.186	0.459	3.461	0.005	0.038	0.111	0.173	0.418[a]	0.386	0.316	0.256
727	(0.335)	(0.795)	(0.629)	(0.006)	(0.094)	(0.067)	(0.061)	(0.059)	(0.154)	(0.139)	(0.169)
1992	10.292	1.841	5.261	0.013[a]	0.044[b]	0.031[a]	0.034[a]	0.133[a]	-0.026	0.039	0.022
732	(0.006)	(0.398)	(0.385)	(0.003)	(0.018)	(0.008)	(0.004)	(0.044)	(0.104)	(0.062)	(0.049)
1993	15.841	2.492	6.538	0.014[a]	0.033[a]	0.037[a]	0.034	0.075[a]	0.025	0.014	0.021
732	(0)	(0.288)	(0.257)	(0.003)	(0.006)	(0.005)	(0.004)	(0.021)	(0.036)	(0.038)	(0.031)
1994	5.782	0.724	8.169	0.01[a]	0.147	0.042[a]	0.028[a]	0.134[a]	-0.707	-0.061	0.023
732	(0.056)	(0.696)	(0.147)	(0.003)	(0.44)	(0.015)	(0.009)	(0.036)	(2.72)	(0.098)	(0.066)
1995	7.519	2.906	10.322	0.017[a]	0.03[a]	0.046[a]	0.057	0.098[a]	0.066[b]	0.028	0.002
732	(0.023)	(0.234)	(0.067)	(0.003)	(0.009)	(0.01)	(0.009)	(0.03)	(0.029)	(0.043)	(0.044)

TABLE 6
Higher Order Moments Estimation Using Reconstructed Subsample

This table presents OLS and GMM results from estimating

$$\frac{I_t}{K_{t-1}} = a + \beta_1 q_t + \beta_2 \frac{CF_t}{K_{t-1}} + e_t$$

with Q_E used as a mismeasured proxy for q for the 595 firms that matches with the Erickson and Whited (2000) sample and COMPUSTAT in all variables. A data discrepancy in the data concerning income and finished goods inventories are corrected in this sample. In Panel A, original EW assumptions are used for variable definitions and observations with negative Q_E values are omitted. Results are reported for the 1982-1995 period. In Panel B, variable definitions that are standard in the literature are used (see Appendix 1). Investment, cash flow and Q_E are winsorized at the bottom and top percentile to remove the outliers. Results are reported for the 1982-2003 period. GMM estimates are obtained using the measurement error-consistent higher-order moments estimators of Erickson and Whited (2000). The set of perfectly measured regressors includes a constant and normalized cash flow. Robust standard errors for OLS and Newey-McFadden (1994) influence function adjusted standard errors for GMM are reported in parentheses. [a], [b], and [c] denote significance at the 1%, 5%, and 10% levels, respectively. Also reported are test statistics for identification tests and *J*-tests of over-identifying restrictions along with their *p*-values. Instances in which both identifying and over-identifying conditions are satisfied have results reported in grey cells. Number of observations is 7,724 in Panel A and 11,543 in Panel B.

Panel A

	ID Test	J-Test		QE				CF			
		GMM4	GMM5	OLS	GMM3	GMM4	GMM5	OLS	GMM3	GMM4	GMM5
1982	2.73	0.631	6.885	-0.079[a]	-0.1	-0.104	-0.104	1.329[a]	1.332	1.333	1.333
485	(0.26)	(0.729)	(0.229)	(0.014)	(0.005)	(0.002)	(0.002)	(0.006)	(0.003)	(0.003)	(0.003)
1983	0.27	2.551	58.324	0.001	0.006	0.015	0	0.179[a]	0.143	0.069	0.187
497	(0.88)	(0.279)	(0)	(0.005)	(0.147)	(0.022)	(0.012)	(0.065)	(1.215)	(0.201)	(0.122)
1984	1.50	3.62	8.204	0.013	0.066	0.035	0.023	0.01	-0.595	-0.25	-0.112
501	(0.47)	(0.164)	(0.145)	(0.009)	(0.031)	(0.019)	(0.005)	(0.13)	(0.548)	(0.344)	(0.127)
1985	6.34	4.769	7.81	0.013[a]	0.022[a]	0.019	0.01[a]	0.102[c]	0.016	0.047	0.128[a]
509	(0.04)	(0.092)	(0.167)	(0.004)	(0.006)	(0.003)	(0.003)	(0.058)	(0.079)	(0.042)	(0.023)
1986	3.11	1.712	6.982	-0.011	-0.132	-0.073	-0.064	0.591[a]	0.655	0.624	0.619
523	(0.21)	(0.425)	(0.222)	(0.012)	(0.13)	(0.015)	(0.006)	(0.049)	(0.101)	(0.035)	(0.038)
1987	3.15	3.323	10.774	0.012[a]	0.032	0.024	0.029	0.254[a]	0.25	0.252	0.251
545	(0.21)	(0.19)	(0.056)	(0.004)	(0.007)	(0.005)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
1988	4.58	3.234	12.937	-0.013[b]	-0.699	-0.128	-1.589	0.649[a]	1.581	0.804	2.79
560	(0.10)	(0.198)	(0.024)	(0.006)	(3.05)	(0.033)	(9.507)	(0.042)	(4.01)	(0.049)	(12.926)
1989	0.71	0.469	12.019	-0.023[c]	-0.221	-0.231	-0.225	0.574[a]	0.751	0.76	0.755
569	(0.70)	(0.791)	(0.035)	(0.013)	(0.179)	(0.131)	(0.009)	(0.145)	(0.233)	(0.201)	(0.092)
1990	5.76	3.331	8.174	0.003	0.093[c]	0.128[a]	0.059[b]	0.363[a]	0.118	0.022	0.209
582	(0.06)	(0.189)	(0.147)	(0.008)	(0.051)	(0.032)	(0.025)	(0.098)	(0.231)	(0.184)	(0.14)
1991	7.86	2.554	5.267	-0.02[b]	-0.249	-0.087[a]	-0.066[a]	0.668[a]	0.675[a]	0.67[a]	0.67[a]
588	(0.02)	(0.279)	(0.384)	(0.009)	(0.158)	(0.012)	(0.013)	(0.009)	(0.011)	(0.006)	(0.007)
1992	4.50	2.04	3.745	0.006	0.051	0.044	0.024	0.237[a]	-0.066	-0.017	0.119
592	(0.11)	(0.361)	(0.587)	(0.004)	(0.041)	(0.025)	(0.01)	(0.055)	(0.292)	(0.192)	(0.09)
1993	12.66	1.166	5.342	0.01[a]	0.032[a]	0.027[a]	0.034[a]	0.126[a]	0.038	0.058[c]	0.03
592	(0.00)	(0.558)	(0.376)	(0.003)	(0.007)	(0.003)	(0.003)	(0.042)	(0.042)	(0.031)	(0.03)
1994	16.42	2.677	3.676	0.009[b]	0.094	0.026[a]	0.022[a]	0.138[a]	-0.233	0.065[b]	0.08[a]
590	0.00	(0.262)	(0.597)	(0.004)	(0.168)	(0.005)	(0.004)	(0.035)	(0.765)	(0.033)	(0.029)
1995	7.27	2.388	5.026	0.018[a]	0.032[a]	0.053[a]	0.051[a]	0.085[a]	0.043	-0.021	-0.014
591	(0.03)	(0.303)	(0.413)	(0.003)	(0.008)	(0.008)	(0.01)	(0.023)	(0.033)	(0.044)	(0.043)

