Panel Cointegration Estimates of the Effect of Interest Rates, Capital Goods Prices, and Taxes on the Capital Stock

May 25, 2007

(Preliminary and incomplete; not for citation; comments welcome)

Abstract: The effect of interest rates, capital goods prices, and taxes on the capital stock is an issue of central importance in economics, with implications for monetary policy, business cycle models, tax policy, economic development, growth, and other areas. For more than 30 years it has been difficult to obtain precise estimates of these effects, and there is little consensus in the profession on their magnitude, despite their importance for both theory and policy. In this paper, we therefore turn to panel data, specifically a newly constructed data set with more than 50 years of firm-level data on the capital stock and with detailed industry-specific data on the interest rate, the price of investment goods, and tax parameters. Using this rich panel data set, we implement recently developed tests for cointegration in panel data. These tests allow us to determine whether the long-run implications of Jorgensonian neoclassical, q, irreversibility, and (s,S) theories are supported by the data. Using the same data, we then use recently developed panel cointegration estimators to assess the quantitative effect of the interest rate, capital goods prices, and taxes on the capital stock.

Keywords: investment, capital stock, user cost elasticity, interest rate, taxes, capital goods prices, panel cointegration

JEL codes: H25,C23,E22,E62

Professor Huntley Schaller, Carleton University and Institute for Advanced Studies (Vienna)

Professor Marcel Voia, Carleton University

Mailing Address:
Department of Economics
Carleton University
1125 Colonel By Drive
Ottawa ON Canada K1S 5B6

Email: Schaller: schaller@ccs.carleton.ca, Voia: mvoia@connect.carleton.ca
Tel: Schaller: (613) 520-3751, Voia: (613) 520-2600 ext. 3546

We would like to thank Mark Blanchette for excellent research assistance. Schaller thanks MIT for providing an excellent environment in which to begin this research and the SSHRC for financial support.

1 Introduction

According to neoclassical growth theory, the capital stock is one of the main determinants of the long-run standard of living. In some versions of endogenous growth theory, the capital stock plays an even more important role by influencing the rate of economic growth.

According to standard economic theory, the long-run capital stock is determined by the interest rate, taxes, and capital goods prices. The quantitative magnitude of these effects is of crucial importance for many areas of economics, including monetary policy, business cycle models, tax policy, trade, economic development, and growth. Unfortunately, there is little consensus on the magnitude of these effects. For example, Chirinko (1993) concludes that "the response of investment to price variables tends to be small and unimportant relative to quantity variables," while Hassett and Hubbard (2002) conclude that the user cost elasticity is probably between -0.5 and -1.

Caballero (1994, 1999) and Schaller (2006) argue that there are serious problems in obtaining unbiased estimates of user cost elasticity from short-run movements in investment, as the great majority of the previous literature has tried to do. Empirical researchers are trying to estimate the elasticity of the demand for capital, but the equilibrium quantity and price depend on both supply and demand. At business cycle frequencies, there are substantial movements in demand. If the supply curve for capital is upward sloping in the short run, as we believe most supply curves are, econometric methods that emphasize high-frequency fluctuations in the data will tend to pick up movements along this supply curve, biasing the elasticity toward more positive values.

On the basis of these economic issues – and their implications for the appropriate econometric techniques – Caballero (1994, 1999) and Schaller (2006) argue that it will be possible to obtain better estimates of user cost elasticity by using low-frequency movements in the variables. To see this, note that shifts in the supply curve for capital are probably due primarily to technological change and productivity shocks, which tend to have persistent effects on the price of investment goods and the real interest rate, and tax reforms, which also tend to be relatively persistent. This implies that techniques that emphasize low-frequency movements will tend to trace out points on the demand curve for capital while techniques that emphasize high-frequency movements are more likely to trace out points on the supply curve.

There are two other ways to address the simultaneity problem that plagues estimates of the user cost elasticity. Schaller (2006) points out that the supply curve for capital in a small, open economy will be flat. Theoretically, this should help to reduce the bias towards positive estimates of the elasticity. Empirically, he finds that the estimated elasticity is substantially larger for a small, open economy (about -1.4) than the Caballero (1994) estimate for a large economy (about -0.9).

The second way to address the simultaneity problem is to use panel data. Theoretically, the simultaneity problem in a large economy arises because shocks to capital demand affect the equilibrium interest rate. To the extent that shocks to capital demand at the firm level are partially idiosyncratic, the use of firm-level panel data should reduce the simultaneity problem. The intuition is the same as that for a small, open economy. Just as shocks to capital demand in a small economy have little effect on the world interest rate, idiosyncratic shocks to capital demand in a particular firm have little effect on the country's interest rate. To the best of our knowledge, this is the first paper to use firm-level data to estimate the user cost elasticity with panel cointegration techniques.

The simultaneity problem provides a strong reason for using cointegration techniques, but there is a second important reason. Economic theories make quite different predictions about investment dynamics.[1] However, a wide variety of theories predict the same long-run relationship between the capital stock and user cost. Again, this suggests that better estimates can be obtained by using techniques that focus on low-frequency movements in the data.

