Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2018)
181, *Part* 3, *pp.* 689–716

On solving endogeneity with invalid instruments: an application to investment equations

Antonio F. Galvao,

University of Arizona, Tucson, USA

Gabriel Montes–Rojas,

Consejo Nacional de Investigaciones Científicas y Tecnicas–Universidad de Buenos Aires, Argentina

Jose Olmo

University of Southampton, UK

and Suyong Song

University of Iowa, Iowa City, USA

[Received July 2016. Revised July 2017]

Summary. Regression models relating investment demand with firms' Tobin's q and cash flow are fraught with measurement errors which, in turn, cause endogeneity bias. We propose an alternative solution to this problem based on modelling the interaction between the endogenous Tobin's q and the error term in the investment equation as a function of lagged values of Tobin's q. We then study the identification conditions and asymptotic properties of the resulting estimator. Our analysis of a panel of US firms reveals a larger effect of Tobin's q on firms' investment demand than that obtained by using available estimators in the literature. Moreover, the estimates highlight the importance of cash flow. We find mixed evidence on the relationship between investment demand and firms' cash flow with respect to different measures of financial constraints. Nevertheless, this evidence is more supportive of the view that firms' cash flows have a weaker correlation to investment demand when financial conditions tighten.

Keywords: Cash flow; Endogeneity; Investment equation; Measurement errors; Tobin's q

1. Introduction

Investment theory suggests that a correct measure for firms' investment demand is *marginal* Tobin's q. Fazzari *et al.* (1988) developed estimators for the investment equation model, where a firm's investment is regressed on a proxy for investment demand (*average* Tobin's q) and cash flow. Following them, investment–cash flow sensitivities became a standard metric in the literature to examine the effect of financing imperfections on corporate investment (Stein, 2003). These empirical sensitivities are also used for drawing inferences about efficiency in internal capital markets (Lamont, 1999; Shin and Stulz, 1998), the effect of agency on corporate spending (Hadlock, 1998; Bertrand and Mullainathan, 2005), the role of business groups in cap-

Address for correspondence: Jose Olmo, Department of Economics, University of Southampton, Highfield Lane, Southampton, SO17 1BH, UK.
E-mail: j.b.olmo@soton.ac.uk

 0964–1998/18/181689

ital allocation (Hoshi *et al.*, 1991) and the effect of managerial characteristics on corporate policies (Bertrand and Schoar, 2003). A related influential literature assesses the effect of financial frictions on economic growth. See Bernanke and Gertler (1989), Holmstrom and Tirole (1998), Kind and Levine (1993) and Kiyotaki and Moore (1997), among many others. Nevertheless, empirical models proposed to assess the sensitivity of investment demand to firm characteristics are usually fraught with the presence of measurement error. A typical example is the use of the average Tobin's q for describing the investment–capital ratio or the choice of proxy variables for capturing financial frictions (Hayashi, 1982). The introduction of error when measuring these variables causes endogeneity bias in least squares estimators and leads to erroneous interpretations of the effect of firm characteristics on investment demand. Thus, Poterba (1988) introduced the idea that errors in measuring Tobin's q may be responsible for the observed investment–cash flow sensitivities. If cash flow were correlated with investment opportunities, which is not well measured by a proxy for the marginal Tobin's q, investment–cash flow sensitivities could arise. This argument minimizes the role of financing constraints in determining the relationship between firms' cash flow and investment.

Many studies intend to control for the measurement error in Tobin's q, while analysing the relationship between investment and cash flow. A common approach has been to use the instrumental variables (IVs) method together with ordinary least squares (OLS) and generalized method-of-moments (GMM) estimators to correct the endogeneity problem (see, for example, Almeida *et al.* (2010) and Lewellen and Lewellen (2016)). In this strand of literature, lags of the observed Tobin's q are used as instruments, by assuming that they are uncorrelated with the error term in the regression equation. Specifically, Almeida *et al.* (2010) employed GMM methods using lagged Tobin's q as instruments for the measurement error problem. Their results show the importance of both Tobin's q and cash flow in investment equation models.

A related method in the literature is to use different proxies for the marginal Tobin's q as alternative instruments. For instance, Cummins *et al.* (2006) found no evidence that cash flow is a statistically significant determinant of investment in US companies. Agca and Mozumdar (2017), in contrast, found that cash flow is a statistically significant cause of investment and investment–cash flow sensitivity is higher for financially constrained firms. (An alternative solution relies on the high order moments. This body of the literature addresses the issue by developing measurement error consistent GMM estimators based on the third and higher order moments of the joint distribution of the observed variables (see, for example, Erickson and Whited (2000)). The high order methods, however, rely on very strong conditions on unobservables and provide unstable and biased coefficient estimates in the presence of fixed effects under heteroscedasticity, or in the absence of a high degree of skewness in the data.)

These methods for correcting measurement errors have several shortcomings for the specific problem of estimating investment equation models. In fact, IV-based estimates might be invalid in the presence of error persistence in the structural investment regression equation, which is very likely if the error term contains an auto-correlated factor. In this case, the persistent measurement error induces correlation between the IV and the error term. Hence, empirically, the standard IV corrections of the measurement error problem are inappropriate for obtaining consistent estimates of the structural parameters in the investment equation.

The main contribution of this paper is to suggest an alternative solution to the measurement error problem in investment models that explicitly exploits persistence in the error term. By doing so we develop an econometric methodology that is suitable for assessing the effect of firm characteristics measured with error on investment demand. Our solution is based on modelling the joint interaction of the endogenous variable, e.g. the average Tobin's q, and the error term as a function of polynomials of the lags of Tobin's q. The framework allows for situations in which no valid standard instruments are available, but there are additional variables that are related to the joint interaction of the endogenous variable and the unobserved causes of the dependent variable. These additional variables are defined as simultaneous variables. The intuition of the main identification condition is that, by using the restriction proposed, the researcher can approximate the endogeneity bias by using the simultaneous variables. We state sufficient conditions on the primitives for the identification of regression coefficients. Motivated by this identification result, we suggest an estimator of the structural parameters based on moment conditions that arise from the use of the structural investment equation and an additional equation proposed under our correction method. We also derive consistency and asymptotic normality of the estimator and develop inference procedures.

We apply this methodology to a panel of US firms over the period 1974–2010 and observe large differences across estimates of the effects of Tobin's average q and firms' cash flow. Our empirical findings invalidate the use of OLS and IV estimators due to serial persistence in the error term of the investment regression equation. In contrast, our novel estimation procedure reports estimates of Tobin's average q-coefficient that are significantly larger than the OLS, IV and GMM counterparts. The parameter that is associated with firms' cash flow is also statistically significant, suggesting that firms' cash flow adds relevant information beyond that provided by Tobin's q for describing firms' investment demand.

To obtain deeper insights into the role of cash flow and the relationship with financial constraints, we also classify firms into constrained and unconstrained by using several of the criteria that were proposed in Almeida *et al.* (2004), Moyen (2004) and Hadlock and Pierce (2010). The results of our empirical analysis provide mixed evidence on the relationship between cash flow and financial constraints that, nevertheless, is consistent with the existing literature on the role of cash flow sensitivities in investment demand. More specifically, we find a higher sensitivity of investment to cash flow for financially constrained firms, as characterized by smaller and younger firms. These results are consistent with Fazzari *et al.* (1988), Almeida *et al.* (2004) and Hadlock and Pierce (2010). Interestingly, we find the opposite result, which is consistent with Kaplan and Zingales (1997), when financial constraints are characterized by variables that are related to firms' pay-out ratios and dividends. An additional robustness exercise consists in analysing the effect of a credit supply shock on the relationship between investment demand and firms' cash flow. This analysis shows that firms' cash flow has a weaker correlation to investment demand when financial conditions tighten.

The paper is organized as follows. Section 2 overviews the measurement error problem in investment models. Section 3 presents our solution to correct for endogeneity, derives a feasible estimator and develops inference. Section 4 presents empirical evidence on the drivers of firms' investment by using a panel of US firms. Finally, Section 5 concludes the paper. Appendix A contains technical proofs and a discussion on why the IV estimator is inconsistent in the presence of measurement errors' serial persistence.

The data that are analysed in the paper and the programs that were used to analyse them can be obtained from

`http://wileyonlinelibrary.com/journal/rss-datasets`

2. Measurement errors and endogeneity

In this section, we discuss why measurement errors on the marginal Tobin's q are common, and why conventional IV methods cannot control for them under persistent measurement errors.

2.1. Measurement errors on marginal Tobin's q

The theory suggests that the correct measure for a firm's investment demand is captured by the marginal Tobin's q. This measure stems from the relationship that equates firms' marginal benefit with marginal cost in equilibrium. (We refer readers to Abel and Eberly (1994) and Erickson and Whited (2000) for a discussion on a microfounded model based on the neoclassical theory of investment that helps in the motivation of the relationship between Tobin's q and firms' investment demand.) Nevertheless, the presence of financial constraints may distort this relationship by introducing other factors that influence the firm's optimal investment level. More specifically, financial constraints create a wedge between internal and external funding that invalidates theoretical arguments in the spirit of Modigliani and Miller's (1958) capital structure irrelevance proposition. In this scenario, a firm's cash flow reflects the presence of financial constraints and may contain information that is relevant for explaining the differences in investment demand across firms.

Fazzari *et al.* (1988) proposed a regression specification of the investment equation that allows the inclusion of additional explanatory variables to explain variation in the investment–capital ratio as follows:

$$y_{it} = \alpha + \beta q^*_{it} + \gamma \mathrm{CF}_{it} + \eta_{it}, \qquad i = 1, \ldots, n, \quad t = 1, \ldots, T, \tag{1}$$

with $y_{it} \equiv I_{it}/K_{it}$ the investment–capital ratio, $\mathrm{CF}_{it} \equiv \mathrm{cf}_{it}/K_{it}$ the cash flow–capital ratio, q^*_{it} represents the quantity 'marginal q' and η_{it} is the idiosyncratic structural error term, which is assumed to be zero-mean white noise.

The q^*_{it}-quantity is unobservable and researchers use instead its measurable counterpart, the average Tobin's q. Hayashi (1982) showed analytically the differences between these quantities for different production and cost functions. More specifically, he showed that

$$q^{\mathrm{a}}_{it} = \lambda_{it} + q^*_{it}, \tag{2}$$

where q^{a}_{it} denotes Tobin's average q and λ_{it} is a quantity that captures the present discounted value of current and future tax deductions attributable to past investments for a production function exhibiting constant returns to scale and a cost function that is homogeneous of order 1. The quantity λ_{it} captures other features of the production function such as the elasticity of demand for the firm's output for different market structures, e.g. when firms are price makers.

Further, following Cummins *et al.* (2006), it is also possible to accommodate the possibility of measurement error in the average Tobin's q. Consider the following specification for the observed Tobin's q:

$$q_{it} = q^{\mathrm{a}}_{it} + \nu_{it},$$

where q_{it} is the observable average Tobin's q, which is measured as the average Tobin's q, q^{a}_{it}, plus the error ν_{it}. The above expressions imply that

$$q^*_{it} = q_{it} - e_{it}, \tag{3}$$

with $e_{it} = \lambda_{it} + \nu_{it}$ denoting a modified measurement error term. Then, plugging equation (3) into the investment model (1), we obtain

$$\begin{aligned} y_{it} &= \alpha + \beta(q_{it} - e_{it}) + \gamma \mathrm{CF}_{it} + \eta_{it}, \\ &= \alpha + \beta q_{it} + \gamma \mathrm{CF}_{it} + \epsilon_{it}, \end{aligned} \tag{4}$$

where $\epsilon_{it} \equiv \eta_{it} - \beta e_{it}$ is correlated with q_{it} by the presence of the measurement error. This correlation leads to endogeneity of the regressors and inconsistent parameter estimates. Furthermore, if q_{it} is correlated with the observable exogenous variables CF_{it} then the above regression equation would also entail the correlation between CF_{it} and ϵ_{it} and the inconsistency of γ. We shall assume hereafter that cash flow is measured without error and, hence, is uncorrelated with the error term ϵ_{it}.

2.2. Failure of instrumental variable methods

It has been common in the literature to employ IV estimators to resolve the statistical problems that are induced by the presence of endogeneity in investment equation models. Almeida *et al.* (2010) showed that, under some conditions, IV methods deliver estimated coefficients that are robust and economically meaningful. These estimators employ lags of q_{it} as instruments for the endogenous variable q_{it}.

Almeida *et al.* (2010) discussed the assumptions on the dynamics of the measurement error to make IV methods valid. They found that, if the measurement error e_{it} in equation (3) is independent and identically distributed across firms and time, and q_{it} is serially correlated, then, lags of the variable with errors (e.g. q_{it-2}, q_{it-3} or $q_{it-2}-q_{it-3}$) are valid instruments for q_{it} since they are correlated with q_{it} (instrument relevance condition) but uncorrelated with the error term ϵ_{it} (instrument exogeneity condition).