Table 6, Panel B

	ID Test	J-Test		QE				CF			
		GMM4	GMM5	OLS	GMM3	GMM4	GMM5	OLS	GMM3	GMM4	GMM5
1982	4.37	5.36	6.853	0.091[b]	0.458	0.629	0.461	0.214[a]	0.03	-0.055	0.029
485	(0.11)	(0.069)	(0.232)	(0.043)	(0.179)	(0.096)	(0.122)	(0.068)	(0.101)	(0.1)	(0.092)
1983	5.46	5.69	12.899	0.033	-0.291	-2.284	-0.003	0.253[a]	0.38[b]	1.164	0.267
497	(0.07)	(0.058)	(0.024)	(0.027)	(0.477)	(2.018)	(0.254)	(0.063)	(0.191)	(0.861)	(0.117)
1984	3.72	0.90	2.412	0.084[a]	0.263	0.29	0.277	0.171[a]	0.046	0.027	0.036
501	(0.16)	(0.637)	(0.79)	(0.025)	(0.046)	(0.023)	(0.031)	(0.049)	(0.051)	(0.053)	(0.06)
1985	5.92	2.63	8.01	0.103[a]	0.287[a]	0.313[a]	0.436[a]	0.095[a]	0.018	0.007	-0.045
509	(0.05)	(0.269)	(0.156)	(0.022)	(0.054)	(0.047)	(0.076)	(0.03)	(0.032)	(0.032)	(0.058)
1986	2.33	0.50	11.431	0.025	0.13	0.198	0.116	0.247[a]	0.195	0.161	0.202
523	(0.31)	(0.78)	(0.043)	(0.018)	(0.122)	(0.081)	(0.033)	(0.039)	(0.075)	(0.061)	(0.044)
1987	2.20	1.40	4.824	0.033[b]	0.154	0.316	0.231	0.128[a]	0.052	-0.05	0.003
545	(0.33)	(0.496)	(0.438)	(0.015)	(0.142)	(0.147)	(0.044)	(0.036)	(0.092)	(0.089)	(0.038)
1988	0.38	2.08	4.246	0.027[b]	-0.297	-0.002	-0.111	0.137[a]	0.323	0.153	0.216
560	(0.83)	(0.354)	(0.515)	(0.011)	(1.73)	(0.245)	(0.507)	(0.024)	(0.992)	(0.145)	(0.294)
1989	0.20	1.35	2.934	0.03[b]	1.044	0.069	0.372	0.142[a]	-0.458	0.119	-0.061
569	(0.91)	(0.508)	(0.71)	(0.012)	(13.914)	(0.653)	(0.123)	(0.024)	(8.261)	(0.391)	(0.091)
1990	5.76	1.12	10.12	0.002	0.091	0.014	0.427[b]	0.187[a]	0.111	0.177[a]	-0.176
582	(0.06)	(0.571)	(0.072)	(0.009)	(0.096)	(0.039)	(0.19)	(0.028)	(0.087)	(0.036)	(0.176)
1991	5.49	2.12	7.088	0.052[a]	0.173[a]	0.077	0.061[b]	0.074[a]	0.015	0.062[b]	0.07[a]
588	(0.06)	(0.346)	(0.214)	(0.013)	(0.049)	(0.047)	(0.027)	(0.022)	(0.03)	(0.027)	(0.022)
1992	2.17	4.14	8.664	0.019[b]	0.062	0.001	0.033	0.142[a]	0.111	0.155	0.132
592	(0.34)	(0.126)	(0.123)	(0.008)	(0.046)	(0.07)	(0.028)	(0.021)	(0.04)	(0.056)	(0.027)
1993	7.49	0.89	14.883	0.015[b]	-0.301	-1.148	-0.002	0.126[a]	0.305	0.785	0.136
592	(0.02)	(0.641)	(0.011)	(0.006)	(0.514)	(1.699)	(0.044)	(0.018)	(0.286)	(0.964)	(0.03)
1994	1.64	5.18	6.487	0.024[a]	-0.102	-0.002	-0.086	0.106[a]	0.159	0.117	0.153
590	(0.44)	(0.075)	(0.262)	(0.007)	(0.223)	(0.107)	(0.099)	(0.015)	(0.094)	(0.047)	(0.043)
1995	2.44	1.44	9.786	0.041[a]	0.127	0.158	0.098	0.088[a]	0.068	0.061	0.075
591	(0.30)	(0.487)	(0.082)	(0.013)	(0.124)	(0.08)	(0.054)	(0.023)	(0.034)	(0.029)	(0.027)
1996	9.43	0.12	10.75	0.062[a]	0.201[a]	0.204[a]	0.109	0.006	-0.014	-0.015	-0.001
595	(0.01)	(0.943)	(0.057)	(0.012)	(0.039)	(0.04)	(0.027)	(0.02)	(0.022)	(0.022)	(0.02)
1997	5.70	2.29	5.765	0.046[a]	0.153[a]	0.129[a]	0.144[a]	0.078[a]	0.058[a]	0.062[a]	0.059[a]
571	(0.06)	(0.318)	(0.33)	(0.013)	(0.027)	(0.045)	(0.028)	(0.017)	(0.018)	(0.017)	(0.017)
1998	0.27	2.97	7.262	0.019[b]	0.313	0.12	0.137	0.083[a]	0.032	0.065	0.062
527	(0.87)	(0.227)	(0.202)	(0.008)	(0.927)	(0.148)	(0.08)	(0.023)	(0.15)	(0.038)	(0.031)
1999	3.20	0.83	5.801	0.015[b]	0.049	0.053	0.023	0.083[a]	0.069	0.067	0.079
485	(0.20)	(0.661)	(0.326)	(0.006)	(0.04)	(0.043)	(0.018)	(0.015)	(0.021)	(0.024)	(0.019)
2000	8.25	0.29	5.981	0.041[a]	0.08[a]	0.079[a]	0.068[a]	0.054[a]	0.042[a]	0.042[a]	0.045[a]
439	(0.02)	(0.866)	(0.308)	(0.008)	(0.012)	(0.01)	(0.018)	(0.016)	(0.014)	(0.015)	(0.017)
2001	10.29	2.53	4.099	0.034[a]	0.09[a]	0.054[a]	0.055[a]	0.038[b]	0.016	0.03[b]	0.03[c]
418	(0.01)	(0.282)	(0.535)	(0.008)	(0.026)	(0.017)	(0.013)	(0.017)	(0.015)	(0.015)	(0.016)
2002	6.07	0.24	35.149	0.021[a]	-0.012	-0.002	0	0.051[a]	0.067[a]	0.062[a]	0.061
400	(0.05)	(0.885)	(0)	(0.005)	(0.051)	(0.025)	(0.019)	(0.011)	(0.024)	(0.015)	(0.014)
2003	5.82	1.32	11.354	0.017[a]	-0.111	-0.037	-0.013	0.042[a]	0.077[b]	0.057[a]	0.051
384	(0.06)	(0.517)	(0.045)	(0.005)	(0.143)	(0.043)	(0.027)	(0.01)	(0.039)	(0.015)	(0.014)