The simultaneity problem is one reason that we turn to panel data. A second motivation is the traditional one: more variation in the data is usually helpful in obtaining better estimates. In panel data, there is considerably more variation – both in the capital/output ratio and user cost – than in aggregate data. For example, the weighted average cost of capital differs across firms because of differences in the relative importance of debt and equity and because of cross-sectional differences in risk. In fact, the additional cross-sectional variation seems to help: our standard errors are generally quite small. There is also Monte Carlo evidence that panel cointegration techniques help modestly in reducing small sample bias in the estimation of cointegrating relationships. According to the existing evidence, the biggest reduction in bias comes from using at least a few cross-sectional units, instead of a single time series.[2]

---

[1] In the neoclassical model without adjustment costs, the capital stock will respond immediately to shocks. In a Q model with convex adjustment costs, the transition path to the new steady state will depend on whether shocks are anticipated (or realized) and transitory (or persistent). In a model with irreversibility at the micro level, the estimated short-run elasticity at a higher level of aggregation will depend on the sequence of previous shocks and the cross-sectional distribution at a lower level of aggregation (e.g., at the plant level) of the gap between the desired and actual capital stock.

[2] See Kao and Chiang (2000).

A more important reason for our recourse to panel data is economic, rather than purely econometric. There are theoretical reasons for suspecting that user cost elasticity is more complex than a single production function parameter (as might be the case if capital markets were frictionless, managers always acted purely in the interest of their shareholders, and fixed adjustment costs never led to nonconvexities). For example, in a model with investment irreversibility, Bertola and Caballero (1994) show that there will be an "irreversibility premium" that will be added to the usual discount rate. We know little about how large this irreversibility premium is and how it covaries with the observable market interest rate. By making cross-sectional comparisons, we can get some understanding of how nonconvexities affect the long-run response of the capital stock to user cost. Similar issues arise from financial market imperfections. As Fazzari, Hubbard, and Petersen (1988) point out, asymmetric information can lead to a "lemons premium" that drives a wedge between the interest rate and the shadow cost of finance. Again, we know relatively little about the time series behaviour – or even the quantitative importance – of the lemons premium. Influential papers have argued that financial market imperfections are an issue of first-order importance for both macroeconomic fluctuations and economic growth.[3] On the other hand, prominent papers have argued that there are serious flaws in the main existing evidence for important financial market imperfections (such as finance constraints).[4] Again, by using cross-sectional comparisons, we may be able to get a better understanding of how financial market imperfections affect the accumulation of capital and economic growth.

[3] See, e.g., Bernanke and Gertler (1989) and Jermann and Quadrini (2003). {To be supplemented by additional references.}
[4] See, e.g., Kaplan and Zingales (1997), Gomes (2001), and Erickson and Whited (2000).

In the empirical work, we make use of an unusually rich panel data set. Our panel data covers the period 1964-2004 and includes firm-level data on the capital stock and output. To get a sense of how long this time dimension is for *panel* data, note that Caballero (1994) uses 31 years of data, while Schaller (2006) uses 38 years of data. Both of these studies use *aggregate* time series data. The firm-level data has been linked to industry-specific data on variables like prices, risk, taxes, and depreciation. The data includes the carefully constructed, firm-specific weighted average cost of capital, rather than an aggregate measure of the cost of capital. The firm-level cost of capital takes into account variation in risk using standard techniques from financial economics. In addition, careful attention has been paid to tax parameters (including industry-specific measures of the present value of depreciation allowances per dollar of capital spending), which are based on state-of-art work by Dale Jorgenson.

The paper uses recently developed econometric techniques for non-stationary panel data, including recently developed panel unit root tests, tests for cointegration in panel data, and panel cointegrating regression estimators. These techniques are discussed in more detail in the section on empirical results.

The paper is organized as follows. Section 2 provides a review of previous estimates of user cost elasticity. Section 3 describes the data. Section 4 describes the cointegrating relationship between capital and user cost, including empirical evidence. Section 5 discusses econometric issues, including the potential for small sample bias if OLS is used to estimate the cointegrating relationship (and how we can overcome this bias) and the potential bias due to the presence of fixed effects. Section 6 presents estimates of the user cost elasticity for all types of firms. Sections 7, 8, and 9 estimate the user cost elasticity for firms that are more likely to face finance constraints, corporate governance problems, and binding irreversibility constraints, respectively.

2 Review of Previous Estimates

2a Studies that emphasize high-frequency movements in the data

The great majority of previous estimates of user cost elasticity come from studies that emphasize high-frequency movements in the data. Within these studies, there is a tendency to find relatively low values of user cost elasticity. (Throughout this paper, we will use "low" to refer to user cost elasticities that are close to zero.) But there is also considerable variation in elasticity estimates based on these types of studies. For example, Cummins and Hassett (1992) obtain an estimate of slightly more than -1 using US firm-level data. In contrast, Chirinko, Fazzari, and Meyer (1999) obtain a preferred elasticity of about -0.25, again using US firm-level data.[5] Clark (1993) finds an estimated elasticity of -0.01 using aggregate US data, while Tevlin and Whelan (2003) estimate the user cost elasticity at -0.18, also using aggregate US data. In a slightly different type of study, Goolsbee (2000) finds that a 10% investment tax credit raises investment by about 4 to 5%, using US data that is differentiated by the type of asset. Using aggregate data for Japan, Kiyotaki and West (1996) obtain an estimate of -0.05 to -0.07.

2b Studies that emphasize low-frequency movements in the data

[5] When the same authors try to avoid the problems with econometric methods that emphasize high-frequency variation in the data (but without estimating the cointegrating relationship), they obtain a slightly larger elasticity estimate of -0.4. See Chirinko, Fazzari, and Meyer (2001).

Studies that are based on the cointegrating relationship between the capital stock and user cost tend to emphasize low-frequency movements in the data. The pioneering study of this type is Caballero (1994), which obtained an estimate of about -0.9 using aggregate US data for equipment capital. Using data from a small, open economy (Canada) and Dynamic OLS (DOLS), Schaller (2006) obtains a preferred estimate of about -1.4 for equipment and 0 for structures. Like Caballero (1994), Schaller (2006) uses aggregate data.