However, when both the marginal Tobin's q and its measurement errors exhibit serial persistence, the IV approach, which employs the lags of the mismeasured Tobin's q as instruments, fails to solve the endogeneity problem. In this case, as further discussed in Appendix B, the instrument exogeneity condition is no longer valid because the IV (the lags of mismeasured Tobin's q) are correlated with the regression error term.

3. Econometric methodology

This section suggests an alternative method to obtain consistent parameter estimates in the investment equation model. The methodology proposed introduces an auxiliary equation that models the interaction term $q_{it}\epsilon_{it}$ as a function of observable covariates. These covariates are determined by polynomials of lagged values of average Tobin's q, and their use is motivated by the persistence of both the measurement error variable and the marginal Tobin's q. The second part of the section suggests an estimator based on the empirical counterpart of the identification result and develops inference procedures.

3.1. Econometric model and identification

3.1.1. *A preview of the solution*

For simplicity, we outline first the identification strategy of the structural parameters for the simple case given by an investment equation that depends on only the marginal Tobin's q. Consider the following simplified model:

$$y_{it}=\beta q_{it}+\epsilon_{it}, \tag{5}$$

where $\epsilon_{it}\equiv\eta_{it}-\beta e_{it}$ and $e_{it}=\lambda_{it}+\nu_{it}$. Assume that there is a set of observable variables $\mathbf{Z}_{it}$ such that

$$E[q_{it}\epsilon_{it}|q_{it},\mathbf{Z}_{it}]=\mathbf{Z}_{it}\phi, \tag{6}$$

with ϕ being a vector of parameters that are different from 0. This condition introduces an auxiliary equation given by

$$q_{it}\epsilon_{it} = \mathbf{Z}_{it}\phi + u_{it}, \tag{7}$$

where u_{it} is an error term that is orthogonal to $\mathbf{Z}_{it}$ and q_{it}^2, by construction.

Given equation (7), the intuition of the solution is as follows. Note that ϵ_{it} is not observable; however, from equations (5) and (7), $q_{it}\epsilon_{it}$ can be rewritten through the equation

$$q_{it}y_{it} = \beta q_{it}^2 + q_{it}\epsilon_{it} = \beta q_{it}^2 + \mathbf{Z}_{it}\phi + u_{it}.$$

The structural parameters of the investment equation can be identified through the introduction of the auxiliary equation (6). To see this note that from equations (5) and (7)

$$\begin{aligned} u_{it} &= q_{it}\epsilon_{it} - \mathbf{Z}_{it}\phi \\ &= q_{it}(y_{it} - \beta q_{it}) - \mathbf{Z}_{it}\phi \\ &= q_{it}y_{it} - \tilde{\mathbf{q}}_{it}(\beta, \phi^{\mathrm{T}})^{\mathrm{T}}, \end{aligned}$$

where $\tilde{\mathbf{q}}_{it} \equiv (q_{it}^2, \mathbf{Z}_{it})$. We then consider the moment equation

$$E[\tilde{\mathbf{q}}_{it}^{\mathrm{T}} u_{it}] = 0,$$

so that we obtain $E[\tilde{\mathbf{q}}_{it}^{\mathrm{T}}\{q_{it}y_{it} - \tilde{\mathbf{q}}_{it}(\beta, \phi^{\mathrm{T}})^{\mathrm{T}}\}] = 0$. Finally, by distributing the expectation, we have that

$$E[\tilde{\mathbf{q}}_{it}^{\mathrm{T}} q_{it}y_{it}] = E[\tilde{\mathbf{q}}_{it}^{\mathrm{T}}\tilde{\mathbf{q}}_{it}](\beta, \phi^{\mathrm{T}})^{\mathrm{T}}.$$

Given non-singularity of $E[\tilde{\mathbf{q}}_{it}^{\mathrm{T}}\tilde{\mathbf{q}}_{it}]$, this moment equation uniquely identifies β and ϕ.

3.1.2. General methodology

In what follows, we extend and formalize the above results and also accommodate an exogenous regressor, i.e. CF_{it}. (Galvao *et al.* (2017) discuss endogeneity bias modelling in a cross-section context.) Identification of the parameters of interest is achieved by explicitly modelling the interaction of the endogenous variable and the unobserved causes of the dependent variable as a function of additional observable variables. In particular, we consider the case where the variable $q_{it}\epsilon_{it}$ can be modelled by using additional variables. For notational simplicity, we suppress the subscripts (i, t) whenever there is no confusion. The following equation formalizes modelling endogeneity:

$$E[q\epsilon|\mathbf{z}, \mathbf{x}] = g(\mathbf{z}), \tag{8}$$

where $g(\cdot)$ is an unknown smooth function $\mathbf{z}$, a k-vector of additional observable variables, and $\mathbf{x}$ is the set of regressors, in this case $\mathbf{x} = [1, \mathrm{CF}, q]$. This is a general formulation to model the endogeneity in the linear parametric model. For simplicity, we assume that $g(\cdot)$ is a known function of $\mathbf{z}$ with unknown parameters ϕ such as $g(\mathbf{z}; \phi)$. But the analysis can be extended to the case of unknown functional form of $g(\cdot)$, as we might approximate the unknown function $g(\cdot)$ with one of the sieve bases (e.g. power series, Fourier series or splines). (See, for example, Chen (2007) for more details on the method of sieves.)

A simple example of equation (8), that is convenient for exposition and estimation purposes, is to assume the following polynomial approximation for $g(\mathbf{z})$:

$$E[q\epsilon|\mathbf{z}, \mathbf{x}] = \mathbf{Z}\phi, \tag{9}$$

where $\mathbf{Z} = (1, \mathbf{z}, \mathbf{z}^2, \ldots, \mathbf{z}^m)$ and $\phi = (\phi_0, \phi_1^{\mathrm{T}}, \ldots, \phi_m^{\mathrm{T}})^{\mathrm{T}}$, which is a non-zero vector with $\phi \neq \mathbf{0}$. Equation (9) is explicitly modelling the endogeneity of q and requires observable variables $\mathbf{z}$, the simultaneous variables. We are interested in identifying and estimating the parameters (β, γ) in equation (1). In practice, ϕ is unknown, and it is important to note that this parameter cannot be directly estimated from equation (9) because ϵ is unobservable. Hence, we consider the joint identification and estimation of both (β, γ) and ϕ.

Consider now the structural model in equation (1) and define $\boldsymbol{\theta} \equiv (\alpha, \gamma, \beta, \boldsymbol{\phi}^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{\theta}_1 \equiv (\alpha, \gamma)^{\mathrm{T}}$ and $\boldsymbol{\theta}_2 \equiv (\beta, \boldsymbol{\phi}^{\mathrm{T}})^{\mathrm{T}}$, $\mathbf{x} \equiv [1, \mathrm{CF}, q]$ and $\mathbf{x}_1 \equiv (1, \mathrm{CF})$. To ease notation, define $\tilde{y}$ and $\tilde{q}$ after netting out the exogenous regressor $\mathbf{x}_1$ and multiplying the resulting objects by q. Let $\tilde{y} = q(y - \mathbf{x}_1 E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1} E[\mathbf{x}_1^{\mathrm{T}} y])$ and $\tilde{\mathbf{x}} = (\tilde{q}, \mathbf{Z})$, with $\tilde{q} = q(q - \mathbf{x}_1 E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1} E[\mathbf{x}_1^{\mathrm{T}} q])$ and $\mathbf{Z} = (1, \mathbf{z}, \mathbf{z}^2, \ldots, \mathbf{z}^m)$. Further, consider the following assumptions.

Assumption 1.

(a) The variable $\mathbf{x}_1$ is exogenous such that $E[\mathbf{x}_1^{\mathrm{T}}\epsilon] = \mathbf{0}$.
(b) The matrices $E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]$ and $E[\tilde{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{x}}]$ are non-singular.
(c) We have $q\epsilon = \mathbf{Z}\boldsymbol{\phi} + u$, where $\boldsymbol{\phi} \neq \mathbf{0}$ and $E[\tilde{\mathbf{x}}^{\mathrm{T}} u] = \mathbf{0}$.

Assumption 1, part (a), states that CF (i.e. cash flow) is an exogenous regressor. In practice, we propose to use polynomials of lagged values of average Tobin's q as $\mathbf{Z}$. Thus, assumption 1, part (b), is satisfied if $\tilde{q}_{it}$ is not perfectly linearly related to q_{it-1} and q_{it-1}^2, and its higher order terms. Finally, assumption 1, part (c), means that $E[(q_{it-1}, q_{it-1}^2, \ldots, q_{it-1}^m) u_{it}] = 0$ and $E[\tilde{q}_{it} u_{it}] = 0$ hold, where $\tilde{q}_{it}$ is the product of q_{it} and q_{it} with CF_{it} netted out. Assumption 1, part (a), is a standard condition in the corporate finance literature, and assumption 1, part (b), is satisfied in practice. The two conditions in assumption 1, part (c), crucially depend on the validity of $\mathbf{Z}$. In particular, the conditions depend on u_{it}, the residual projection of $q\epsilon$ on $\mathbf{Z}$, being uncorrelated with $\tilde{\mathbf{x}}_{it}$, where $\tilde{\mathbf{x}} = (\tilde{q}, \mathbf{Z})$ with $\tilde{q} = q(q - \mathbf{x}_1 E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1} E[\mathbf{x}_1^{\mathrm{T}} q])$. The identification condition 1, part (c), requires that $\mathbf{Z}$ captures as much information as possible on $q\epsilon$ so that the remainder is not further correlated with the square of q (after CF has been netted out). We expect that lags of q_{it} contain useful information on q_{it} in the presence of auto-regressive time structure or any other persistent process. The following theorem formalizes the identification results.

Theorem 1. Suppose that assumption 1 holds. Then, $\boldsymbol{\theta}$ is uniquely identified as

$$\boldsymbol{\theta}_1 = E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1} E[\mathbf{x}_1^{\mathrm{T}} y] - E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1} E[\mathbf{x}_1^{\mathrm{T}} q]\beta,$$
$$\boldsymbol{\theta}_2 = E[\tilde{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{x}}]^{-1} E[\tilde{\mathbf{x}}^{\mathrm{T}} \tilde{y}].$$

It remains to address the empirical question of choosing an appropriate set of observable variables for $\mathbf{Z}$. The strategy that is pursued in this paper allows, contrary to IV methods, both q^* and e to be auto-correlated such that lags of q can be used to achieve identification and to produce consistent estimates of the parameters of interest. A correlation condition seems very natural in this context if we consider that Tobin's q and, potentially, λ_t in equation (2) are both persistent over time. Thus, the main objective is to verify assumption 1, part (c), under a correlation condition. We formalize the result in the following proposition for the investment equation where both q^* and e are allowed to be auto-correlated. We consider the investment equation (5) that depends only on Tobin's q, and the exogenous regressor, cash flow CF, can be straightforwardly included by changing the notation *à la* theorem 1.

Proposition 1. Suppose that the sequence of measurement errors follows $e_{it} = h(e_{it-1}, q_{it-1}^*)$ for a measurable function h, q_{it}^* is auto-correlated and $g(\mathbf{z}_{it})$ can be approximated by a polynomial of order m. Then, assumption 1, part (c), is satisfied, i.e.

$$q_{it}\epsilon_{it} = \mathbf{Z}_{it}\boldsymbol{\phi} + u_{it}$$

where $E[\tilde{\mathbf{x}}_{it}^{\mathrm{T}} u_{it}] = 0$ with $\tilde{\mathbf{x}}_{it} \equiv (q_{it}^2, \mathbf{Z}_{it})$, $\mathbf{Z}_{it} \equiv (1, q_{it-1}, q_{it-1}^2, \ldots, q_{it-1}^m)$ and $u_{it} \equiv q_{it}\eta_{it} - \beta q_{it} e_{it} - E[\beta q_{it} e_{it} | \mathbf{z}_{it}]$.

First, the assumption on e_{it} in proposition 1 imposes an auto-correlation model on the measurement errors. It also allows for the dependence between the lags of true Tobin's q and the measurement error. The interpretation of this assumption is intuitive. In $e_{it} = h(e_{it-1}, q^*_{it-1})$ the variable q^*_{it-1} is unobserved and could be interpreted as an innovation shock, which means that the shock to the measurement error process of Tobin's q is driven by the past values of the true unobserved Tobin's q. The assumption on e_{it} also encompasses non-classical measurement errors in that the true Tobin's q is correlated with measurement error (or true Tobin's q affects levels of measurement error). This is very useful in practice since the size of the measurement error could depend on the level of true Tobin's q (or firms' expected stream of future marginal benefits from using capital). For instance, the larger the true q^*, the lower the measurement error since larger firms usually have more refined accounting systems. Note that the condition on e_{it} invalidates the instrument exogeneity condition in the IV model, but it is a key element in the estimator proposed. Second, the assumption on q^*_{it} being auto-correlated is mild and only used to guarantee that the lags of observed Tobin's q are valid simultaneous variables. Finally, the condition that $g(\mathbf{z}_{it})$ can be approximated by a polynomial is commonly used in empirical applications.