TABLE 7

Higher Order Moments Estimation Using Extended Sample of Manufacturing Firms

This table presents OLS and GMM results from estimating

$$\frac{I_t}{K_{t-1}} = a + \beta_1 q_t + \beta_2 \frac{CF_t}{K_{t-1}} + e_t$$

with Q_E used as a mismeasured proxy for q for a sample of all manufacturing firms in COMPUSTAT with at least 5 years of continuous data. Observations in each year are given under the year in italics. Panel A uses original EW assumptions for this extended data and reports results for the 1982-1995 period. Observations with negative Q_E values are omitted. Panel B uses variable definitions that are standard in the literature (see Appendix 1) and reports results for the 1982-2003 period. Investment, cash flow and Q_E are winsorized at the bottom and top percentile to remove the outliers. GMM estimates are obtained using the measurement error-consistent higher-order moments estimators of Erickson and Whited (2000). The set of perfectly measured regressors includes a constant and normalized cash flow. The number of observations in each year is reported in the first column under the corresponding year. Robust standard errors for OLS and Newey-McFadden (1994) influence function adjusted standard errors for GMM are reported in parentheses. [a], [b], and [c] denote significance at the 1%, 5%, and 10% levels, respectively. Also reported are test statistics for identification tests and *J*-tests of over-identifying restrictions along with their *p*-values. Instances in which both identifying and over-identifying conditions are satisfied have results reported in grey cells. Number of observations is 21,742 in Panel A and 37,617 in Panel B.

Panel A

	ID Test	J-Test		QE				CF			
		GMM4	GMM5	OLS	GMM3	GMM4	GMM5	OLS	GMM3	GMM4	GMM5
1982	8.79	1.793	6.539	0.021[a]	0.066[a]	0.07[a]	0.077[a]	0.259[a]	0.055	0.04	0.009
1449	(0.01)	(0.408)	(0.257)	(0.004)	(0.008)	(0.006)	(0.002)	(0.039)	(0.048)	(0.049)	(0.047)
1983	15.87	0.528	12.983	0.012[a]	0.04[a]	0.041[a]	0.06[a]	0.21[a]	0.071	0.067	-0.03
1511	0.00	(0.768)	(0.024)	(0.002)	(0.006)	(0.006)	(0.005)	(0.035)	(0.047)	(0.049)	(0.063)
1984	11.17	7.372	9.748	0.011[a]	0.035[a]	0.029	0.033	0.157[a]	0.067	0.086	0.073
1504	(0.00)	(0.025)	(0.083)	(0.002)	(0.005)	(0.003)	(0.001)	(0.031)	(0.042)	(0.035)	(0.041)
1985	15.05	4.189	11.103	0.015[a]	0.038[a]	0.035[a]	0.037	0.158[a]	0.106[a]	0.115[a]	0.108
1544	(0.00)	(0.123)	(0.049)	(0.003)	(0.004)	(0.003)	(0.002)	(0.03)	(0.039)	(0.036)	(0.038)
1986	10.38	5.145	11.92	0.012[b]	0.07[a]	0.072	0.044	0.352[a]	0.197[c]	0.193	0.267
1546	(0.01)	(0.076)	(0.036)	(0.005)	(0.012)	(0.009)	(0.007)	(0.056)	(0.104)	(0.087)	(0.067)
1987	18.61	2.148	6.163	0.013[a]	0.043[a]	0.04[a]	0.03[a]	0.219[a]	0.167[a]	0.172[a]	0.19[a]
1585	0.00	(0.342)	(0.291)	(0.003)	(0.006)	(0.006)	(0.003)	(0.042)	(0.053)	(0.053)	(0.043)
1988	23.87	2.161	12.621	0.011[a]	0.033[a]	0.031[a]	0.024	0.195[a]	0.169[a]	0.171[a]	0.179
1567	0.00	(0.339)	(0.027)	(0.002)	(0.007)	(0.007)	(0.003)	(0.029)	(0.036)	(0.034)	(0.031)
1989	14.83	0.554	15.583	0.011[a]	0.032[a]	0.031[a]	0.026	0.151[a]	0.11[a]	0.111[a]	0.122
1546	(0.00)	(0.758)	(0.008)	(0.002)	(0.004)	(0.004)	(0.003)	(0.021)	(0.028)	(0.029)	(0.023)
1990	11.99	6.439	11.651	0.009[a]	0.048[a]	0.038	0.038	0.231[a]	0.207[a]	0.213	0.213
1600	(0.00)	(0.04)	(0.04)	(0.002)	(0.007)	(0.005)	(0.005)	(0.037)	(0.051)	(0.045)	(0.044)
1991	19.53	1.468	12.477	0.013[a]	0.037[a]	0.037[a]	0.031	0.144[a]	0.111[a]	0.111[a]	0.12
1682	0.00	(0.48)	(0.029)	(0.002)	(0.004)	(0.004)	(0.003)	(0.026)	(0.033)	(0.034)	(0.028)
1992	35.47	13.898	31.839	0.008[a]	0.014[a]	0.021	0.031	0.027[b]	0.035[b]	0.044	0.056
1641	0.00	(0.001)	(0)	(0.001)	(0.001)	(0.002)	(0.003)	(0.013)	(0.016)	(0.024)	(0.034)
1993	29.73	8.372	24.95	0.008[a]	0.014[a]	0.018	0.016	0.049[a]	0.069[a]	0.082	0.077
1573	0.00	(0.015)	(0)	(0.001)	(0.001)	(0.001)	(0.001)	(0.011)	(0.013)	(0.018)	(0.015)
1994	31.83	16.507	30.024	0.008[a]	0.013[a]	0.021	0.025	0.054[a]	0.068[a]	0.092	0.102
1524	0.00	(0)	(0)	(0.001)	(0.001)	(0.002)	(0.002)	(0.011)	(0.013)	(0.022)	(0.028)
1995	17.31	10.004	16.981	0.012[a]	0.024[a]	0.023	0.044	0.074[a]	0.051[a]	0.052	0.014
1470	0.00	(0.007)	(0.005)	(0.001)	(0.002)	(0.002)	(0.006)	(0.013)	(0.015)	(0.015)	(0.033)