At the time we started this paper, the only previous study that estimates the cointegrating relationship between the capital stock and user cost using disaggregated data was Caballero, Engel, and Haltiwanger (1995), which uses data for 20 two-digit SIC industries. Unlike our paper, however, they do not treat the time-series-cross-sectional data as a panel. (The first papers describing panel cointegration estimators were published after their work.) Instead, they treat each industry as an individual time series. Using this approach, they obtain estimates that range from -0.01 for transportation to -2.0 for textiles.[6] The wide dispersion of estimated elasticities across industries is striking. In this paper, we test several explanations that might account for differences in user cost elasticity across different types of firms.

3 Data

The capital stock is constructed using a standard perpetual inventory method and is based primarily on firm-level financial statement data from CompuStat. Output is calculated using firm-specific sales data from CompuStat. This firm-specific data is

[6] Interestingly, when they estimate the corresponding short-run elasticities, they obtain estimates that are much smaller, typically only one-tenth of the corresponding long-run elasticity estimated using the cointegrating relationship.

linked with the Bureau of Economic Analysis data that we use to construct sector-specific, time-varying depreciation rates and capital goods price indexes.

User cost is calculated as follows

$$\tilde{R}_{f,t} = (r_{f,t} + \delta_{s,t})\left(\frac{1 - u_{s,t} - ITC_{s,t}}{1 - \tau_t}\right)\frac{p^K_{s,t}}{p^Y_{s,t}} \quad (1)$$

where r is the real, risk-adjusted interest rate, u is the sector-specific present value of depreciation allowances, ITC is the sector-specific investment tax credit rate, τ is the corporate tax rate, p^K is the price of capital goods, and p^Y is the price of output. The data on u were provided by Dale Jorgenson.

The real interest rate is calculated using a weighted average of the costs of debt and equity (with sector-specific leverage ratios). We adjust for differences in risk using a standard CAPM technique (with sector-specific CAPM βs).

In general, the data is of very high quality because it comes directly from firms' financial statements, but our work and the previous literature have identified certain instances in which data problems can arise. We deal with this in two ways. First, our own careful analysis of the data showed that there are more frequent data problems with extremely small firms, so firms with initial book value of capital of less than \$1,000,000 are omitted. Second, the first observation for each firm is excluded. Third, consistent with other papers that use firm-specific panel data, we trim extreme observations as a way of removing data that is contaminated by accounting problems (e.g., those that arise from acquisitions) and reporting errors. Further details are provided in the Data Appendix.

Table 1 reports summary statistics for the full sample.

4 Cointegrating Relationship

Suppose that

$$k_t = \alpha_0 + \alpha_R R_t + u_t \quad (2)$$
$$\Delta R_t = v_t$$

where k is the log capital/output ratio, R (=ln $\tilde{R}$) is the log of user cost, and u and v are stationary.[7] [8] The variables k and R will then be cointegrated.

Cointegration between k and R is a good description of the data. First, Levin-Lin-Chu (2002) panel unit root tests suggest unit roots in both k and R, as shown in Table 2. The Levin-Lin-Chu test is a one-sided test, so a sufficiently large negative value of t* would lead to rejection of the null hypothesis of a unit root. In fact, in our data, the p-value is 1.00 for both k and R, so there is no evidence against the existence of a unit root. We focus on the Levin-Lin-Chu test because it is very popular, but it does not allow for cross-sectional dependency. We therefore also report the results of a second panel unit root test [Chang (2004)], which allows for cross-sectional dependency. As shown in Table 3, the results are very similar, yielding p-values of 1.00 for the null hypothesis of a unit root for both k and R.

Second, panel cointegration tests show that k and R are cointegrated. We consider three different panel cointegration tests that have been proposed by Kao (1999) – the Kao Dickey-Fuller test, the Kao Phillips-Perron test, and the Kao Augmented

[7] As discussed by Caballero (1999, p. 816-821), this relationship can be obtained by solving the firm's problem (under the consumption of Cobb-Douglas technology) for the frictionless capital stock and then relaxing the unit user cost elasticity constraint.
[8] To keep the notation simple and straightforward, we only include the time subscript (suppressing the firm subscript) in this section and the next section (where we provide intuition for small sample bias and how dynamic OLS reduces the bias).

Dickey-Fuller test. As shown in Table 4, all three tests strongly reject the null hypothesis of no cointegration.

5 Econometric Issues

5a Small sample bias

Asymptotically, Static OLS (SOLS) yields consistent estimates of the coefficients in the cointegrating regression. (SOLS is OLS estimation of a cointegrating relationship.) In the presence of adjustment frictions, however, SOLS will tend to produce biased estimates. Analytical results in Caballero (1994) show that SOLS could be downward biased (i.e., biased towards 0) in time series data by 50 to 60% for a sample of 120 observations and 70 to 80% for a sample of 50 observations, if adjustment frictions are important. There are no analytical results, Monte Carlo simulations, or empirical evidence on the bias of SOLS in panel data in situations where adjustment frictions are important.