We note that the result in proposition 1 implies that under the stated conditions assumption 1, part (c), is satisfied, and hence a polynomial of the lags of average q serves as covariates to model the endogeneity that is implied by the measurement errors in the investment equation model. The intuition behind this result is that the first-order lag of mismeasured q_{it} contains sufficient information on the interaction of mismeasured q_{it} and the regression error term ϵ_{it}.

3.2. *Estimation and inference*

In this section we construct an estimator which is simple to implement in practice and derive its asymptotic properties. We define this estimator as the S-estimator as we are specifically modelling the simultaneous covariance between the endogenous variable and the error term.

An estimator of $\boldsymbol{\theta}$ motivated by results in theorem 1 is

$$\hat{\boldsymbol{\theta}}_1 = \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}y_{it} - \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}q_{it}\right)\hat{\beta},$$

$$\hat{\boldsymbol{\theta}}_2 = \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{\mathbf{x}}_{it}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{y}_{it},$$

where $\hat{x}$ and $\hat{y}$ are sample analogues of $\tilde{x}$ and $\tilde{y}$, which are obtained by replacing the expectations with sample means, and where $\hat{\beta}$ is the first element of $\hat{\boldsymbol{\theta}}_2$. Implementation of the estimator is simple and can be carried through a sequence of OLS estimations. First, compute $\hat{x}$ and $\hat{y}$ and estimate $\hat{\boldsymbol{\theta}}_2$ by using OLS. These generated variables affect the asymptotic variance–covariance matrix (see for example Pagan (1984)). Finally, given $\hat{\beta}$, $\hat{\boldsymbol{\theta}}_1$ can be estimated by OLS.

Define $Q \equiv E[\tilde{\mathbf{x}}_{it}^{\mathrm{T}}\tilde{\mathbf{x}}_{it}]$, $\mathbf{C}_1 \equiv E[\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}]$ and $C_2 \equiv E[\mathbf{x}_{1it}^{\mathrm{T}}q_{it}]$. To establish the asymptotic properties of the estimator, consider the following assumptions. For simplicity, we consider a balanced panel and the case of large n and fixed T.

Assumption 2.

(a) The data $\{(y_{it}, \mathbf{x}_{it}, \mathbf{Z}_{it}); i = 1, 2, \ldots, n, t = 1, 2, \ldots, T\}$ are independent and identically distributed across i.
(b) We have $E[\|y_{it}\|^4] < \infty$, $E[\|\mathbf{x}_{it}\|^4] < \infty$ and $E[\|\mathbf{Z}_{it}\|^4] < \infty$.
(c) The matrices Q and $\mathbf{C}_1$ are non-singular.

The asymptotic properties of the S-estimator are summarized in the following result.

Theorem 2. Let assumptions 1 and 2 hold with $E[u|\tilde{\mathbf{x}}]=0$ in place of $E[\tilde{\mathbf{x}}^{\mathrm{T}}u]=0$. Then, as $n\rightarrow\infty$,

$$\hat{\boldsymbol{\theta}}_1 \xrightarrow{\mathrm{p}} \boldsymbol{\theta}_1,$$
$$\hat{\boldsymbol{\theta}}_2 \xrightarrow{\mathrm{p}} \boldsymbol{\theta}_2,$$

and

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_1) \xrightarrow{\mathrm{d}} N(0, \mathbf{C}_1^{-1}C_2V_{\hat{\beta}}C_2^{\mathrm{T}}\mathbf{C}_1^{-1}),$$
$$\sqrt{n}(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_2) \xrightarrow{\mathrm{d}} N(0, Q^{-1}MQ^{-1}),$$

with $V_{\hat{\beta}}=\mathrm{var}(\hat{\beta})$ and $M=\mathrm{var}\{\tilde{\mathbf{x}}^{\mathrm{T}}u-Gr(\delta_q)+Hs(\delta_y)\}$, where G, $r(\delta_q)$, H and $s(\delta_y)$ are defined in the proof.

Given the result in theorem 2, general hypotheses on the vector $\boldsymbol{\theta}$ can be easily accommodated by Wald-type tests. The Wald process and associated limiting theory provide a natural foundation for testing the linear null hypothesis $H_0: R\boldsymbol{\theta}=r$, when r is known. In practice, to carry out inference and to apply a Wald-type test we need a consistent estimator of the asymptotic variance matrix. As described in the above result, to estimate the asymptotic variance–covariance matrix, we need to estimate both $\mathrm{var}(\hat{\boldsymbol{\theta}}_2)=Q^{-1}MQ^{-1}/n$ and $\mathrm{var}(\hat{\boldsymbol{\theta}}_1)=\mathbf{C}_1^{-1}C_2V_{\beta}C_2\mathbf{C}_1^{-1}/n$. The latter is easily recovered from its sample counterparts, i.e.

$$\mathbf{C}_1=\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it},$$
$$C_2=\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}q_{it}.$$

In addition, $V_{\hat{\beta}}$ is the first element of the variance–covariance matrix $\mathrm{var}(\hat{\boldsymbol{\theta}}_2)$. Finally, for the estimation of the variance–covariance matrix of $\hat{\boldsymbol{\theta}}_2$ we can consider its sample counterpart such as $\widehat{\mathrm{var}}(\hat{\boldsymbol{\theta}}_2)=\hat{Q}^{-1}\hat{M}\hat{Q}^{-1}/n$ with

$$\hat{Q}=\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{\mathbf{x}}_{it},$$

$$\hat{M}=\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}(\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{u}_{it}-\hat{G}\hat{r}_{it}(\delta_q)+\hat{H}\hat{s}_{it}(\delta_y))(\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{u}_{it}-\hat{G}\hat{r}_{it}(\delta_q)+\hat{H}\hat{s}_{it}(\delta_y))^{\mathrm{T}}$$

where

$$\hat{u}_{it}=\hat{y}_{it}-\hat{\mathbf{x}}_{it}\hat{\boldsymbol{\theta}}_2,$$

$$\hat{G}=\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{\mathbf{x}}_{it}^{\mathrm{T}}\nabla_{\delta_q}\tilde{\mathbf{x}}_{it}\hat{\boldsymbol{\theta}}_2=\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}(-q_{it}\mathbf{x}_{1it},0,0)\hat{\boldsymbol{\theta}}_2,$$

$$\hat{H}=\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{\mathbf{x}}_{it}^{\mathrm{T}}\nabla_{\delta_y}\tilde{y}_{it}=\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}(-q_{it}\mathbf{x}_{1it}),$$

$$\hat{r}_{it}(\hat{\delta}_y)=\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\mathbf{x}_{1it}^{\mathrm{T}}(q_{it}-\mathbf{x}_{1it}\hat{\delta}_q),$$

$$\hat{s}_{it}(\hat{\delta}_y) = \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} \mathbf{x}_{1it} \right)^{-1} \mathbf{x}_{1it}^{\mathrm{T}} (y_{it} - \mathbf{x}_{1it} \hat{\delta}_y),$$

$$\hat{\delta}_q = \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} \mathbf{x}_{1it} \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} q_{it}$$

and

$$\hat{\delta}_y = \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} \mathbf{x}_{1it} \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} y_{it}.$$

4. Empirical analysis for a panel of US firms

The purpose of this empirical analysis is twofold. First, we compare and statistically assess the parameter estimates determining the relationship between corporate investment demand, Tobin's q and firms' cash flow obtained under various estimation procedures. We highlight the differences between existing procedures such as OLS, IV and the GMM and our novel S-estimator. Second, we carry out an exhaustive exercise to determine the effect of firms' cash flow on investment demand and assess whether this relationship depends on the extent of financial constraints that are faced by firms. To do this, we estimate the investment regression equation for several subsamples of financially constrained and unconstrained firms classified according to different criteria as set out in Almeida *et al.* (2004), Moyen (2004) and Hadlock and Pierce (2010).

4.1. Data

The data are taken from the COMPUSTAT® database and cover the period 1974–2010. The data collection process follows an extension of that of Almeida and Campello (2007). The sample consists of manufacturing firms with fixed capital of more than $5 million (with 1974 as the base year for the consumer price index). Firms in the sample have growth of less than 100% in both assets and sales. We keep observations with at least three lags. Summary statistics for investment and cash flow are presented in Table 1. These statistics are similar to those reported by Almeida and Campello (2007), among others. The sample corresponds to an unbalanced panel of 3085 firms, with an average number of observations per firm of 10.60. For brevity, we omit the discussion of these descriptive statistics.

Table 1. Investment model: descriptive statistics†

Variable	*Mean*	*Standard deviation*	*Minimum*	*Maximum*	*Observations*
Investment	0.2070	0.1465	0.0000	8.5031	32698
q	0.9295	0.4565	0.1786	22.7962	32698
CF	0.3512	0.6666	−83.0667	17.9674	32698

†Data from COMPUSTAT, 1974–2010. The sample consists of manufacturing firms with fixed capital of more than $5 million (with 1976 as the base year for the consumer price index), and the sample firms have growth of less than 100% in both assets and sales: observations with at least three lags.

4.2. *Empirical results*

4.2.1. *Comparison of estimation methods*

We use first simple estimators that do not correct for measurement errors like OLS and least squares fixed effects (FEs). We add FE methods to control for unobserved firm-specific effects such as manager's ability or geographical characteristics, and also for the unbalanced nature of the sample. Second, we employ estimators that seek to correct for measurement error, which are IV two-stage least squares and the GMM, and our proposed S-estimator where we model the joint interaction of Tobin's q and the error term. For the second set of estimators, we use data in levels as well as the demeaned version of the estimators for controlling for FEs. (For the presentation, we do not report the results for the GMM and FE GMM cases. For both methods, the results are very similar to the IV counterpart estimators. The results for these methods are, nevertheless, available from the authors on request.)

The results for OLS and FE models are summarized in Table 2. These are the benchmark estimates. Table 3 collects the results for the standard IV estimator and Table 4 for the FE version of the IV estimator. The results that were obtained from our simultaneous estimator are collected in Tables 5 and 6, for the model in levels and demeaned data respectively. All the reported methods use different combinations of lags of the endogenous variables (Tobin's q) as instruments. In our model, these combinations of lags are interpreted as candidates to model the joint interaction between the endogenous variable and the error term. The results for different choices of lagged values of Tobin's q assess the robustness of our procedure to such choice.

The OLS estimates show that both q and CF have positive coefficients that are statistically significant (see Table 2). In the OLS specification, we obtain a Tobin's q-coefficient of 0.0726, and a cash flow coefficient of 0.0521, which are likely to be biased. These results are also consistent with those found in the OLS-related literature, in particular Agca and Mozumdar (2017) and Almeida *et al.* (2010). Next, we consider firm-specific FEs. The results are also in Table 2. The estimate for Tobin's q increases to 0.0809 and for cash flow decreases to 0.0432, both being statistically significant. These are, however, not statistically different from the OLS estimates.

We move our attention to models that correct for the measurement error in Tobin's q. First, we consider the IV estimates, which are represented by a two-stage least squares estimator in Table 3. We report the value of the parameter estimates that are associated with q and CF for various choices of instruments based on lags of q_t and q_t^2 such that each estimator builds on the previous

Table 2. Investment model: least squares models†

Variable	*OLS*	*FEs*
q	0.0725‡	0.0809‡
	(0.0089)	(0.0078)
CF	0.0521‡	0.0432‡
	(0.0153)	(0.0122)
Constant	0.1213‡	0.1167‡
	(0.0047)	(0.0059)
Observations	32698	32698
R^2	0.188	0.187
Number of firms	3085	3085

†Standard errors are in parentheses. For OLS and FEs we use cluster robust standard errors by firm.
‡$p < 0.01$.