Table 7, Panel B

	ID Test	J-Test		QE				CF			
		GMM4	GMM5	OLS	GMM3	GMM4	GMM5	OLS	GMM3	GMM4	GMM5
1982	19.90	3.69	5.56	0.096[a]	0.407[a]	0.483[a]	0.372[a]	0.173[a]	0.047	0.016	0.061[c]
1449	(0.00)	(0.158)	(0.352)	(0.014)	(0.076)	(0.104)	(0.052)	(0.022)	(0.036)	(0.049)	(0.034)
1983	16.62	2.29	4.877	0.084[a]	0.189[a]	0.183[a]	0.2[a]	0.159[a]	0.117[a]	0.119[a]	0.113[a]
1511	(0.00)	(0.318)	(0.431)	(0.011)	(0.019)	(0.017)	(0.013)	(0.022)	(0.023)	(0.023)	(0.024)
1984	14.01	6.60	9.86	0.084[a]	0.306[a]	0.341	0.348	0.104[a]	0.068[a]	0.062	0.061
1504	(0.00)	(0.037)	(0.079)	(0.012)	(0.046)	(0.048)	(0.04)	(0.019)	(0.024)	(0.026)	(0.028)
1985	26.95	14.55	34.24	0.108[a]	0.245[a]	0.273	0.242	0.082[a]	0.088[a]	0.089	0.088
1544	(0.00)	(0.001)	(0)	(0.012)	(0.024)	(0.02)	(0.024)	(0.013)	(0.014)	(0.015)	(0.014)
1986	19.06	2.53	10.67	0.095[a]	0.27[a]	0.276[a]	0.2	0.11[a]	0.113[a]	0.113[a]	0.112
1546	(0.00)	(0.282)	(0.058)	(0.013)	(0.031)	(0.029)	(0.026)	(0.02)	(0.023)	(0.023)	(0.021)
1987	19.88	1.73	16.402	0.069[a]	0.192[a]	0.153[a]	0.085	0.061[a]	0.076[a]	0.072[a]	0.063
1585	(0.00)	(0.421)	(0.006)	(0.011)	(0.032)	(0.022)	(0.016)	(0.021)	(0.021)	(0.021)	(0.02)
1988	16.61	2.71	16.01	0.078[a]	0.202[a]	0.191[a]	0.38	0.053[a]	0.05[a]	0.05[a]	0.046
1567	(0.00)	(0.258)	(0.007)	(0.011)	(0.017)	(0.013)	(0.075)	(0.011)	(0.012)	(0.012)	(0.02)
1989	14.82	0.82	19.335	0.054[a]	0.188[a]	0.168[a]	0.567	0.056[a]	0.075[a]	0.072[a]	0.129
1546	(0.00)	(0.665)	(0.002)	(0.01)	(0.039)	(0.03)	(0.233)	(0.012)	(0.016)	(0.015)	(0.055)
1990	14.57	2.14	10.85	0.039[a]	0.16[a]	0.198[a]	0.147	0.043[a]	0.076[a]	0.086[a]	0.072
1600	(0.00)	(0.343)	(0.054)	(0.007)	(0.031)	(0.037)	(0.027)	(0.007)	(0.016)	(0.02)	(0.014)
1991	20.87	2.43	3.165	0.073[a]	0.233[a]	0.274[a]	0.255	0.011	0.041[a]	0.048[a]	0.045
1682	(0.00)	(0.297)	(0.675)	(0.01)	(0.031)	(0.039)	(0.031)	(0.009)	(0.012)	(0.014)	(0.012)
1992	17.43	1.90	3.15	0.075[a]	0.165[a]	0.168[a]	0.164	-0.005	0.013	0.014	0.013
1641	(0.00)	(0.386)	(0.677)	(0.008)	(0.014)	(0.011)	(0.01)	(0.011)	(0.013)	(0.013)	(0.013)
1993	23.15	4.94	16.07	0.077[a]	0.313[a]	0.25	0.233	0.002	0.048[a]	0.036	0.032
1573	(0.00)	(0.084)	(0.007)	(0.01)	(0.041)	(0.041)	(0.037)	(0.012)	(0.013)	(0.014)	(0.013)
1994	23.72	0.59	15.42	0.065[a]	0.242[a]	0.263[a]	0.187	0.02[b]	0.06[a]	0.065[a]	0.048
1524	(0.00)	(0.744)	(0.009)	(0.008)	(0.038)	(0.037)	(0.025)	(0.009)	(0.014)	(0.015)	(0.011)
1995	16.45	3.06	19.409	0.07[a]	0.269[a]	0.35[a]	0.23	0.052[a]	0.063[a]	0.067[a]	0.061
1470	(0.00)	(0.217)	(0.002)	(0.011)	(0.038)	(0.051)	(0.041)	(0.009)	(0.011)	(0.014)	(0.01)
1996	25.51	6.13	14.44	0.07[a]	0.163[a]	0.131	0.207	-0.008	0.008	0.002	0.015
1928	(0.00)	(0.047)	(0.013)	(0.008)	(0.022)	(0.019)	(0.017)	(0.006)	(0.007)	(0.006)	(0.007)
1997	20.44	5.06	8.282	0.052[a]	0.162[a]	0.148	0.167[a]	-0.005[c]	0.009[a]	0.007	0.009[c]
2042	(0.00)	(0.08)	(0.141)	(0.006)	(0.024)	(0.02)	(0.018)	(0.003)	(0.005)	(0.005)	(0.005)
1998	18.12	2.76	4.63	0.045[a]	0.336[a]	0.301[a]	0.26[a]	-0.002	0.051[a]	0.044[a]	0.037[a]
2078	(0.00)	(0.252)	(0.463)	(0.006)	(0.083)	(0.056)	(0.037)	(0.005)	(0.015)	(0.011)	(0.007)
1999	20.12	0.49	4.658	0.042[a]	0.177[a]	0.183[a]	0.206[a]	0.002	0.028[a]	0.029[a]	0.034[a]
2173	(0.00)	(0.781)	(0.459)	(0.006)	(0.029)	(0.026)	(0.022)	(0.004)	(0.007)	(0.007)	(0.006)
2000	21.52	0.02	3.98	0.058[a]	0.18[a]	0.179[a]	0.186[a]	-0.002	0.025[a]	0.025[a]	0.026[a]
2058	(0.00)	(0.988)	(0.552)	(0.007)	(0.021)	(0.018)	(0.017)	(0.005)	(0.007)	(0.007)	(0.006)
2001	8.16	12.44	12.902	0.021[a]	0.124[a]	0.151	0.122	-0.006[b]	0.013	0.018	0.013
1955	(0.02)	(0.002)	(0.024)	(0.005)	(0.03)	(0.047)	(0.017)	(0.003)	(0.008)	(0.011)	(0.005)
2002	9.85	7.66	8.88	0.022[a]	0.273	0.149	0.132	-0.001	0.056	0.028	0.024[a]
1864	(0.01)	(0.022)	(0.114)	(0.005)	(0.186)	(0.042)	(0.027)	(0.002)	(0.041)	(0.01)	(0.006)
2003	2.95	6.65	13.422	0.017[b]	0.114	0.124	0.121	0	0.007	0.007	0.007
1774	(0.23)	(0.036)	(0.02)	(0.007)	(0.055)	(0.028)	(0.024)	(0.001)	(0.004)	(0.003)	(0.002)