To explain the intuition for the SOLS bias, let k^* be the frictionless capital stock (measured in logs and normalized by the log of output) and let it be a linear function of user cost:

$$k_t^* = \alpha_R R_t \quad (3)$$

(For convenience, we ignore the constant term and explain the intuition for a single time series.) Adjustment frictions (broadly defined) will cause a gap u_t between the actual capital stock k_t and the frictionless capital stock. Thus the actual capital stock will be equal to the frictionless capital stock plus u_t:

$$k_t = \alpha_R R_t + u_t \quad (4)$$

In the presence of adjustment frictions, k^* will typically fluctuate more than k, since k will respond only slowly and partially to shocks. Since k is a sum of the random variables k^* and u.

$$\text{var}(k) = \text{var}(k^*) + \text{var}(u) + 2\,\text{cov}(k^*, u) \quad (5)$$

so the variance of k can be smaller than the variance of k^* only if $\text{cov}(k^*,u)$ is negative. However, the OLS estimates of k^* and u (i.e., $\hat{k}^* = \hat{\alpha}_R R$ and $\hat{z} = k - \hat{\alpha}_R R$) are orthogonal by construction, which implies $\text{var}(\hat{k}^*)$ is less than var(k). In order to achieve this, OLS will tend to bias the estimate of α_R toward 0.[9]

5b Dynamic OLS

The necessary condition for *unbiased* SOLS estimation of α_0 and α_R is that u_t be uncorrelated with v_s for all *s* and *t*. This strong condition arises because it is only under this condition that R will be uncorrelated with the error term u since:

$$\begin{aligned} \text{cov}(R_t, u_t) &= \text{cov}(R_0 + \Delta R_1 + \Delta R_2 + \ldots + \Delta R_t, u_t) \quad (6) \\ &= \text{cov}(v_1 + v_2 + \ldots + v_t, u_t) \end{aligned}$$

One solution to the problem of small sample bias in SOLS is the DOLS estimator proposed by Stock and Watson (1993).[10] Dynamic OLS (DOLS) addresses the problem of finite sample bias by replacing the original error term u by a new error term η, which is constructed to be orthogonal to R. The intuition is straightforward. OLS projects the dependent variable onto the space spanned by the right hand side variables. The remaining variation in the dependent variable is orthogonal to the right hand side variables. Suppose u were projected onto the space spanned by all leads and lags of ΔR

[9] This argument follows Caballero (1994, 1999).
[10] Kao and Chiang (2000) provide the panel cointegration DOLS counterpart to the original Stock and Watson (1993) DOLS estimator.

(which is equivalent to the space spanned by v). The error term v_t from this regression will be orthogonal to R_s since:

$$\operatorname{cov}(R_s, v_t) = \operatorname{cov}(R_0 + \Delta R_1 + ... + \Delta R_s, v_t) = 0 \quad (7)$$

The last equality follows from the fact that v_t is orthogonal to all leads and lags of ΔR_t by construction.

In practice, it is not possible to include all leads and lags of ΔR_t in the regression. Instead, a finite number p are included, resulting in the following empirical specification:

$$k_t = \alpha_0 + \alpha_R R_t + \sum_{s=-p}^{p} \beta_s \Delta R_{t-s} + \varepsilon_t \quad (8)$$

5c Common Intercept or Fixed Effects?

An issue that is frequently important in panel data econometrics is the specification of the intercept. Many panel data studies allow for different intercepts for each individual unit. When applied to panel data, equation (2), the specification used by Cabellero (1994) and Schaller (2006) in aggregate time series data, assumes that all firms have the same intercept α_0. An alternative possibility is that the true model might be

$$k_{it} = \alpha_i + \alpha_R R_{it} + u_{it} \quad (9)$$

The intercepts are nuisance parameters, but the choice between equation (2), the common intercept specification, and equation (9), the fixed effects specification, could have an effect on the estimate of α_R, the user cost elasticity. The problem is that, if we assume the common intercept specification but the true DGP involves fixed effects, the fixed effects will become part of the regression error. If R is correlated with the fixed effects, the estimate of α_R will be biased.

On the other hand, the dummy variable (DV) estimator might worsen the bias described earlier in this section. To see this, note that the DV estimator partials out the firm-specific intercepts. More precisely, we can think of the DV estimator as partialling out the firm-specific intercepts by multiplying each variable in the regression by $I-\tilde{P}_1$, where $\tilde{P}_1$ is a projection matrix of the following form:

$$\begin{bmatrix} P_1 & 0 & 0 \\ 0 & P_1 & 0 \\ 0 & 0 & P_1 \end{bmatrix}$$

where

$$P_1 = 1(1'1)^{-1}1'$$

and 1 is a vector of T ones, where T is the time dimension of the balanced panel. Multiplying R by $I-\tilde{P}_1$ replaces R_{it} with:

$$R_{it} - \frac{1}{T}\sum_{t=1}^{T} R_{it}$$

which includes all past and future values of R_{it} in the sample. Since the source of the bias described earlier in this section is correlation between R and u at various leads and lags, the DV estimator has the potential to exacerbate the bias discussed earlier.

In stationary panel data econometrics, the potential correlation between right-hand side variables and fixed effects is often a central issue, and there are well-established tests to determine whether such a correlation is present. In particular, a Hausman test is often used. The Hausman test is based on the idea that the random effects estimator is efficient – and unbiased as long as the right-hand side variables are not correlated with the fixed effects. In contrast, the within estimator is unbiased even in the presence of correlation between the right-hand side variables and the fixed effects. The difference in point estimates provides an informal indication of the presence of fixed effects and the Hausman statistic provides a formal test.

Unfortunately, in non-stationary panel econometrics, there does not yet appear to be a counterpart to the Hausman test. To the best of our knowledge, this potentially central issue in non-stationary panel econometrics has not been addressed in the previous literature, either theoretical or applied. We therefore develop a novel approach in this paper, which may be of potential interest to other researchers.