Table 3. IVs approach

Variable	*IV 1*	*IV 2*	*IV 3*	*IV 4*	*IV 5*	*IV 6*
Second stage						
q	0.0463‡	0.0520‡	0.0529‡	0.0529‡	0.0524‡	0.0524‡
	(0.0058)	(0.0062)	(0.0062)	(0.0062)	(0.0062)	(0.0062)
CF	0.0513‡	0.0514‡	0.0515‡	0.0515‡	0.0515‡	0.0515‡
	(0.0183)	(0.0176)	(0.0175)	(0.0175)	(0.0176)	(0.0176)
Constant	0.146‡	0.141‡	0.140‡	0.140‡	0.140‡	0.140‡
	(0.0047)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)
Observations	32698	32698	32698	32698	32698	32698
R^2	0.095	0.098	0.098	0.098	0.098	0.098
First stage						
CF	−0.114	−0.128§	−0.129§	−0.129§	−0.129§	−0.129§
	(0.0713)	(0.0692)	(0.0683)	(0.0683)	(0.0684)	(0.0684)
q_{t-1}	0.618‡	0.826‡	0.876‡	0.867‡	0.877‡	0.877‡
	(0.0190)	(0.0440)	(0.0323)	(0.0585)	(0.0572)	(0.0569)
q_{t-1}^2		−0.0251‡	−0.0261‡	−0.0254‡	−0.0259‡	−0.0258‡
		(0.0054)	(0.0055)	(0.0077)	(0.0078)	(0.0077)
q_{t-2}			−0.0412§§	−0.0295	−0.0982‡	−0.112‡
			(0.0183)	(0.0252)	(0.0193)	(0.0284)
q_{t-2}^2				−0.0009	0.0005	0.0016
				(0.0034)	(0.0036)	(0.0043)
q_{t-3}					0.0527‡	0.0712‡
					(0.0129)	(0.0119)
q_{t-3}^2						−0.00158
						(0.00135)
Constant	0.369‡	0.204‡	0.200‡	0.198‡	0.199‡	0.196‡
	(0.0400)	(0.0147)	(0.0149)	(0.0117)	(0.0115)	(0.0117)
Observations	32698	32698	32698	32698	32698	32698
R^2	0.674	0.716	0.718	0.718	0.720	0.720

†Cluster robust standard errors by firm are in parentheses.
‡$p < 0.01$.
§$p < 0.1$.
§§$p < 0.05$.

set of instruments added by one further lag of q or q^2. The results in Table 3 highlight the role of adding more lags up to q_{t-3}. The parameter estimates are very similar across estimators, yielding values oscillating about 0.0524 and 0.0515 for Tobin's q and cash flow respectively, for the IV estimator. Table 4 reports the estimates of the IV estimator that controls for FEs. The results of this exercise are very similar to those obtained for the previous IV case.

Before presenting the results for the S-estimator, we assess the potential presence of correlation in the investment equation by implementing Arellano and Bond (1991) tests for auto-correlation. Table 7 reports the tests for auto-correlation of orders 1, 2 and 3 for the residuals of the OLS and IV models, and of the corresponding counterpart models accounting for the presence of FEs (FE and FE–IV). The evidence points to strong persistence in the error term even for the residuals of IV models, which signals persistence in measurement errors in Tobin's q and, hence, provides evidence that invalidates the IV methods.

The presence of persistence in measurement errors, thus, motivates our proposed estimator. To provide further evidence on the suitability of our method for estimating the investment equation model we discuss the following example. Suppose that we compute OLS residuals from Table 2, column OLS, i.e. $\hat{\varepsilon}_t = y_{it} - (\hat{\alpha}_{\mathrm{OLS}} + \hat{\beta}_{\mathrm{OLS}} q_{it} + \hat{\gamma}_{\mathrm{OLS}} \mathrm{CF}_{it})$. The evidence in the above paragraph determines that this would be correlated with the lags of Tobin's q. Fig. 1(a) shows that q_{t-1}

Table 4. FEs IVs†

Variable	*FE IV 1*	*FE IV 2*	*FE IV 3*	*FE IV 4*	*FE IV 5*	*FE IV 6*
Second stage						
q	0.0420‡	0.0490‡	0.0506‡	0.0506‡	0.0496‡	0.0495‡
	(0.0027)	(0.0024)	(0.0026)	(0.0026)	(0.0026)	(0.0026)
CF	0.038‡	0.039‡	0.039‡	0.039‡	0.039‡	0.0394‡
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Constant	0.154‡	0.147‡	0.146‡	0.146‡	0.147‡	0.147‡
	(0.0027)	(0.0026)	(0.0026)	(0.0026)	(0.0027)	(0.0027)
Observations	32698	32698	32698	32698	32698	32698
R^2	0.005	0.067	0.068	0.068	0.068	0.068
First stage						
CF	−0.162‡	−0.167‡	−0.168‡	−0.168‡	−0.168‡	−0.168‡
	(0.0022)	(0.0021)	(0.0021)	(0.0022)	(0.0021)	(0.0021)
q_{t-1}	0.511‡	0.674‡	0.712‡	0.719‡	0.728‡	0.728‡
	(0.0026)	(0.045)	(0.006)	(0.0068)	(0.0068)	(0.0068)
q^2_{t-1}		−0.016‡	−0.017‡	−0.0179‡	−0.0183‡	−0.0183§
		(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
q_{t-2}			−0.032‡	−0.0426‡	−0.0982‡	−0.100‡
			(0.0034)	(0.0058)	(0.0070)	(0.0076)
q^2_{t-2}				0.0008	0.0021	0.0022
			(0.0004)	(0.0043)	(0.0004)	(0.0046)
q_{t-3}					0.0452‡	0.0481‡
					(0.0032)	(0.0055)
q^2_{t-3}						−0.0002
						(0.0004)
Constant	0.354‡	0.196‡	0.352‡	0.352‡	0.353‡	0.353‡
	(0.003)	(0.0041)	(0.0042)	(0.0044)	(0.0044)	(0.0046)
Observations	32698	32698	32698	32698	32698	32698
R^2	0.577	0.604	0.605	0.605	0.720	0.720

†Cluster robust standard errors by firm are in parentheses.
‡$p<0.01$.
§$p<0.05$.

is correlated with both q_t and OLS residuals. A visual inspection reveals that the graph has a clear negative slope in the direction given by Tobin's q but it is also related to the OLS residuals. Fig. 1(b) displays the quantity $\hat{\varepsilon}_t \times q_t$ as a function of q_{t-1}. The plot reveals a positive correlation between the joint interaction of Tobin's q and the OLS residual with the first lag of Tobin's q.

We complete the section by presenting results for two different versions (one in levels and one with demeaned data using the within transformation) of the proposed S-estimator. The results are presented in Tables 5 and 6. The coefficients for both Tobin's q and CF are statistically different from 0. In fact, Tobin's q-coefficient is almost twice as large as the standard OLS estimate across models, and the cash flow coefficient is about 10% smaller than the corresponding OLS value. Thus, although the theory does not pin down the exact values that this coefficient should take, one could argue that an estimator that solves measurement error in the true q^* in a standard investment equation should return a higher estimate for β when compared with standard OLS estimates. This is so because the measurement error causes an attenuation bias on the estimate for this coefficient. On the contrary, the IV procedure returns estimates for Tobin's q that do not satisfy this condition, although they are statistically significant. Further, the values of the parameter estimates are quite stable across exercises but the comparison of the estimates across models highlights the relevance of including q_{t-1} and q^2_{t-1} in the set of observables Z compared with including q_{t-2} and beyond. (The choice of an

Table 5. *S*-estimator†

Variable	*S 1*	*S 2*	*S 3*	*S 4*	*S 5*	*S 6*
q	0.1519‡	0.1748‡	0.1751‡	0.1760‡	0.1762‡	0.1764‡
	(0.042)	(0.031)	(0.031)	(0.030)	(0.030)	(0.030)
CF	0.0555‡	0.0554‡	0.0554‡	0.0554‡	0.0554‡	0.0554‡
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.0466‡	0.0250‡	0.0248‡	0.0239‡	0.0238‡	0.0235‡
	(0.040)	(0.030)	(0.029)	(0.029)	(0.029)	(0.028)
q_{t-1}	−0.0950§	0.0020	0.0759	−0.0303	−0.0390	−0.0488
	(0.034)	(0.021)	(0.043)	(0.054)	(0.054)	(0.057)
q^2_{t-1}		−0.0349‡	−0.0362‡	−0.0288‡	−0.0284‡	−0.0278§
		(0.013)	(0.012)	(0.015)	(0.015)	(0.015)
q_{t-2}			−0.0682	0.0399	0.0883‡	0.0469
			(0.055)	(0.045)	(0.049)	(0.040)
q^2_{t-2}				−0.0088	−0.0099	−0.0067‡
				(0.008)	(0.008)	(0.007)
q_{t-3}					−0.0372‡	0.0144
					(0.014)	(0.022)
q^2_{t-3}						−0.0044
						(0.003)
Observations	32698	32698	32698	32698	32698	32698

†Standard errors are in parentheses.
‡$p < 0.01$.
§$p < 0.05$.

appropriate polynomial $\mathbf{Z}_{it}$ is analogous to the IV approach of which consistency heavily relies on an appropriate selection of valid instruments. This empirical result highlights the importance of studying the optimal number of lags for this novel estimation method. We leave this study for future research. As a check of robustness, we also considered in unreported exercises higher order moments such as the cubes of the lags of q but the results do not substantially change.) The analysis of the results for the *S*-estimator using demeaned variables reported in Table 6 sheds similar findings to those of Table 5. The magnitudes of the parameter estimates for both Tobin's q and firms' cash flows are, however, slightly higher. The role of q_{t-2} and higher order lags are more relevant than in the version of the *S*-estimator from the data in levels.

The empirical findings in Tables 5 and 6 suggest that the effect of Tobin's q on the investment capital ratio is greater than that estimated by OLS and IV. These findings are in themselves evidence that the *S*-estimator is doing a good job of capturing the true variation in investment opportunities. Low observed Tobin's q-coefficients have been noted in the literature as a diagnostic for the failure of the empirical investment model; however, the estimates that were obtained from our proposed method are significantly higher. Furthermore, the sensitivity of cash flow does not vanish after correcting for measurement error bias. This empirical observation was also noted in Agca and Mozumdar (2017) and challenges previous literature minimizing the role of cash flows in explaining the rate of investment. This interesting empirical finding is analysed in more detail in the following section.

4.2.2. The sensitivity of cash flow and financial constraints

The literature on the influence of cash flow on investment is very controversial. In this empiri-

Table 6. *S*-estimator using demeaned variables†

Variable	*S FE 1*	*S FE 2*	*S FE 3*	*S FE 4*	*S FE 5*	*S FE 6*
q	0.173‡	0.184‡	0.184‡	0.185‡	0.185‡	0.186‡
	(0.041)	(0.036)	(0.036)	(0.036)	(0.036)	(0.035)
CF	0.0551‡	0.0563‡	0.0564‡	0.0565‡	0.0565‡	0.0565‡
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Constant	0.0115	0.0126	0.0126	0.0127	0.0128	0.0128
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
q_{t-1}	−0.0701‡	0.0249	0.0623‡	0.0225	0.0131	0.0118
	(0.027)	(0.026)	(0.021)	(0.025)	(0.025)	(0.025)
q_{t-1}^2		−0.0221‡	−0.0230‡	−0.0193§	−0.0188§	−0.0185§
		(0.012)	(0.011)	(0.012)	(0.011)	(0.011)
q_{t-2}			−0.0384	0.0164	0.0465‡	0.0320‡
			(0.024)	(0.014)	(0.018)	(0.018)
q_{t-2}^2				−0.0071‡	−0.0080§	−0.0065
				(0.005)	(0.005)	(0.005)
q_{t-3}					−0.0238‡	−0.0026‡
					(0.011)	(0.012)
q_{t-3}^2						−0.0028
						(0.002)
Observations	32698	32698	32698	32698	32698	32698

†Standard errors are in parentheses. All variables are previously demeaned with the within transformation.
‡$p<0.01$.
§$p<0.1$.

cal exercise, we separate the analysis into firms that are financially constrained and those that are not, and we assess the sensitivity of investment ratio to cash flow in each case. We follow seminal contributions in the literature such as Almeida *et al.* (2004), Moyen (2004) and Hadlock and Pierce (2010) to characterize constrained and unconstrained firms. (There are other influential studies proposing indices to characterize firms in terms of financial conditions such as Kaplan and Zingales (1997) or Whited and Wu (2006); however, the nature of our database does not allow us to compute these indices. We, nevertheless, believe that the empirical exercise that is carried out below is quite comprehensive as it covers most of the distinctive features characterizing each of the different indices.) In particular, we discuss four alternative schemes to distinguish between constrained and unconstrained firms. The first two schemes have been taken from Almeida *et al.* (2004), who proposed five schemes to distinguish between financially constrained and unconstrained firms but, for brevity and availability of data, we focus on only those schemes that are determined by the pay-out ratio and firms' asset size. For completeness, we reproduce here schemes 1 and 2.