TABLE 8
Dynamic Model Estimation Using Lagged Measures of q as Instruments

This table reports results from estimating ρ, β_1, and β_2 in the restricted model

$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = \rho\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + \beta_1\Delta q_{it} - \rho\beta_1\Delta q_{it-1} + \beta_2\Delta\left(\frac{CF_{it}}{K_{it-1}}\right) - \rho\beta_2\Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it}$$

from the unrestricted model

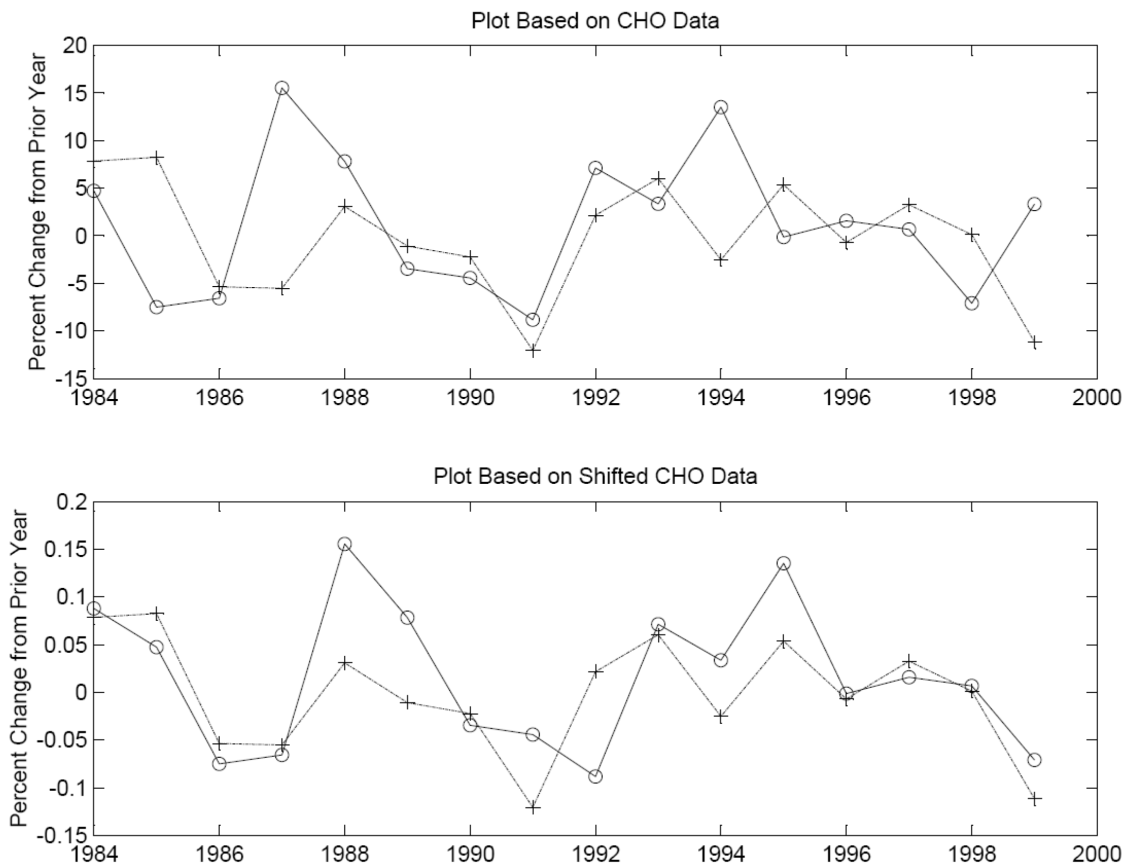
$$\Delta\left(\frac{I_{it}}{K_{it-1}}\right) = b_0\Delta\left(\frac{I_{it-1}}{K_{it-2}}\right) + b_1\Delta q_{it} + b_2\Delta q_{it-1} + b_3\Delta\left(\frac{CF_{it}}{K_{it-1}}\right) + b_4\Delta\left(\frac{CF_{it-1}}{K_{it-2}}\right) + YEAR^*_t + \zeta^*_{it}$$