We begin by comparing the point estimates of the DOLS estimators of equation (2) and (9). The DOLS-CI (common intercept estimate of α_R) is -1.107 (for the T=20 sample, with p=1). The DOLS-W (within) estimate is -0.654. This informal comparison of point estimates suggests that fixed effects might result in an important bias if we used the DOLS-CI estimator. Another possible interpretation is that the true model involves a common intercept, but the DOLS-W estimator exacerbates bias by effectively including all leads and lags of R in the regression. (Note that the DOLS-W estimator is numerically identical to the DOLS-DV estimator.)

Since there is no formal test in the literature, we develop two bootstrapping approaches. The key idea of our bootstrapping approaches is to examine the properties of the DOLS-CI and DOLS-W estimators under the two different specifications of the true model – equations (2) and (9). We use a bootstrapping approach, rather than a Monte Carlo simulation of an underlying economic model, for two main reasons. First and most important, standard economic models of capital choice are largely silent on the key question – the extent to which R might be correlated with the α_i. The bootstrap approach allows us to incorporate information contained in the data about the nature of this correlation. Second, in Sections 7, 8, and 9, we will examine three variations on the standard economic model of capital choice, each of which has different implications for the correct specification of the capital choice problem. By using the bootstrap, we can be agnostic about these modeling issues while taking into account their effects on the variables that appear in the cointegrating regression.

The first case considered in the bootstrap is based on the assumption that equation (9) is the true model and that R may be correlated with the α_i. We allow for a general autoregressive specification of u (the error term in the cointegrating regression) and v (the shocks to user cost) and use the residuals from the cointegrating regression estimated by the appropriate estimator (i.e., DOLS-W, since we are assuming that equation (9) is the true model) to estimate these autoregressive processes. We then assume that the shocks to the autoregressive processes for u and v are jointly normally distributed (using the empirical moments to estimate the variances and correlation of the distribution). On this basis, we can simulate R and k and examine the properties of the DOLS-CI and DOLS-W estimators under the assumption that equation (9) is the true model. We use a similar approach to examine the properties of the DOLS-CI and DOLS-W estimators under the assumption that equation (2) is the true model. The description above applies to the fully parametric bootstrap. We also consider a semi-parametric bootstrap. The details of both bootstrap approaches are described in an appendix.

The results are presented in Table 5. Panel A is based on the first bootstrap approach. The first approach is part of the fully parametric bootstrap family, since it assumes both a parametric model and a known distribution of the shocks. The results are striking. If the true model involves fixed effects, the DOLS-CI estimator will be severely biased, but, if the true model involves a common intercept, the DOLS-W estimator will have a negligible bias. Panel B is based on a semi-parametric bootstrap, in which we continue to assume a parametric model but relax the assumptions on the distribution of the shocks to the autoregressive processes for u and v.[11] The results for the semi-parametric bootstrap are qualitatively similar to those for the fully parametric bootstrap. Relaxing the distributional assumption used in the parametric bootstrap does, however, imply an even larger bias for the DOLS-CI estimator.

The conclusion of our bootstrap test is clear: to avoid a potentially serious bias, we should use the DOLS-W estimator rather than the DOLS-CI estimator.

6 Estimates of the User Cost Elasticity

In this section, we focus on a sample of 2356 firms for which we have at least 20 years of continuous data.[12] The first column of Table 6 presents SOLS estimates of user cost elasticity for this panel. The estimated elasticity is -0.476. The second column of Table 6 presents the DOLS estimate of user cost elasticity (for p=1). The DOLS estimate is -0.654. Consistent with our analysis of the SOLS attenuation bias, the DOLS estimate is about 40% larger than the SOLS estimate.

[11] In addition, our semi-parametric bootstrap differs from the parametric bootstrap because, in the latter, instead of assuming starting values equal to zero in constructing the bootstrapped u and v series, we draw a block of starting values from the actual u and v series for each firm (a common procedure in time series bootstraps).

[12] This choice of minimum time dimension reflects a tradeoff. We need a reasonably large value of T in order to use cointegration techniques. But, as we will see in Sections 7, 8, and 9, the estimated elasticity depends on the composition of the sample. As we increase T, the cross-sectional dimension of the sample declines dramatically (e.g., to N=308 for T=40), increasing the possibility that the subsample may not be representative of the universe of firms.

The DOLS estimate provides support for the relatively high estimates of user cost elasticity obtained by Caballero (1994) and Schaller (2006). As noted above, their estimates were based on aggregate data. Our study is the first to estimate the cointegrating relationship between capital stock and user cost using firm-level panel data.

The Caballero (1994) estimate of the user cost elasticity is for equipment capital only; Caballero's preferred estimate of the elasticity is about -0.9. Schaller (2006) presents estimates for both equipment and structures capital. His preferred estimate is about -1.4 for equipment and 0 for structures. The firm-level panel data estimates are consistent with this evidence, since the firm-level capital stock data includes both structures and equipment.

In explaining the intuition for why DOLS tends to yield less biased estimates, we discuss the case where all leads and lags of ΔR_t are included in the empirical specification to illustrate how this guarantees the orthogonality of R_t and ε_t, the error term in the regression. In Table 6, however, we set p=1, so only one lead and lag of ΔR_t are included. Table 7 shows that the elasticity estimate is reasonably robust to other choices of p. If anything, setting p=1 leads to a relatively conservative estimate of user cost elasticity.