Scheme 1. Firms are ranked on the basis of their pay-out ratio and assigned to the financially constrained or unconstrained group of those firms in respectively the bottom or top three deciles of the annual pay-out distribution. The pay-out ratio is computed as the ratio of total distributions (dividends plus stock repurchases) to operating income. Financially constrained firms have lower pay-out ratios (see Fazzari *et al.* (1988) and Almeida *et al.* (2004) for further motivation on this finding).

Scheme 2. Firms are ranked on the basis of their asset size over the 1974–2010 period and

Table 7. Arellano and Bond (1991) tests for auto-correlation in the residuals†

	OLS		*FEs*		*IV 6*		*FE IV 6*	
	z	*p-value*	*z*	*p-value*	*z*	*p-value*	*z*	*p-value*
1	73.14	0.000	47.66	0.000	18.78	0.000	18.82	0.000
2	43.23	0.000	11.80	0.000	12.71	0.000	7.61	0.000
3	30.74	0.000	−0.30	0.764	10.61	0.000	−1.12	0.265

†Arellano and Bond (1991) tests for auto-correlation in the equation $y_{it} = \alpha + \beta q_{it} + \gamma \mathrm{CF}_{it} + \varepsilon_{it}$.

assigned to the financially constrained or unconstrained group of those firms in respectively the bottom or top three deciles of the size distribution. The rankings are again performed annually. Small firms are typically young, less well known and thus more vulnerable to capital market imperfections so they are identified as financially constrained as opposed to large firms that are identified as financially unconstrained.

We also propose a third scheme advocated in Moyen (2004) based on firms' dividends rather than on the pay-out ratios.

Scheme 3. Firms are ranked on the basis of dividends (sums of dividends on common and preferred stocks) over the 1974–2010 period and assigned to the financially constrained or unconstrained group of those firms in respectively the bottom or top three deciles of the dividend payment distribution. The rankings are again performed annually.

As a further check of robustness, we also consider the classification of firms by using the index proposed in Hadlock and Pierce (2010). (We are grateful to a referee for suggesting this analysis.) This index is defined as a combination of firms' asset size and age; more specifically, the index is constructed as $-0.737\mathrm{size} + 0.043\mathrm{size}^2 - 0.040\mathrm{age}$.

Scheme 4. Firms are ranked on the basis of the Hadlock and Pierce (2010) index computed over the 1974–2010 period. The financially constrained or unconstrained group is comprised of those firms in respectively the top or bottom three deciles of the index distribution. The rankings are again performed annually. In contrast with the previous schemes, and due to the definition of the index, the constrained group is found at the top deciles of the distribution.

We proceed now to discuss the empirical findings that were obtained from running the investment equation separately for each group of firms.

Table 8 presents the estimates of the q- and CF-coefficients for scheme 1. This scheme classifies firms with high and low pay-out ratios as financially unconstrained and constrained respectively. There are significant differences in the magnitude of the Tobin's q-parameter estimates across groups. In particular, the sensitivity of Tobin's q to investment demand is several times higher for constrained firms than for unconstrained firms. In contrast, we observe a greater sensitivity of investment to firms' cash flows for unconstrained firms than for constrained firms. This result contrasts with the literature on investment that follows from Fazzari *et al.* (1988) and agrees with the findings that were obtained in Kaplan and Zingales (1997). The comparison between estimation methods also highlights important differences between the IV methodology and our proposed S-estimators. These differences are particularly large for the analysis of Tobin's q for constrained firms. For example, the FE IV 6 estimator for low pay-out ratios reports parameter

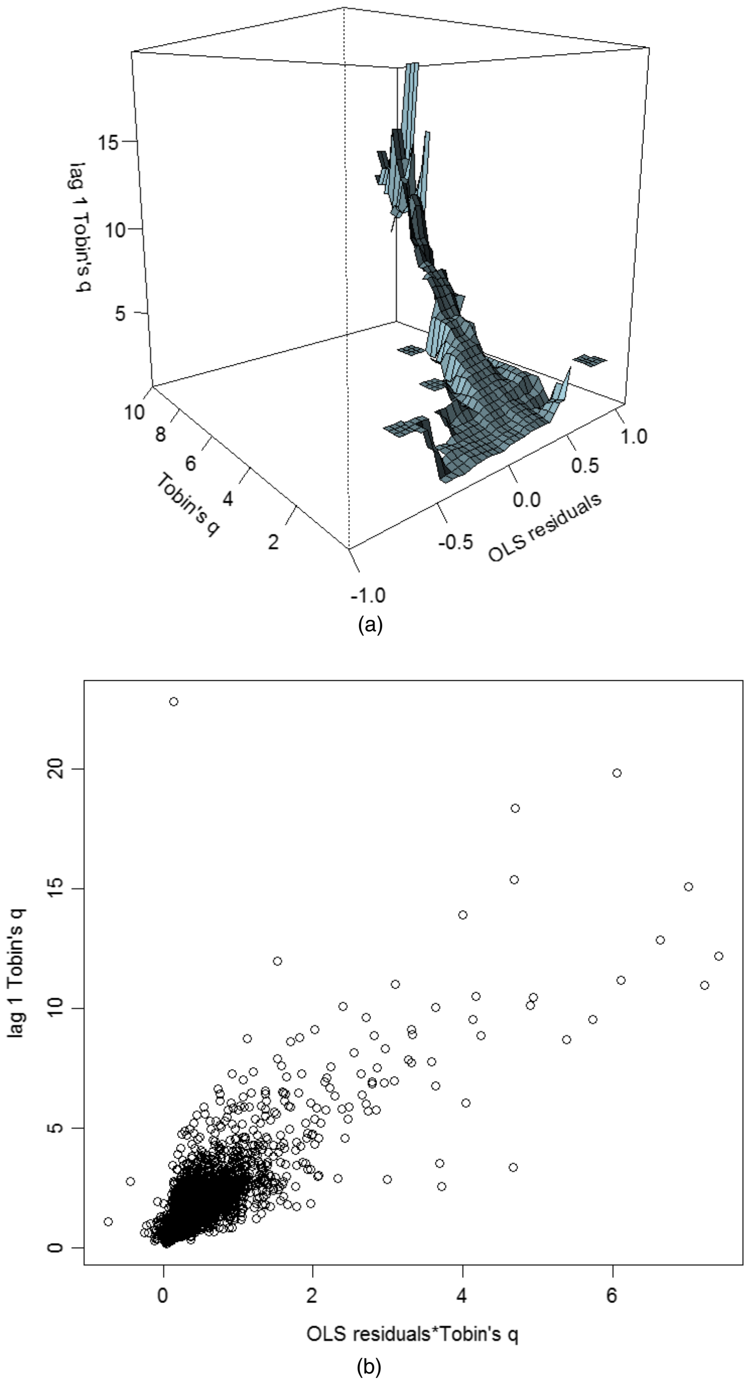

Fig. 1. Lag of Tobin's q as a function of OLS residuals and Tobin's q: (a) consider $\mathcal{G}_q$ and $\mathcal{G}_\varepsilon$, grids of size 50 of spaced points on the range of q_t and the OLS residuals $\hat{\varepsilon}_t$—then on the plane given by $\mathcal{G}_q \times \mathcal{G}_\varepsilon$ with element $(g_q(h), g_\varepsilon(j))$, $h, j = 1, 2, \ldots, 50$, we compute the average q_{t-1} of all observations with $g_q(h-2) \leqslant q_{t-1} < g_q(h+2)$ and $g_\varepsilon(j) \leqslant q_{t-1}\hat{\varepsilon} < g_\varepsilon(j)$ (empty areas correspond to $(\hat{\varepsilon}, q)$ cells with no observations); (b) q_t multiplied by the OLS residuals $\hat{\varepsilon}_t$ *versus* q_{t-1}

Table 8. Investment model for unconstrained firms characterized by high and low pay-out ratios†

Variable	*OLS*	*FEs*	*IV 6*	*FE IV 6*	*SE 6*	*S FE 6*
High pay-out ratios (unconstrained)						
q	0.0291‡	0.0330‡	0.0233‡	0.0230‡	0.0333‡	0.0244‡
	(0.0048)	(0.0053)	(0.0055)	(0.0050)	(0.0049)	(0.0050)
CF	0.0802‡	0.0679‡	0.0817‡	0.0898‡	0.0790‡	0.0689‡
	(0.0102)	(0.0107)	(0.0106)	(0.0099)	(0.0013)	(0.0010)
Constant	0.135‡	0.136‡	0.140‡	0.136‡	0.131‡	−0.008‡
	(0.0044)	(0.0058)	(0.0049)	(0.0041)	(0.0042)	(0.0007)
Observations	11643	11643	11643	11643	11643	11643
Low pay-out ratios (constrained)						
q	0.1393‡	0.1432‡	0.1140‡	0.1231‡	0.2243‡	0.2209‡
	(0.0144)	(0.0213)	(0.0122)	(0.0133)	(0.0498)	(0.0478)
CF	0.0520‡	0.0462‡	0.0478‡	0.0670‡	0.0664‡	0.0625‡
	(0.0100)	(0.0079)	(0.0118)	(0.0199)	(0.0083)	(0.0090)
Constant	0.0625‡	0.0602‡	0.0850‡	0.0721‡	−0.0146‡	−0.0028‡
	(0.0115)	(0.0195)	(0.0101)	(0.0093)	(0.0450)	(0.0040)
Observations	9383	9383	9383	9383	9383	9383

†Cluster robust standard errors by firm are in parentheses.
‡$p < 0.01$.

Table 9. Investment model for unconstrained firms characterized by large and small size†

Variable	*OLS*	*FEs*	*IV 6*	*FE IV 6*	*SE 6*	*S FE 6*
Large size (unconstrained)						
q	0.0425‡	0.0484‡	0.0406‡	0.0417‡	0.0350‡	0.0343‡
	(0.0061)	(0.0058)	(0.0067)	(0.0028)	(0.0055)	(0.0042)
CF	0.0772‡	0.0571‡	0.0775‡	0.0574‡	0.0784‡	0.0578‡
	(0.0125)	(0.0119)	(0.0126)	(0.0023)	(0.0009)	(0.0002)
Constant	0.128‡	0.129‡	0.130‡	0.135‡	0.135‡	0.002‡
	(0.0059)	(0.0067)	(0.0061)	(0.0030)	(0.0050)	(0.0003)
Observations	10839	10839	10839	10839	10839	10839
Small size (constrained)						
q	0.0707‡	0.0709‡	0.0450‡	0.0250‡	0.0513‡	0.0222‡
	(0.0102)	(0.0116)	(0.0081)	(0.0057)	(0.0103)	(0.0070)
CF	0.119‡	0.142‡	0.127‡	0.154‡	0.1252‡	0.1539‡
	(0.0110)	(0.0146)	(0.0116)	(0.0153)	(0.0033)	(0.0018)
Constant	0.112‡	0.104‡	0.132‡	0.140‡	0.1273‡	−0.0048‡
	(0.0079)	(0.0099)	(0.0064)	(0.0050)	(0.0079)	(0.0007)
Observations	9805	9805	9805	9805	9805	9805

†Cluster robust standard errors by firm are in parentheses.
‡$p < 0.01$.

estimates equal to 0.123, whereas the estimators S 6 and S FE 6 report values of 0.2243 and 0.2209 respectively.

Table 9 presents the estimates of the q- and CF-coefficients for scheme 2. This scheme classifies large firms as financially unconstrained and small firms as financially constrained. The results that are reported in Table 9 suggest a greater sensitivity of cash flows to investment for constrained than for unconstrained firms. These results are in line with Fazzari *et al.* (1988),

Table 10. Investment model for unconstrained firms characterized by high and low dividend payments†

Variable	*OLS*	*FEs*	*IV 6*	*FE IV 6*	*SE 6*	*S FE 6*
High dividend payments (unconstrained)						
q	0.0333‡	0.0412‡	0.0251‡	0.0271‡	0.0369‡	0.0028
	(0.0047)	(0.0048)	(0.0051)	(0.0024)	(0.0040)	(0.0041)
CF	0.0918‡	0.0777‡	0.0946‡	0.0807‡	0.0905‡	0.0807‡
	(0.0100)	(0.0137)	(0.0102)	(0.0032)	(0.0013)	(0.0008)
Constant	0.132‡	0.130‡	0.140‡	0.143‡	0.1290‡	0.0004
	(0.0050)	(0.0068)	(0.0053)	(0.0028)	(0.0036)	(0.0005)
Observations	11292	11292	11292	11292	11292	11292
Low dividend payments (constrained)						
q	0.1362‡	0.1590‡	0.1011‡	0.1160‡	0.2179‡	0.2007‡
	(0.0136)	(0.0180)	(0.0127)	(0.0124)	(0.0251)	(0.0159)
CF	0.0504‡	0.0464‡	0.0457‡	0.0384‡	0.0614‡	0.0549‡
	(0.0080)	(0.0055)	(0.0109)	(0.0028)	(0.0033)	(0.0027)
Constant	0.0761‡	0.0571‡	0.108‡	0.0968‡	0.0023‡	−0.0035‡
	(0.0105)	(0.0167)	(0.0106)	(0.0114)	(0.0226)	(0.0010)
Observations	9597	9597	9597	9597	9597	9597

†Cluster robust standard errors by firm are in parentheses.
‡$p<0.01$.