as in Blundell, Bond, Devereaux, and Schiantarelli (1992). The sample period is 1982-2003. Financial and utility firms are excluded. Negative cash flow observations are excluded. Q_E and/or $\hat{Q}$ are used as proxies for q. The subscripts i and t indicate firm and year, respectively, in the data panel. The sample in Panels A and B consists of 1223 firms and 7188 observations, while that in Panels C, D, and E consists of 1547 firms and 11829 observations. Panel A reports results for the Anderson and Hsiao (1981) instrumental variables estimator. Panel B reports results for the mixed model of Arellano and Bond (1991, 1998) that uses time-disaggregated moment conditions for lagged investment in levels and aggregated ones for the other instruments in first-differences. Panels C, D, and E report results for the complete panel data GMM estimator of Arellano and Bond (1991, 1998). Panels A, B, and C, and the top row in Panel E use 3rd and 4th lags of normalized investment, cash flow, and q as instruments. Panel D and the bottom row in Panel E use all available lags of length greater than 2 of normalized investment, cash flow, and q as instruments. Robust standard errors are reported in parentheses. [a], [b], and [c] denote significance at the 1%, 5%, and 10% levels, respectively. Also reported are p-values for J-tests for over-identifying conditions, m_2 tests for second order serial correlation as in Arellano and Bond (1991), and the BBDS comfac test of common factor restrictions.

	Q_E	$\hat{Q}$	CF/K	I/K	J	m_2	*Comfac*
Instruments							
Panel A							
$\Delta(I/K)_{t-i}$, $\Delta(CF/K)_{t-i}$, ΔQ_{Et-i} $i=3,4$	0.074[a] (0.027)		-0.211 (0.131)	-0.338[a] (0.053)	0.336	0.906	0.889
$\Delta(I/K)_{t-i}$, $\Delta(CF/K)_{t-i}$, $\Delta\hat{Q}_{t-i}$ $i=3,4$	-0.035 (0.077)		-0.194 (0.151)	0.396[a] (0.086)	0.700	0.323	0.833
$\Delta(I/K)_{t-i}$, $\Delta(CF/K)_{t-i}$, $\Delta\hat{Q}_{t-i}$ $i=3,4$		0.069 (0.128)	-0.011 (0.105)	-0.114[a] (0.021)	0.821	0.688	0.975
$\Delta(I/K)_{t-i}$, $\Delta(CF/K)_{t-i}$, ΔQ_{Et-i} $i=3,4$		0.377 (0.268)	-0.025 (0.289)	-0.417[c] (0.219)	0.801	0.558	0.558
Panel B							
$(I/K)_{t-i}$, $\Delta(CF/K)_{t-i}$, ΔQ_{Et-i} $i=3,4$	0.060[a] (0.020)		0.198[a] (0.061)	0.230[a] (0.027)	0.025	0.770	0.557
$(I/K)_{t-i}$, $\Delta(CF/K)_{t-i}$, $\Delta\hat{Q}_{t-i}$ $i=3,4$	0.067[a] (0.025)		0.180[a] (0.061)	0.147[a] (0.017)	0.010	0.480	0.784
$(I/K)_{t-i}$, $\Delta(CF/K)_{t-i}$, $\Delta\hat{Q}_{t-i}$ $i=3,4$		0.012 (0.043)	0.218[a] (0.065)	0.164[a] (0.015)	0.011	0.713	0.715
$(I/K)_{t-i}$, $\Delta(CF/K)_{t-i}$, ΔQ_{Et-i} $i=3,4$		0.038 (0.042)	0.239[a] (0.067)	0.177[a] (0.017)	0.009	0.619	0.724
Panel C							
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, Q_{Et-i} $i=3,4$	0.047[a] (0.010)		0.107[a] (0.034)	0.156[a] (0.023)	0.079	0.197	0.479
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, $\hat{Q}_{t-i}$ $i=3,4$	0.077[a] (0.017)		0.129[a] (0.039)	0.280[a] (0.043)	0.192	0.894	0.184
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, $\hat{Q}_{t-i}$ $i=3,4$		-0.009 (0.020)	0.122[a] (0.035)	0.212[a] (0.018)	0.103	0.173	0.215
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, Q_{Et-i} $i=3,4$		0.025 (0.033)	0.122[a] (0.033)	0.148[a] (0.019)	0.043	0.097	0.471
Panel D							
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, Q_{Et-i} $i=3,4,\ldots,t-1$	0.053[a] (0.008)		0.083[a] (0.021)	0.212[a] (0.028)	0.136	0.302	0.119
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, $\hat{Q}_{t-i}$ $i=3,4,\ldots,t-1$	0.056[a] (0.010)		0.075[a] (0.022)	0.240[a] (0.030)	0.157	0.574	0.099
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, $\hat{Q}_{t-i}$ $i=3,4,\ldots,t-1$		0.014 (0.012)	0.102[a] (0.022)	0.266[a] (0.019)	0.079	0.193	0.005
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, Q_{Et-i} $i=3,4,\ldots,t-1$		0.048[a] (0.015)	0.103[a] (0.022)	0.226[a] (0.023)	0.240	0.104	0.053
Panel E							
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, Q_{Et-i}, $\hat{Q}_{t-i}$ $i=3,4$	0.037[a] (0.011)	-0.026 (0.018)	0.100[a] (0.032)	0.204[a] (0.022)	0.343	0.228	0.349
$(I/K)_{t-i}$, $(CF/K)_{t-i}$, Q_{Et-i}, $\hat{Q}_{t-i}$ $i=3,4,\ldots,t-1$	0.047[a] (0.008)	-0.005 (0.012)	0.086[a] (0.020)	0.250[a] (0.023)	0.295	0.296	0.035

FIGURE 1
Annual Changes in Aggregate Investment and Cash Flow
Using Unshifted and Shifted CHO Data

The plots show annual percent changes from previous years in aggregate investment (+) and cash flow (o) for the Cummins, Hassett, and Oliner (2006) sample of firms. The top panel uses investment and cash flow series as reported in their data set, while the bottom panel uses the cash flow series shifted forward by one year.