7 Financial Market Imperfections and User Cost Elasticity

Figure 1 presents a simplified diagram of the supply and demand of finance for a firm that faces a binding finance constraint.[13] Under asymmetric information, there may

[13] The diagram is adapted from Fazzari, Hubbard, and Petersen's (1988) classic paper on finance constraints. It is strictly applicable only to a one-period model where investment is the same as the capital stock but provides helpful intuition for the more general case.

be a difference between the cost of internal finance (the opportunity cost; i.e., the interest rate at which the firm lends) and the cost of external finance, leading to a step function in the supply of finance with the step at the point where the firm exhausts its internal finance. If the firm's demand for finance intersects the supply of finance along this step, there will be a wedge between the observable market interest rate r and the shadow cost of finance $r+\omega$.

Relatively little is known about ω. In fact, there has been an extensive debate over the evidence on the existence of finance constraints. Fazzari, Hubbard, and Petersen (1988) and a series of subsequent papers used differences across classes of firms in the coefficient on cash flow in a Q investment equation as evidence of finance constraints. This line of research has been criticized by Kaplan and Zingales (1997), Erickson and Whited (2000), and Gomes (2001), among others.

If ω is small, it should make little difference to the estimated user cost elasticity.[14] If shocks to ω are transitory, they will have little effect on the long-run elasticity estimated from the cointegrating regression, even for firms that were constrained in some years. Finally, the effect of finance constraints on the estimated user cost elasticity depends on the covariance between ω and R. If the shocks to ω and R are orthogonal, the estimated elasticity with respect to measured user cost will be biased towards 0 compared to the elasticity with respect to the true user cost (which incorporates ω).

Our approach is to compare user cost elasticity across classes of firms, focusing on classes of firms that are more likely to be finance constrained. Whenever we compare

[14] Direct estimates of ω by Whited (1992), Ng and Schaller (1996), Chirinko and Schaller (2004), and Whited and Wu (2006), however, suggest that it may be substantial for some firms – on the order of several hundred basis points.

classes of firms, we construct balanced panels with 20 years of data for each class of firms. One reason for doing this is because of the potential for small sample bias. By maintaining a consistent time dimension across subsamples, we ensure that we will induce no difference in the estimated elasticity between classes of firms through differential small sample bias.

By definition, firms are finance constrained if they have good investment opportunities but not enough internal finance (or access to external finance) to be able to carry out their investment projects. Our first measure of finance constraints uses the firm's Tobin's Q to measure investment opportunities, averaging Tobin's Q over the 20 years for which data are continuously available. The theory of finance constraints is based on asymmetric information. To the extent that firms are finance constrained, this theory suggests that firms that find it difficult to credibly communicate private information are more likely to face constraints. One way of capturing information asymmetries is the size of the firm, since potential suppliers of finance will pay a higher fixed cost (relative to the size of the firm) to gather information about small firms. We therefore use the firm's size at the end of the 20 years of continuous data as a way of identifying firms for which asymmetric information is likely to be more important. Our first classification of firms that are more likely to be finance constrained is based on firms that are in the top quintile for Tobin's Q (averaged over the 20 years of continuous data) and that are small, where small firms are defined as those that are in the bottom output quintile (at the end of the 20 years of continuous data).

The first row in Table 8 presents user cost elasticity estimates for small firms with good investment opportunities and the complement of this class of firms (i.e., all the firms that are not small firms with good investment opportunities). The results suggest that finance constraints exist and that they have an effect on the estimated user cost elasticity. For small firms with good investment opportunities, the estimated user cost elasticity is about -0.4. For the remaining firms in the sample, the estimated user cost elasticity is about -0.8, roughly twice as large.

We use a second approach to examine the effect of finance constraints. Young firms will typically find it more difficult to credibly communicate private information than mature firms. For this reason, comparisons between the investment behaviour of young and mature firms have been used to test for finance constraints. We therefore focus on a subset of firms for which we have 40 years of continuous data. We define the firms as young in the first 20 years and mature in the last 20 years. This approach has the advantage of holding all other time-invariant characteristics (such as industry) constant between the two classes of firms we are comparing. The results again suggest that finance constraints exist. The estimated user cost elasticity is about -0.2 for young firms and about -0.9 for mature firms. The standard errors for each of these estimates are quite small (less than 0.1).

8 Corporate Governance and User Cost Elasticity

Corporate governance problems can also introduce a wedge between the observed market interest rate and the discount rate used by a firm. To see the intuition for this, consider an empire building manager whose utility function puts some weight on the size of his firm and some weight on the firm’s profit. Chirinko and Schaller (2004) have shown that such a manager will use a lower discount rate in evaluating investment projects. Specifically, such a manager will set the marginal product of capital (in the absence of taxes, which we ignore for simplicity) as follows.

$$F_K = r - \frac{\gamma}{\beta}\frac{\pi}{K} + \delta = r - \phi + \delta$$

where F_K is the marginal product of capital, γ is the weight on size in manager's utility function, β is the weight on profit (π) in manager's utility function, and the other variables have already been defined in equation (1). Since γ, β, and K are positive, corporate governance problems will introduce a "corporate governance discount" ϕ (for $\pi > 0$). The corporate governance discount will be larger, the larger the weight on size in the manager's utility function and the smaller the weight on profit.

To the best of our knowledge, the only direct estimates of ϕ are provided by Chirinko and Schaller (2004), whose estimate is in the range of 300 to 400 basis points. Nothing is known about the persistence of ϕ and its covariance with user cost. If ϕ is sufficiently large, persistent, and less than perfectly correlated with user cost, estimates of user cost elasticity may be biased toward 0, compared to the elasticity with respect to the "user cost" that is actually being used by the manager of the firm.