Table 11. Investment model for unconstrained and constrained firms characterized by the Hadlock–Pierce index†

Variable	*OLS*	*FEs*	*IV 6*	*FE IV 6*	*SE 6*	*S FE 6*
Low Hadlock–Pierce index (unconstrained)						
q	0.0522‡	0.0608‡	0.0433‡	0.0460‡	0.0450‡	0.0398‡
	(0.0071)	(0.0074)	(0.0068)	(0.0035)	(0.0070)	(0.0047)
CF	0.076‡	0.057‡	0.078‡	0.058‡	0.077‡	0.065‡
	(0.0102)	(0.0114)	(0.0107)	(0.0030)	(0.0012)	(0.0004)
Constant	−0.123‡	0.131‡	0.132‡	0.136‡	0.129‡	−0.0013‡
	(0.0069)	(0.0082)	(0.0062)	(0.0037)	(0.0065)	(0.0005)
Observations	11012	11012	11012	11012	11012	11012
High Hadlock–Pierce index (constrained)						
q	0.0575‡	0.0624‡	0.0353‡	0.0127‡	0.0398‡	0.0336‡
	(0.0098)	(0.0096)	(0.0085)	(0.0085)	(0.0113)	(0.0127)
CF	0.132‡	0.142‡	0.136‡	0.1523‡	0.1357‡	0.1324‡
	(0.110)	(0.0131)	(0.0114)	(0.0061)	(0.0023)	(0.0015)
Constant	0.110‡	0.102‡	0.123‡	0.141‡	0.123‡	−0.0039‡
	(0.0078)	(0.0081)	(0.0068)	(0.0069)	(0.0087)	(0.0009)
Observations	8435	8435	8435	8435	8435	8435

†Cluster robust standard errors by firm are in parentheses.
‡$p<0.01$.

Agca and Mozumdar (2017) and most of the empirical investment literature highlighting the differences between external and internal financing for determining investment demand.

The contradictory results that are obtained from comparing Table 8 and Table 9 suggest that the criteria that are used to define financial constraints can reflect different firms' features. To provide further empirical evidence on the effect of firms' cash flows as a function of financial constraints, we use the classification that is proposed in scheme 3 based on dividend payments. The results reported in Table 10 provide empirical evidence that is similar to the analysis using

Table 12. Arellano and Bond (1991) tests for auto-correlation in the residuals†

	OLS		*FEs*		*IV*		*FE IV*	
	z	*p-value*	*z*	*p-value*	*z*	*p-value*	*z*	*p-value*
High pay-out								
1	37.56	0.000	26.33	0.000	14.52	0.000	13.98	0.000
2	21.27	0.000	8.25	0.000	10.32	0.000	5.46	0.000
3	13.69	0.000	2.57	0.764	7.93	0.000	2.13	0.000
Low pay-out								
1	35.26	0.000	22.13	0.000	9.35	0.000	12.74	0.000
2	19.34	0.000	4.51	0.000	5.40	0.000	3.62	0.000
3	15.43	0.000	−0.15	0.882	4.56	0.000	−0.41	0.685
Large size								
1	61.06	0.000	45.62	0.000	13.19	0.000	12.48	0.000
2	40.14	0.000	15.94	0.000	11.48	0.000	8.00	0.000
3	30.90	0.000	4.44	0.000	9.74	0.000	2.56	0.010
Small size								
1	34.63	0.000	21.43	0.000	11.30	0.000	10.89	0.000
2	17.88	0.000	3.69	0.000	7.73	0.000	2.68	0.007
3	11.95	0.000	−1.75	0.081	5.92	0.000	−1.78	0.076
High dividends								
1	58.61	0.000	45.32	0.000	13.13	0.000	13.41	0.000
2	36.93	0.000	16.86	0.000	10.73	0.000	8.64	0.000
3	27.14	0.000	5.79	0.000	8.81	0.000	3.51	0.000
Low dividends								
1	33.60	0.000	17.44	0.000	10.97	0.000	10.19	0.000
2	20.04	0.000	2.34	0.019	6.46	0.000	1.50	0.133
3	15.00	0.000	−0.92	0.360	5.66	0.000	−1.25	0.211
High Hadlock–Pierce index								
1	25.31	0.000	16.10	0.000	12.91	0.000	7.93	0.000
2	10.58	0.000	0.69	0.492	7.62	0.000	0.38	0.707
3	5.39	0.000	–2.92	0.003	4.68	0.000	−2.10	0.000
Low Hadlock–Pierce index								
1	46.87	0.000	30.72	0.000	12.19	0.000	5.96	0.000
2	31.63	0.000	14.07	0.000	11.27	0.000	8.70	0.000
3	25.73	0.000	6.89	0.000	9.56	0.000	4.83	0.000

†Arellano and Bond (1991) tests for auto-correlation for subsamples classified according to different criteria.

pay-out ratios. Note, for example, the parameter estimates of Tobin's q-coefficient. These findings are very supportive of the idea that cash flow of firms paying larger dividends have stronger predictive power for determining investment demand than low dividend firms' cash flows.

As a further check of robustness, we consider scheme 4 based on the Hadlock and Pierce (2010) index that classifies firms as a function of asset size and age. The classification of firms as constrained and unconstrained according to this index is similar to the classification that was obtained in scheme 2. Table 11 reports empirical results for the unconstrained and constrained firms according to this index that are consistent with the classification of firms in terms of asset size. In particular, the effect of firms' cash flows is higher (around 0.1357) for constrained firms than for unconstrained firms (0.077).

These empirical findings are robust across estimation methods but exhibit large parameter variability that challenges the validity of the various methods. To assess this formally, we compute in Table 12 the Arellano and Bond (1991) test for serial auto-correlation of the residuals that were obtained from the OLS and IV models, and the corresponding FE counterpart models obtained from the within transformation. The results show overwhelming evidence of serial correlation in the residuals and shed important doubts on the validity of the IV methodology in this context. These results provide further empirical support to choosing the alternative family of simultaneous estimators that is proposed in this paper.

As a final separate exercise to assess the relationship between investment demand, Tobin's q and firms' cash flow empirically, we study the effect of a credit supply shock on investment–cash flow sensitivities for groups of financially constrained firms. (We are grateful to a referee for suggesting this analysis.) Recent literature focuses on an alternative method to measure the effect of liquidity on investment, which exploits financial shocks in a difference-in-differences framework (see, for example, Lemmon and Roberts (2010), Almeida *et al.* (2011) and Duchin *et al.* (2010)). In this paper, we do not consider a difference-in-differences framework but, instead, compute the investment regression equation for two different periods separated by a structural break defined by the passage of the Financial Institutions Reform, Recovery, and Enforcement Act of 1989; see Lemmon and Roberts (2010). The passage of the Act led to an immediate cessation of the $12 billion annual flow of capital to speculative grade firms from savings and loans, while simultaneously forcing a sell-off of all junk bond holdings. Thus, whereas the period before 1989 is characterized by the absence of credit and liquidity constraints, after 1989 financially constrained firms faced a potentially disrupting financing environment that could have led them to use more of the cash flows to finance investment.

As our database does not contain information on firms' bond ratings, we proxy firms' financial constraints by using asset size and the pay-out ratio. (For brevity, we do not entertain constrained firms characterized by dividends and the Hadlock–Pierce index. It is worth noting, though, the strong positive correlation between the classification of firms in terms of the pay-out ratio

Table 13. Investment model for low pay-out firms before and after 1989†

Variable	*OLS*	*FEs*	*IV 6*	*FE IV 6*	*SE 6*	*S FE 6*
Before 1989						
q	0.1075‡	0.1000‡	0.0710‡	0.0427‡	0.0852‡	0.2184‡
	(0.0226)	(0.0245)	(0.0204)	(0.0128)	(0.0261)	(0.0186)
CF	0.184‡	0.131‡	0.193‡	0.131‡	0.190‡	0.0588‡
	(0.0512)	(0.0494)	(0.0565)	(0.0093)	(0.0059)	(0.0037)
Constant	0.078‡	0.097‡	0.110‡	0.151‡	0.098‡	−0.0151‡
	(0.0173)	(0.0228)	(0.0152)	(0.0123)	(0.0230)	(0.0015)
Observations	3616	3616	3616	3616	3616	3616
After 1989						
q	0.132‡	0.1615‡	0.0938‡	0.0970‡	0.2323‡	0.2184‡
	(0.0169)	(0.0396)	(0.0113)	(0.0185)	(0.0128)	(0.0186)
CF	0.044‡	0.037‡	0.127‡	0.0341‡	0.0633‡	0.059‡
	(0.0071)	(0.0088)	(0.0116)	(0.0043)	(0.0024)	(0.0037)
Constant	0.058‡	0.092‡	0.132‡	0.090‡	−0.0278‡	−0.0150‡
	(0.0131)	(0.0095)	(0.0064)	(0.0162)	(0.0111)	(0.0015)
Observations	5767	5767	5767	5767	5767	5767

†Cluster robust standard errors by firm are in parentheses.
‡$p < 0.01$.

Table 14. Investment model for small firms before and after 1989†

Variable	*OLS*	*FEs*	*IV 6*	*FE IV 6*	*SE 6*	*S FE 6*
Before 1989						
q	0.0428‡	0.0309‡	0.0174‡	−0.0002‡	0.0385‡	0.0145‡
	(0.0107)	(0.0099)	(0.0073)	(0.0075)	(0.0115)	(0.0069)
CF	0.220‡	0.303‡	0.242‡	0.324‡	0.224‡	0.292‡
	(0.0181)	(0.0200)	(0.0186)	(0.0137)	(0.0098)	(0.0047)
Constant	0.115‡	0.097‡	0.132‡	0.120‡	0.118‡	0.007‡
	(0.0092)	(0.0100)	(0.0071)	(0.0073)	(0.0076)	(0.0007)
Observations	4140	4140	4140	4140	4140	4140
After 1989						
q	0.0901‡	0.1268‡	0.0522‡	0.0384‡	0.0837‡	0.0088
	(0.0127)	(0.0160)	(0.0146)	(0.0174)	(0.0187)	(0.0214)
CF	0.097‡	0.105‡	0.104‡	0.117‡	0.098‡	0.1104‡
	(0.0108)	(0.0133)	(0.0117)	(0.0057)	(0.0036)	(0.0034)
Constant	0.096‡	0.063‡	0.124‡	0.131‡	0.1006‡	−0.0091‡
	(0.0096)	(0.0126)	(0.0107)	(0.0135)	(0.0139)	(0.0021)
Observations	5665	5665	5665	5665	5665	5665

†Cluster robust standard errors by firm are in parentheses.
‡$p < 0.01$.

and dividends, and in terms of asset size and the Hadlock–Pierce index.) Table 13 reports the estimates of the cash flow sensitivities and Tobin's q to investment by using the pay-out ratio (scheme 1) as the classification criterion. Table 14 shows the counterpart estimates before and after 1989 using firms' asset size (scheme 2).

The results corresponding to both firms' classifications are robust across estimation methods and reveal a decrease on the effect of firms' cash flows after 1989. These results provide further empirical evidence suggesting that the positive correlation between investment demand and firms' cash flows decreases in periods that are characterized by tighter financial conditions. This empirical finding is broadly consistent with the findings that were obtained from the analysis of pay-out ratios and firms' dividends, and the conclusions in Kaplan and Zingales (1997).

5. Conclusion

This paper addresses the measurement error problem arising in investment equations relating firms' investment demand with Tobin's q and cash flow. We have shown that serial correlation in the measurement error variable invalidates standard corrections based on IV methods. To solve the problem, we have proposed an alternative methodology that is based on modelling the interaction between the endogenous regressor (the average Tobin's q) and the error term as a function of lags of the endogenous regressor. The solution yields a consistent and asymptotically normal estimator that works under serial correlation of the structural equation error term and enables us to make correct statistical inference.

An application to a panel of US firms reveals stark differences between the novel estimator and existing competitors, highlighting the relevance of Tobin's q and firms' cash flow for explaining investment demand. More specifically, the S-estimator shows that the effect of these variables on predicting investment demand is greater than those estimated by the OLS and IV methodologies. These findings are in themselves evidence that the S-estimator is doing a good job of capturing the true variation in investment opportunities.