APPENDIX

A1

We access raw accounting and stock market data from the COMPUSTAT database. These are then combined to yield the regression variables of interest using the following definitions that are standard in the literature:

Capital stock (K) = Net property, plant, and equipment (DATA8)

Investment (I) = Capital expenditure (DATA128)

Q_E = (Total assets (DATA6) + Year-end share price (DATA25)*Number of shares outstanding (DATA24) – Book value of equity (DATA60))/Total assets

Cash flow (CF) = Net income before extraordinary items (DATA18) + Depreciation and amortization (DATA14).

Investment and cash flow are scaled by beginning-of-period capital stock. Q_E is measured at the beginning of the period.

A2

This table gives two examples that income before extraordinary items and finished goods inventories are reverse sorted in the original EW dataset. The reverse sort is detected by matching EW sample with COMPUSTAT with all variables for 595 firms. The columns named COMPUSTAT gives the values from COMPUSTAT dataset and the columns named EW gives the values from the original EW dataset. The first three values from COMPUSTAT are highlighted in both the COMPUSTAT and EW sample to highlight the reverse sort.

						Income Before Extraordinary Items		Finished Goods Inventories	
	CNUM	GVKEY	SIC	CUSIP	YEAR	COMPUSTAT	EW	COMPUSTAT	EW
Example 1	002824	001078	2834	002824100	1977	**117.838**	1688.7	**101.624**	560.637
	002824	001078	2834	002824100	1978	**148.626**	1516.68	**122.876**	514.715
	002824	001078	2834	002824100	1979	**178.981**	1399.13	**149.635**	476.548
	002824	001078	2834	002824100	1980	214.413	1239.06	200.6	421.932
	002824	001078	2834	002824100	1981	247.283	1088.68	225.857	406.026
	002824	001078	2834	002824100	1982	289.123	965.774	226.704	405.727
	002824	001078	2834	002824100	1983	347.617	859.832	240.007	343.367
	002824	001078	2834	002824100	1984	402.575	752.027	230.142	306.874
	002824	001078	2834	002824100	1985	465.335	632.559	240.201	321.509
	002824	001078	2834	002824100	1986	540.46	540.46	272.334	272.334
	002824	001078	2834	002824100	1987	632.559	465.335	321.509	240.201
	002824	001078	2834	002824100	1988	752.027	402.575	306.874	230.142
	002824	001078	2834	002824100	1989	859.832	347.617	343.367	240.007
	002824	001078	2834	002824100	1990	965.774	289.123	405.727	226.704
	002824	001078	2834	002824100	1991	1088.677	247.283	406.026	225.857
	002824	001078	2834	002824100	1992	1239.057	214.413	421.932	200.6
	002824	001078	2834	002824100	1993	1399.126	**178.981**	476.548	**149.635**
	002824	001078	2834	002824100	1994	1516.683	**148.626**	514.715	**122.876**
	002824	001078	2834	002824100	1995	1688.7	**117.838**	560.637	**101.624**

	CNUM	GVKEY	SIC	CUSIP	YEAR	Income Before Extraordinary Items		Finished Goods Inventories	
						COMPUSTAT	EW	COMPUSTAT	EW
Example 2	004644	001099	3612	004644100	1977	**2.092**	0.992	**2.416**	5.543
	004644	001099	3612	004644100	1978	**2.985**	-5.659	**2.802**	3.793
	004644	001099	3612	004644100	1979	**1.985**	-0.673	**3.218**	3.483
	004644	001099	3612	004644100	1980	2.834	0.624	4.627	4.159
	004644	001099	3612	004644100	1981	3.054	-7.604	4.223	4.965
	004644	001099	3612	004644100	1982	2.386	3.69	4.663	9.034
	004644	001099	3612	004644100	1983	0.126	3.482	3.365	7.15
	004644	001099	3612	004644100	1984	2.26	0.759	3.519	4.752
	004644	001099	3612	004644100	1985	2.651	0.663	4.015	3.622
	004644	001099	3612	004644100	1986	1.883	1.883	4.062	4.062
	004644	001099	3612	004644100	1987	0.663	2.651	3.622	4.015
	004644	001099	3612	004644100	1988	0.759	2.26	4.752	3.519
	004644	001099	3612	004644100	1989	3.482	0.126	7.15	3.365
	004644	001099	3612	004644100	1990	3.69	2.386	9.034	4.663
	004644	001099	3612	004644100	1991	-7.604	3.054	4.965	4.223
	004644	001099	3612	004644100	1992	0.624	2.834	4.159	4.627
	004644	001099	3612	004644100	1993	-0.673	**1.985**	3.483	**3.218**
	004644	001099	3612	004644100	1994	-5.659	**2.985**	3.793	**2.802**
	004644	001099	3612	004644100	1995	0.992	**2.092**	5.543	**2.416**

A3

Robustness check for Table 7, Panel B. Negative cash flows excluded. Number of observations is 29,896.