Based on Jensen (1986), the firms that are most likely to suffer from corporate governance problems are those with high free cash flow and poor investment opportunities. Table 9 presents user cost elasticity estimates for firms with high free cash flow and poor investment opportunities.[15] The firms that have high free cash flow and poor investment opportunities have a substantially lower estimated elasticity (about -0.3) than the remaining firms (about -0.8).

[15]To focus on corporate governance problems, we classify firms as having poor investment opportunities if Tobin's Q is below the median for their industry.

A necessary condition for corporate governance problems is that managers are entrenched. It is sometimes argued that this is more likely to be the case for old firms. The estimated elasticity is about -0.4 for old firms with poor investment opportunities, compared with about -0.8 for the remaining firms.

The model in Chirinko and Schaller (2004) shows that firms require some degree of market power in order to give reign to corporate governance problems (i.e., to have non-zero ϕ). We do not have data on the market power of the firms in our sample, but we can use size as a rough measure of market power. Despite the fact that size is an imperfect measure of market power, we find that large firms with poor investment opportunities have an estimated elasticity of about -0.2, compared to an elasticity of about -0.8 for the remaining firms.

9 Non-convex Adjustment Costs and User Cost Elasticity

Bertola and Caballero (1994) show that there will be a wedge between the market interest rate and the discount rate used by firms when investments in capital stock are irreversible. They derive the following condition for the optimal capital stock in the presence of non-convex adjustment costs (specifically, irreversibility).

$$F_K = r + \frac{1}{2}\Sigma^2 A + \delta = r + \theta + \delta$$

where Σ^2 is a variance (specifically, of the ratio of the state of demand/technology Z to the price of capital goods P, in their notation) and A is a non-negative scalar. Thus, in the presence of non-convex adjustment costs, there is an "irreversibility premium" θ that increases the discount rate used by firms in choosing their desired capital stock.

Little is known about the magnitude of θ, its covariance with r, or its persistence. However, if θ is sufficiently large, not too strongly correlated with r, and persistent, the estimated user cost elasticity may be biased towards 0, compared with the elasticity to the true user cost (including θ).

The literature on irreversible investment suggests several characteristics that make it more likely that firms will encounter a binding irreversibility constraint. Two of the most important are the drift rate of the stochastic process for demand/technology (Z) and the depreciation rate. A low drift rate means that a firm that inadvertently acquires too much capital will find it difficult to grow out of the problem. In contrast, a firm with rapid growth in demand for its product is less likely to encounter a binding irreversibility constraint and far less likely to encounter a persistently binding constraint. The depreciation rate works in a similar way. If a firm's capital stock depreciates rapidly, a shock that leaves it with too much capital will quickly be overcome by depreciation of the capital stock.

We use two characteristics to identify firms that are more likely to encounter persistent binding irreversibility constraints. First, we use the mean growth rate of real sales over the 20 years of continuous data as our measure of the drift rate. We classify firms with real sales growth below the median for the sample as low drift rate firms. Similarly, we divide firms into classes based on the mean depreciation rate over our sample period; firms with a mean depreciation rate below the median for the sample are classified as low depreciation firms.

As shown in Table 10, the results are dramatic. The estimated user cost elasticity for firms with a low drift parameter is close to 0 (-0.029), precisely estimated, and insignificantly different from 0. In contrast, the remaining firms have an elasticity of about -0.9. The results are similar for firms with a low depreciation rate. Their estimated user cost elasticity is also close to 0 (-0.025), precisely estimated, and insignificantly different from 0.

10 Conclusion

The results in this paper add to a growing body of evidence that the user cost elasticity is substantially larger than suggested by much of the earlier literature, which was based on stationary econometrics. Our preferred elasticity estimate is about -0.7. This is much larger than the estimates of Kiyotaki and West (1996) [-0.1], Tevlin and Whaelan (2003) [-0.2], or Chirinko, Fazzari, and Meyer (1999) [-0.3], all of which are based on stationary econometrics.

Our estimate fits nicely with the small number of previous studies based on cointegration econometrics. Caballero's (1994) estimate [-0.9] for equipment investment is based on aggregate data for a large economy. Because the simultaneity problem tends to be more important when aggregate data for a large economy are used to estimate the elasticity, one would expect the elasticity estimate to be larger for a small, open economy – or for an estimate based on firm-level panel data. Our firm-level estimate of -0.7 for total capital (i.e., equipment and structures taken together) lies between the only cointegration-based estimate for structures capital – Schaller (2006) [0] – and the Caballero (1994) and Schaller (2006) estimates for equipment capital. If equipment and structures capital have roughly equal weight over the sample period, our firm-level estimate of -0.7 is somewhat higher than the implied aggregate estimate for a large economy: about -0.4 or -0.5 based on Caballero's (1994) estimate for equipment capital.

Our paper makes a technical contribution by introducing what appears to be the first test for the importance of fixed effects in the context of panel cointegration econometrics. This turns out to be of first-order importance, since the standard panel DOLS estimator (as described and implemented by Kao and Chiang (2000)) overestimates the user cost elasticity by about 40%.

Our paper also provides new evidence on the importance of finance constraints, corporate governance problems, and non-convexities. Our estimates suggest that each of these issues is of considerable importance.