The empirical analysis also contributes to the discussion on the differences of firms' cash flow on investment demand between financially constrained and unconstrained firms. We find mixed evidence on the effect of firms' cash flows on investment demand between financially constrained and unconstrained firms. More specifically, whereas we observe higher sensitivity to cash flows for financially constrained firms defined by smaller and younger firms, we find a similar finding for the group of unconstrained firms characterized by high pay-out ratios and dividend payments. Nevertheless, the natural experiment consisting of studying the effect of a credit supply shock on the relationship between investment demand and firms' cash flow supports the view that cash flow has less predictive power under tighter financial conditions.

Appendix A: Proof of the theorems

A.1. Proof of theorem 1

We prove theorem 1 by showing that solutions to a system of equations are unique if and only if assumption 1, part (b), holds. We first note that from assumption 1, part (a), $E[\mathbf{x}_1^{\mathrm{T}}\epsilon]=0$ we have

$$E[\mathbf{x}_1^{\mathrm{T}}y]-E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]\boldsymbol{\theta}_1-E[\mathbf{x}_1^{\mathrm{T}}q]\beta=0. \tag{10}$$

From assumption 1, part (c), and the definition of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, we also obtain

$$\begin{aligned}
u&=q\epsilon-\mathbf{Z}\phi\\
&=q(y-\mathbf{x}_1\boldsymbol{\theta}_1-q\beta)-\mathbf{Z}\phi\\
&=qy-q\mathbf{x}_1(E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1}E[\mathbf{x}_1^{\mathrm{T}}y]-E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1}E[\mathbf{x}_1^{\mathrm{T}}q]\beta)-q^2\beta-\mathbf{Z}\phi\\
&=q(y-\mathbf{x}_1E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1}E[\mathbf{x}_1^{\mathrm{T}}y])-q(q-\mathbf{x}_1E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1}E[\mathbf{x}_1^{\mathrm{T}}q])\beta-\mathbf{Z}\phi\\
&=\tilde{y}-\tilde{\mathbf{x}}\boldsymbol{\theta}_2.
\end{aligned}$$

Then from $E[\tilde{\mathbf{x}}^{\mathrm{T}}u]=E[\tilde{\mathbf{x}}^{\mathrm{T}}(\tilde{y}-\tilde{\mathbf{x}}\boldsymbol{\theta}_2)]=0$, we have

$$E[\tilde{\mathbf{x}}^{\mathrm{T}}\tilde{y}]-E[\tilde{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{x}}]\boldsymbol{\theta}_2=0. \tag{11}$$

The conclusion immediately follows by equations (10) and (11) and assumption 1, part (b).

A.2. Proof of proposition 1

We first recall that $\epsilon_{it}\equiv\eta_{it}-\beta e_{it}$ such that

$$\begin{aligned}
q_{it}\epsilon_{it}&=q_{it}(\eta_{it}-\beta e_{it})\\
&=q_{it}\eta_{it}-\beta q_{it}e_{it}+E[\beta q_{it}e_{it}|\mathbf{z}_{it}]-E[\beta q_{it}e_{it}|\mathbf{z}_{it}]\\
&=g(\mathbf{z}_{it})+u_{it}
\end{aligned}$$

where $u_{it}\equiv q_{it}\eta_{it}-\beta q_{it}e_{it}-E[\beta q_{it}e_{it}|\mathbf{z}_{it}]$, and $g(\mathbf{z}_{it})\equiv E[\beta q_{it}e_{it}|\mathbf{z}_{it}]\neq 0$ because q_{it}^* is auto-correlated. In addition, by assumption $g(\mathbf{z}_{it})$ can be approximated by a polynomial of order m; hence we have that $q_{it}\epsilon_{it}=\mathbf{Z}_{it}\phi+u_{it}$.

Now, from the above equation, it remains to show that $E[\tilde{\mathbf{x}}_{it}^{\mathrm{T}}u_{it}]=0$ where $\tilde{\mathbf{x}}_{it}=(q_{it}^2,\mathbf{Z}_{it})$ with $\mathbf{Z}_{it}=(1,q_{it-1},q_{it-1}^2,\ldots,q_{it-1}^m)$. We do this by proving $E[u_{it}|q_{it},\mathbf{z}_{it}]=0$ because this implies that $E[\tilde{\mathbf{x}}_{it}^{\mathrm{T}}u_{it}]=0$. From the definition of u_{it} above

$$\begin{aligned}
E[u_{it}|q_{it},\mathbf{z}_{it}]&\equiv E[q_{it}\eta_{it}-\beta q_{it}e_{it}-E[\beta q_{it}e_{it}|\mathbf{z}_{it}]|q_{it},\mathbf{z}_{it}]\\
&=q_{it}E[\eta_{it}|q_{it},\mathbf{z}_{it}]-E[\beta q_{it}e_{it}|q_{it},\mathbf{z}_{it}]-E[\beta q_{it}e_{it}|\mathbf{z}_{it}]\\
&=q_{it}E[\eta_{it}|q_{it},\mathbf{z}_{it}]-\beta q_{it}(E[e_{it}|q_{it},\mathbf{z}_{it}]-E[e_{it}|\mathbf{z}_{it}])\\
&=0,
\end{aligned}$$

since $E[\eta_{it}|q_{it},\mathbf{z}_{it}]=0$ by the zero-mean white noise process of η_{it}, and since $E[e_{it}|q_{it},\mathbf{z}_{it}]=E[e_{it}|\mathbf{z}_{it}]$ by the fact that by assumption 2 $e_{it}=h(e_{it-1},q_{it-1}^*)$ implies that $e_{it}=\tilde{h}(q_{it-1})$ for a measurable function $\tilde{h}(\cdot)$. This completes the proof.

A.3. Proof of theorem 2

We show the asymptotic distribution of the proposed estimator by taking into account the problem of the generated variables. This is an extension of the problem of the generated regressor in Pagan (1984). Recall that

$$\hat{\boldsymbol{\theta}}_1 = \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}y_{it} - \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}q_{it}\right)\hat{\beta}$$

and

$$\hat{\boldsymbol{\theta}}_2 = \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{\mathbf{x}}_{it}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{y}_{it}.$$

Also recall that $\tilde{\mathbf{x}} \equiv (q(q-\mathbf{x}_1\delta_q), \mathbf{Z})$ with $\delta_q \equiv E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1}E[\mathbf{x}_1^{\mathrm{T}}q]$, and $\tilde{y} \equiv q(y-\mathbf{x}_1\delta_y)$ with $\delta_y \equiv E[\mathbf{x}_1^{\mathrm{T}}\mathbf{x}_1]^{-1} \times E[\mathbf{x}_1^{\mathrm{T}}y]$. In addition, define the sample analogues $\hat{\mathbf{x}} \equiv (q(q-\mathbf{x}_1\hat{\delta}_q), \mathbf{Z})$ with

$$\hat{\delta}_q \equiv \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}q_{it},$$

and $\hat{y} \equiv q(y-\mathbf{x}_1\,\hat{\delta}_y)$ with

$$\hat{\delta}_y \equiv \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}y_{it}.$$

We first show the asymptotic properties of $\hat{\boldsymbol{\theta}}_2$. We note that, since $\tilde{y}_{it} = \tilde{\mathbf{x}}_{it}\boldsymbol{\theta}_2 + u_{it}$, we have

$$\hat{y}_{it} = \hat{\mathbf{x}}_{it}\boldsymbol{\theta}_2 + u_{it} - (\hat{\mathbf{x}}_{it} - \tilde{\mathbf{x}}_{it})\boldsymbol{\theta}_2 + \hat{y}_{it} - \tilde{y}_{it}.$$

By plugging the equation above into the definition of $\hat{\boldsymbol{\theta}}_2$, we obtain

$$\hat{\boldsymbol{\theta}}_2 = \boldsymbol{\theta}_2 + \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{\mathbf{x}}_{it}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\{u_{it} - (\hat{\mathbf{x}}_{it} - \tilde{\mathbf{x}}_{it})\boldsymbol{\theta}_2 + (\hat{y}_{it} - \tilde{y}_{it})\},$$

so that

$$\begin{aligned}\sqrt{n}(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2) &= \hat{Q}^{-1}\sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\{u_{it} - (\hat{\mathbf{x}}_{it} - \tilde{\mathbf{x}}_{it})\boldsymbol{\theta}_2 + (\hat{y}_{it} - \tilde{y}_{it})\} \\ &=: \hat{Q}^{-1}(A_1 - A_2 + A_3) \end{aligned} \tag{12}$$

with $\hat{Q} \equiv (1/n)\Sigma_{i=1}^{n}\Sigma_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{\mathbf{x}}_{it}$, where $A_1 \equiv \sqrt{n}(1/n)\Sigma_{i=1}^{n}\Sigma_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}u_{it}$, $A_2 \equiv \sqrt{n}(1/n)\Sigma_{i=1}^{n}\Sigma_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}(\hat{\mathbf{x}}_{it} - \tilde{\mathbf{x}}_{it})\boldsymbol{\theta}_2$ and $A_3 \equiv \sqrt{n}(1/n)\Sigma_{i=1}^{n}\Sigma_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}(\hat{y}_{it} - \tilde{y}_{it})$. We note that

$$\hat{Q} \equiv \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\mathbf{x}}_{it}^{\mathrm{T}}\hat{\mathbf{x}}_{it} \xrightarrow{\mathrm{p}} E[\tilde{\mathbf{x}}_{it}^{\mathrm{T}}\tilde{\mathbf{x}}_{it}] \equiv Q, \tag{13}$$

by Slutsky's theorem and the law of large numbers. We also note that, by a Taylor series expansion,

$$A_1 = \sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{\mathbf{x}}_{it}^{\mathrm{T}}u_{it} + \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\nabla_{\delta_q}\tilde{\mathbf{x}}_{it}^{\mathrm{T}}u_{it}\right)\sqrt{n}(\hat{\delta}_q - \delta_q) + o_p(1),$$

with $\nabla_{\delta_q}\tilde{\mathbf{x}}_{it} = (-q\mathbf{x}_1, \mathbf{0}_{k\times 1}^{\mathrm{T}})$. Since $E[u_{it}|\tilde{\mathbf{x}}_{it}] = 0$, we have $E[\nabla_{\delta_q}\tilde{\mathbf{x}}_{it}^{\mathrm{T}}u_{it}] = 0$. Thus, we obtain

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\nabla_{\delta_q}\tilde{\mathbf{x}}_{it}^{\mathrm{T}}u_{it} = o_p(1).$$

Since $\sqrt{n}(\hat{\delta}_q - \delta_q) = O_p(1)$, it follows that

$$A_1 = \sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{\mathbf{x}}_{it}^{\mathrm{T}}u_{it} + o_p(1). \tag{14}$$

By a similar argument, we have that

$$A_2 = \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}^{\mathrm{T}} \nabla_{\delta_q} \tilde{\mathbf{x}}_{it} \boldsymbol{\theta}_2 \right) \sqrt{n}(\hat{\delta}_q - \delta_q) + o_p(1) = \sqrt{n} G(\hat{\delta}_q - \delta_q) + o_p(1),$$

with $G = E[\tilde{\mathbf{x}}_{it}^{\mathrm{T}} \nabla_{\delta_q} \tilde{\mathbf{x}}_{it} \boldsymbol{\theta}_2]$. Since

$$\hat{\delta}_q \equiv \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} \mathbf{x}_{1it} \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} q_{it}$$

and $\delta_q \equiv E[\mathbf{x}_1^{\mathrm{T}} \mathbf{x}_1]^{-1} E[\mathbf{x}_1^{\mathrm{T}} q]$, A_2 can be rewritten as

$$A_2 = G \sqrt{n} \sum_{i=1}^{n} \sum_{t=1}^{T} r_{it}(\delta_q) + o_p(1), \tag{15}$$

with

$$r_{it}(\delta_q) = \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} \mathbf{x}_{1it} \right)^{-1} \mathbf{x}_{1it}^{\mathrm{T}} (q_{it} - \mathbf{x}_{1it} \delta_q)$$

and $E[r_{it}(\delta_q)] = 0$ by the law of iterated expectations. Similarly, we obtain

$$A_3 = \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}^{\mathrm{T}} \nabla_{\delta_y} \tilde{y}_{it} \right) \sqrt{n}(\hat{\delta}_y - \delta_y) + o_p(1) = \sqrt{n} H(\hat{\delta}_y - \delta_y) + o_p(1),$$

with $\nabla_{\delta_y} \tilde{y}_{it} = -q \mathbf{x}_1$ and $H = E[\tilde{\mathbf{x}}_{it}^{\mathrm{T}} \nabla_{\delta_y} \tilde{y}_{it}]$. Since

$$\hat{\delta}_y \equiv \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} \mathbf{x}_{1it} \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} y_{it}$$

and $\delta_y \equiv E[\mathbf{x}_1^{\mathrm{T}} \mathbf{x}_1]^{-1} E[\mathbf{x}_1^{\mathrm{T}} y]$, A_3 can be rewritten as

$$A_3 = H \sqrt{n} \sum_{i=1}^{n} \sum_{t=1}^{T} s_{it}(\delta_y) + o_p(1), \tag{16}$$

with

$$s_{it}(\delta_y) = \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{1it}^{\mathrm{T}} \mathbf{x}_{1it} \right)^{-1} \mathbf{x}_{1it}^{\mathrm{T}} (y_{it} - \mathbf{x}_{1it} \delta_y),$$

and $E[s_{it}(\delta_y)] = 0$ by the law if iterated expectations.