	ID Test	J-Test		QE				CF			
		GMM4	GMM5	OLS	GMM3	GMM4	GMM5	OLS	GMM3	GMM4	GMM5
1982	7.41	2.14	2.50	0.085[a]	0.413[a]	0.376[a]	0.374[a]	0.225[a]	-0.007	0.02	0.021
1292	(0.03)	(0.343)	(0.777)	(0.017)	(0.141)	(0.066)	(0.05)	(0.031)	(0.102)	(0.061)	(0.056)
1983	10.48	2.45	5.986	0.077[a]	0.172[a]	0.198[a]	0.199[a]	0.21[a]	0.148[a]	0.131[a]	0.131[a]
1341	(0.01)	(0.294)	(0.308)	(0.013)	(0.026)	(0.018)	(0.011)	(0.027)	(0.032)	(0.031)	(0.028)
1984	5.88	1.20	7.53	0.053[a]	0.311[c]	0.329[b]	0.308[a]	0.242[a]	0.013	-0.002	0.016
1322	(0.05)	(0.548)	(0.184)	(0.013)	(0.167)	(0.15)	(0.032)	(0.03)	(0.145)	(0.133)	(0.048)
1985	6.60	8.86	16.23	0.066[a]	0.146	0.001	0.154	0.218[a]	0.171[a]	0.258	0.166
1291	(0.04)	(0.012)	(0.006)	(0.012)	(0.107)	(0.097)	(0.053)	(0.026)	(0.065)	(0.058)	(0.039)
1986	4.13	8.45	8.92	0.047[a]	0.146	0.249	0.244	0.243[a]	0.178[a]	0.11	0.114
1299	(0.13)	(0.015)	(0.112)	(0.018)	(0.095)	(0.034)	(0.032)	(0.033)	(0.062)	(0.046)	(0.043)
1987	6.70	7.57	13.483	0.027[c]	0.049	0.176	0.183	0.24[a]	0.224[a]	0.133	0.128
1327	(0.04)	(0.023)	(0.019)	(0.015)	(0.086)	(0.04)	(0.041)	(0.035)	(0.067)	(0.046)	(0.043)
1988	4.99	0.69	14.20	0.047[a]	0.179[b]	0.145[a]	0.139	0.142[a]	0.071[c]	0.089[a]	0.093
1311	(0.08)	(0.709)	(0.014)	(0.01)	(0.07)	(0.023)	(0.021)	(0.017)	(0.038)	(0.021)	(0.019)
1989	2.77	0.77	7.044	0.035[a]	0.395	0.248	0.35	0.148[a]	-0.013	0.053	0.007
1270	(0.25)	(0.68)	(0.217)	(0.012)	(0.32)	(0.085)	(0.113)	(0.019)	(0.145)	(0.047)	(0.062)
1990	1.02	3.22	12.00	0.02[c]	0.065	0.103	0.054	0.151[a]	0.126	0.105	0.132
1304	(0.60)	(0.2)	(0.035)	(0.011)	(0.063)	(0.03)	(0.012)	(0.02)	(0.038)	(0.026)	(0.019)
1991	3.21	0.51	7.141	0.039[a]	0.208	0.195	0.319	0.12[a]	0.027	0.034	-0.034
1338	(0.20)	(0.774)	(0.21)	(0.011)	(0.117)	(0.049)	(0.072)	(0.019)	(0.066)	(0.037)	(0.06)
1992	7.22	2.46	6.00	0.046[a]	0.119[a]	0.154[a]	0.16[a]	0.124[a]	0.069[b]	0.043	0.039
1349	(0.03)	(0.293)	(0.307)	(0.01)	(0.04)	(0.021)	(0.017)	(0.022)	(0.029)	(0.027)	(0.024)
1993	2.93	15.86	17.339	0.019[b]	-0.254	0	-0.427	0.179[a]	0.329	0.19	0.424
1291	(0.23)	(0)	(0.004)	(0.008)	(0.419)	(0.193)	(0.179)	(0.026)	(0.225)	(0.109)	(0.107)
1994	4.22	4.39	7.69	0.039[a]	0.146	0.251	0.202	0.137[a]	0.088	0.041	0.063
1302	(0.12)	(0.112)	(0.174)	(0.01)	(0.089)	(0.066)	(0.04)	(0.016)	(0.043)	(0.038)	(0.027)
1995	1.29	5.51	7.271	0.039[a]	0.325	0.287	0.413	0.159[a]	0.049	0.064	0.015
1263	(0.52)	(0.064)	(0.201)	(0.014)	(0.216)	(0.053)	(0.1)	(0.019)	(0.087)	(0.03)	(0.051)
1996	1.73	3.01	4.71	0.057[a]	0.138	0.222	0.29	0.145[a]	0.101	0.055	0.018
1544	(0.42)	(0.222)	(0.452)	(0.012)	(0.065)	(0.038)	(0.047)	(0.02)	(0.038)	(0.031)	(0.042)
1997	8.71	3.40	12.899	0.049[a]	0.074[a]	0.101[a]	0.087	0.104[a]	0.093[a]	0.081[a]	0.087
1592	(0.01)	(0.183)	(0.024)	(0.007)	(0.027)	(0.018)	(0.015)	(0.013)	(0.015)	(0.015)	(0.014)
1998	0.95	9.25	12.27	0.018[b]	0.24	0.179	0.241	0.158[a]	0.061	0.087	0.061
1564	(0.62)	(0.01)	(0.031)	(0.007)	(0.344)	(0.06)	(0.065)	(0.016)	(0.148)	(0.034)	(0.041)
1999	7.42	8.64	8.386	0.031[a]	0.076	0.001	0.284[a]	0.107[a]	0.087[a]	0.12	-0.004
1609	(0.02)	(0.013)	(0.136)	(0.007)	(0.073)	(0.056)	(0.04)	(0.012)	(0.033)	(0.023)	(0.031)
2000	5.80	3.03	8.99	0.049[a]	0.114[a]	0.079[c]	0.105[a]	0.094[a]	0.064[a]	0.08[a]	0.068[a]
1503	(0.06)	(0.22)	(0.109)	(0.009)	(0.029)	(0.043)	(0.013)	(0.016)	(0.018)	(0.019)	(0.017)
2001	7.17	1.20	9.669	0.032[a]	0.119[a]	0.113[a]	0.071	0.105[a]	0.055[b]	0.059	0.082
1287	(0.03)	(0.55)	(0.085)	(0.008)	(0.042)	(0.023)	(0.02)	(0.015)	(0.025)	(0.023)	(0.017)
2002	3.27	3.02	12.66	0.024[a]	0.017	0.175	0.001	0.121[a]	0.124	0.055	0.131
1244	(0.20)	(0.221)	(0.027)	(0.008)	(0.082)	(0.057)	(0.201)	(0.018)	(0.038)	(0.039)	(0.085)
2003	2.40	10.80	7.745	0.016[c]	-0.077	-0.001	-0.282	0.066[a]	0.088	0.07	0.136
1253	(0.30)	(0.005)	(0.171)	(0.009)	(0.222)	(0.151)	(0.121)	(0.012)	(0.05)	(0.035)	(0.031)