By plugging equations (13)–(16) into equation (12), we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2) = Q^{-1} \left[\sqrt{n} \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \{ \tilde{\mathbf{x}}_{it}^{\mathrm{T}} u_{it} - G r_{it}(\delta_q) + H s_{it}(\delta_y) \} \right] + o_p(1).$$

As a result, we show the consistency of $\hat{\boldsymbol{\theta}}_2$:

$$\hat{\boldsymbol{\theta}}_2 \xrightarrow{\mathrm{p}} \boldsymbol{\theta}_2 + Q^{-1} \times 0 = \boldsymbol{\theta}_2.$$

To show asymptotic normality, we have that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2) \xrightarrow{\mathrm{d}} Q^{-1} N(0, M) \equiv N(0, Q^{-1} M Q^{-1}),$$

with $M = \mathrm{var}\{ \tilde{\mathbf{x}}_{it}^{\mathrm{T}} u_{it} - G r_{it}(\delta_q) + H s_{it}(\delta_y) \}$, by the Lindeberg–Lévy central limit theorem.

We now show the asymptotic properties of $\hat{\boldsymbol{\theta}}_1$. We note that

$$\hat{\boldsymbol{\theta}}_1 = \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}(\mathbf{x}_{1it}\boldsymbol{\theta}_1 + q_{it}\beta + \epsilon)$$
$$-\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}q_{it}\right)\hat{\beta}$$
$$= \boldsymbol{\theta}_1 - \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}q_{it}\right)(\hat{\beta}-\beta) + o_p(1).$$

By the law of large numbers we have

$$\hat{\mathbf{C}}_1 \equiv \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it} \xrightarrow{\mathrm{p}} E[\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}] \equiv \mathbf{C}_1,$$
$$\hat{C}_2 \equiv \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}q_{it} \xrightarrow{\mathrm{p}} E[\mathbf{x}_{1it}^{\mathrm{T}}q_{it}] \equiv C_2,$$

so that we show the consistency of $\hat{\boldsymbol{\theta}}_1$:

$$\hat{\boldsymbol{\theta}}_1 \xrightarrow{\mathrm{p}} \boldsymbol{\theta}_1 - \mathbf{C}_1^{-1}C_2Q_{\beta}^{-1}\times 0 = \boldsymbol{\theta}_1,$$

where Q_β is the element in the Q-matrix that corresponds to the estimation of β. To show asymptotic normality, we note that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1) = -\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}\mathbf{x}_{1it}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbf{x}_{1it}^{\mathrm{T}}q_{it}\right)\sqrt{n}(\hat{\beta}-\beta).$$

Then, we obtain

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1) \xrightarrow{\mathrm{d}} \mathbf{C}_1^{-1}C_2N(0, V_{\hat{\beta}}) \equiv N(0, \mathbf{C}_1^{-1}C_2V_{\hat{\beta}}C_2^{\mathrm{T}}\mathbf{C}_1^{-1}),$$

where $V_{\hat{\beta}}$ is the variance of $\hat{\beta}$.

Appendix B: Inconsistency of instrumental variables estimator in the presence of serial persistence

Consider, for instance, auto-regressive models of order 1 (AR(1)) for driving the marginal Tobin's q:

$$q_{it}^* = \rho^q q_{it-1}^* + w_{it}^q, \tag{17}$$

where $|\rho^q| < 1$ and w_{it}^q is an independent and identically distributed and zero-mean process, and the measurement errors

$$e_{it} = h(e_{it-1}, q_{it-1}^*), \tag{18}$$

as stated in proposition 1. (There is no constant term in the process for q_{it}^* without loss of generality.)

From equation (3), the endogenous variable q_{it} can be rewritten as

$$\begin{aligned} q_{it} &= q_{it}^* + e_{it} \\ &= \rho^q q_{it-1}^* + w_{it}^q + e_{it} \\ &= \rho^q q_{it-1} + w_{it}^q + h(e_{it-1}, q_{it-1}^*). \end{aligned} \tag{19}$$

As a result, q_{it} is correlated with q_{it-1} so q_{it-1} satisfies the instrument relevance condition. Namely, from equation (19), q_{it-1} is

$$q_{it-1} = \rho^q q_{it-2} + w_{it-1}^q + h(e_{it-2}, q_{it-2}^*).$$

If $\{q_{it-2}, w_{it-1}^q, q_{it-2}^*, e_{it-2}\}$ are uncorrelated with $\epsilon_{it} \equiv \eta_{it} - \beta e_{it}$, the variable q_{it-1} satisfies the instrument exogeneity condition. Then, IV methods which employ q_{it-1} as an instrument produce consistent estimators of β and γ. We note that higher order lags of q_{it} also satisfy the exogeneity condition and, hence, can be also used as IVs.

However, in practice, it is highly likely that current period measurement error is correlated with the first-order or higher order lags of the measurement error. This is so if λ_{it} or ν_{it} exhibit some persistence. For example, consider the process of the measurement errors as in equation (18). The measurement errors are persistent in the sense that the last period measurement error e_{it-1} affects the current period e_{it}. In this case, the instrument exogeneity condition is no longer valid because the instrument $q_{it-1} = \rho^q q_{it-2} + w^q_{it-1} + h(e_{it-2}, q^*_{it-2})$ is correlated with the error term,

$$\epsilon_{it} \equiv \eta_{it} - \beta e_{it} = \eta_{it} - \beta h(e_{it-1}, q^*_{it-1}), \tag{20}$$

through the lag of Tobin's q, q^*_{it-1}, and the lag of the measurement error, e_{it-1}. As a result, the IV approach that was proposed by Almeida *et al.* (2010) also fails to obtain consistent parameter estimates.

This is partly addressed in the study that was carried out by Agca and Mozumdar (2017) by including longer lags of the instruments in the dynamic GMM approaches that were employed by Cummins *et al.* (2006) and Almeida *et al.* (2010). However, the dynamic GMM approaches still fail to work if the measurement errors are persistent. For instance, if we assume, instead, the second-order lag of the average Tobin's q as instrument, $q_{it-2} = \rho^q q_{it-3} + w^q_{it-2} + h(e_{it-3}, q^*_{it-3})$, then, from equation (20), we have

$$\epsilon_{it} = \eta_{it} - \beta h(e_{it-1}, q^*_{it-1}) = \eta_{it} - \beta h\{h(e_{it-2}, q^*_{it-2}), q^*_{it-1}\}.$$

This algebra reveals that the IV is still correlated with the regression error term ϵ_{it} through q^*_{it-2} and e_{it-2}. In fact, the dynamic GMM approach is more efficient in that it provides smaller variance of the estimator by imposing a proper weighting matrix when the error term is heteroscedastic and auto-correlated. However, the estimator is inconsistent if the IVs are correlated with the error term. Thus, employing higher order lags of the IV and imposing a proper weighting matrix are not sufficient conditions to control for measurement errors in the average Tobin's q.

References

Abel, A. B. and Eberly, J. C. (1994) A unified model of investment under uncertainty. *Am. Econ. Rev.*, **84**, 1369–1384.

Agca, S. and Mozumdar, A. (2017) Investment cash flow sensitivity: fact or fiction? *J. Finan. Quant. Anal.*, **3**, 1111–1141.

Almeida, H., Bruno Laranjeira, M. C. and Weisbenner, S. (2011) Corporate debt maturity and the real effects of the 2007 credit crisis. *Crit. Finan. Rev.*, **1**, 3–58.

Almeida, H. and Campello, M. (2007) Financial constraints, asset tangibility and corporate investment. *Rev. Finan. Stud.*, **20**, 1429–1460.

Almeida, H., Campello, M. and Galvao, A. (2010) Measurement errors in investment equations. *Rev. Finan. Stud.*, **23**, 3279–3328.

Almeida, H., Campello, M. and Weisbach, M. S. (2004) The cash flow sensitivity of cash. *J. Finan.*, **59**, 1777–1804.

Arellano, M. and Bond, S. (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev. Econ. Stud.*, **58**, 277–279.

Bernanke, B. and Gertler, M. (1989) Agency costs, net worth, and business fluctuations. *Am. Econ. Rev.*, **79**, 14–31.

Bertrand, M. and Mullainathan, S. (2005) Bidding for oil and gas leases in the Gulf of Mexico: a test of the free cash flow model? *Mimeo*. University of Chicago, Chicago.

Bertrand, M. and Schoar, A. (2003) Managing with style: the effect of managers on firm policies. *Q. J. Econ.*, **118**, 1169–1208.

Chen, X. (2007) Large sample sieve estimation of semi-nonparametric models. In *Handbook of Econometrics*, vol. 6B (eds J. J. Heckman and E. E. Leamer). Amsterdam: North-Holland.

Cummins, J. G., Hasset, K. A. and Oliner, S. D. (2006) Investment behavior, observable expectations, and internal funds. *Am. Econ. Rev.*, **96**, 796–810.

Duchin, R., Ozbas, O. and Sensoy, B. A. (2010) Costly external finance, corporate investment, and the subprime mortgage credit crisis. *J. Finan. Econ.*, **97**, 418–435.

Erickson, T. and Whited, T. (2000) Measurement error and the relationship between investment and *Q*. *J. Polit. Econ.*, **108**, 1027–1057.

Fazzari, S., Hubbard, R. G. and Petersen, B. (1988) Financing constraints and corporate investment. *Brook. Paps Econ. Activ.*, **1**, 141–195.

Galvao, A. F., Montes-Rojas, G. and Song, S. (2017) Endogeneity bias modeling using observables. *Econ.Lett.*, **152**, 41–45.
Hadlock, C. (1998) Ownership, liquidity, and investment. *RAND J. Econ.*, **29**, 487–508.
Hadlock, C. J. and Pierce, J. R. (2010) New evidence on measuring financial constraints: moving beyond the KZ index. *Rev. Finan. Stud.*, **23**, 1909–1940.
Hayashi, F. (1982) Tobin's marginal q and average q: a neoclassical interpretation. *Econometrica*, **50**, 213–224.
Holmstrom, B. and Tirole, J. (1998) Private and public supply of liquidity. *J. Polit. Econ.*, **106**, 1–40.
Hoshi, T., Kashyap, A. and Scharfstein, D. (1991) Corporate structure, liquidity, and investment: evidence from Japanese industrial groups. *Q. J. Econ.*, **106**, 33–60.
Kaplan, S. and Zingales, L. (1997) Do financing constraints explain why investment is correlated with cash flow? *Q. J. Econ.*, **112**, 169–215.
Kind, R. and Levine, R. (1993) Finance and growth: Schumpeter might be right. *Q. J. Econ.*, **108**, 717–737.
Kiyotaki, N. and Moore, J. (1997) Credit cycles. *J. Polit. Econ.*, **105**, 211–248.
Lamont, O. (1999) Cash flow and investment: evidence from internal capital markets. *J. Finan.*, **52**, 83–110.
Lemmon, M. and Roberts, M. R. (2010) The response of corporate financing and investment to changes in the supply of credit. *J. Finan. Quant. Anal.*, **45**, 555–587.
Lewellen, J. and Lewellen, K. (2016) Investment and cashflow: new evidence. *J. Finan. Quant. Anal.*, **51**, 1135–1164.
Modigliani, F. and Miller, M. (1958) The cost of capital, corporation finance and the theory of investment. *Am. Econ. Rev.*, **48**, 261–297.
Moyen, N. (2004) Investment–cash flow sensitivities: constrained versus unconstrained firms. *J. Finan.*, **59**, 2061–2092.
Pagan, A. (1984) Econometric issues in the analysis of regressions with generated regressors. *Int. Econ. Rev.*, **25**, 221–247.
Poterba, J. (1988) Financing constraints and corporate investment: comment. *Brook. Paps Econ. Activ.*, **1**, 200–204.
Shin, H. and Stulz, R. (1998) Are internal capital markets efficient? *Q. J. Econ.*, **113**, 531–552.
Stein, J. (2003) Agency information and corporate investment. In *Handbook of the Economics of Finance* (eds M. Harris and R. Stulz). Amsterdam: North-Holland.
Whited, T. M. and Wu, G. (2006) Financial constraints risk. *Rev. Finan. Stud.*, **19**, 531–559.