Appendix 1: Data

Capital Stock and Investment

For the first observation for firm f, the capital stock is based on the net plant (NPLANT), the nominal book value of net property, plant, and equipment (CompuStat item 8). To convert this to real terms, we divide by the sector-specific price index for capital goods (p^K). Since book value is not adjusted for changes in the value of capital goods purchased in the past, we adjust the initial capital stock using the sector-specific ratio of nominal replacement cost to historical cost:

$$K_{f,t_0^f} = \frac{NPLANT_{f,t_0^f}}{p^K_{s,t_0^f}} \frac{K\$_{s,t_0^f}}{KHIST_{s,t_0^f}} \qquad \text{(A1)}$$

where K$ is the current cost net stock of private fixed assets by sector, KHIST is historical-cost net stock of private fixed assets by sector, s is a NAICS sector index (for firm f's sector), and t_0^f is the year of the first observation for firm f.

For subsequent observations, a standard perpetual inventory method is used to construct the capital stock,

$$K_{f,t+1} = (1-\delta_{s,t})K_{f,t} + \frac{I_{f,t+1}}{p^K_{s,t+1}} \qquad \text{(A2)}$$

where δ is the depreciation rate and I is capital expenditures in the firm's financial statements (CompuStat item 128). The firm reports the additions in nominal terms, so we divide by p^K to convert to real terms.

In some cases, there is a data gap for a particular firm. In this case, we treat the first new observation for that firm in the same way as we would if it were the initial observation. This avoids any potential sample selection bias that would result from dropping firms with gaps in their data.

We construct sector-specific, time-varying depreciation rates using data from the BEA. Specifically,

$$\delta_{s,t} = \frac{D\$_{s,2000} DQUANT_{s,t}}{K\$_{s,2000} KQUANT_{s,t}} \qquad \text{(A3)}$$

where D$ is current-cost depreciation of private fixed assets by sector (BEA, Table 3.4ES), DQUANT is the chain-type quantity index of depreciation of private fixed assets by sector (BEA, Table 3.5ES), K$ is the current cost net stock of private fixed assets by sector (as defined above), and KQUANT is the chain-type quantity index of the net stock of private fixed assets by sector (BEA, Table 3.2ES).

We construct the sector-specific price index for capital goods using BEA data:

$$p^K_{s,t} = \frac{100(I\$_{s,t} / I\$_{s,2000})}{IQUANT_{s,t}} \qquad \text{(A4)}$$

where I$ is historical-cost investment in nonresidential private fixed assets by sector (BEA, Nonresidential Detailed Estimates: Investment, historical cost) and IQUANT is the chain-type quantity index of investment in private fixed assets by sector (BEA, Table 3.8ES).

After constructing the capital stock, firms with a value of GPLANT less than $1 million are dropped, where GPLANT is gross property, plant, and equipment (CompuStat item 7), and the first observation for each firm is excluded. We then trim the sample, eliminating the 1% most extreme observations in each tail for the following four variables: I/K, Sales/K, Cost/K, and real sales growth.

Cost of Capital

The cost of capital is calculated as follows

$$R_{f,t} = (r_{f,t} + \delta_{s,t}) \left(\frac{1 - z_{s,t} - ITC_{s,t}}{1 - \tau_t} \right) \frac{p^K_{s,t}}{p^Y_{s,t}} \qquad \text{(A5)}$$

where r is the real, risk-adjusted interest rate, z is the sector-specific present value of depreciation allowances, ITC is the sector-specific investment tax credit rate, τ is the corporate tax rate, p^K is the price of capital goods, and p^Y is the price of output. R is expressed as an annual rate, so r and δ are both expressed as annual rates. Where variables are available at a monthly or quarterly frequency, we take the average for the calendar year. The corporate tax rate is the U.S. federal tax rate on corporate income. The present value of depreciation allowances – for non-residential equipment and structures, respectively – were provided by Dale Jorgenson. (The data provided by Dale Jorgenson end in 2001: for 2002-04, we use 2001 values.) To calculate z, we took the weighted sum of Jorgenson's z's for equipment and structures, where the weights are the share of equipment investment and the share of structures investment (for a given year) in nominal gross private non-residential investment in fixed assets from the Bureau of Economic Analysis (from table 1IHI, where equipment investment is referred to as equipment and software). Because the investment tax credit applies only to equipment, u=0 for structures, we multiply the statutory ITC rate for each year by the ratio of equipment investment to the sum of structures and equipment investment for that year. The corporate tax rates were provided directly by the Treasury Department, and investment tax credit rates are drawn from Pechman (1987, p.160-161). The sector-specific price index for output is the "Chain-Type Price Index for Value Added by Industry" from the BEA GDP-by-Industry Accounts, normalized to 1 in 2000.

The Real Risk-Adjusted Market Discount Rate

The real, risk-adjusted market discount rate is defined as follows,

$$r_{f,t} \;=\; ((1+r^{NOM}_{f,t}) \,/\, (1+\pi^e_t)) - 1.0. \qquad \text{(A6)}$$

The equity risk premium is calculated using CAPM. The components of $r_{f,t}$ are defined and constructed as follows,

$r_{f,t}^{NOM}$	=	Nominal, short-term, risk-adjusted cost of capital
	=	$\lambda_s (1-\tau_t)\ r_t^{NOM,DEBT} + (1-\lambda_s)\ r_{s,t}^{NOM,EQUITY}$.
$r_t^{NOM,DEBT}$	=	Nominal corporate bond rate (Moody's Seasoned Baa Corporate Bond Yield)
$r_{s,t}^{NOM,EQUITY}$	=	Nominal, short-term, risk-adjusted cost of equity capital for firms

in sector s.

	=	$r_t^{NOM,F} + \sigma_s$.
$r_t^{NOM,F}$	=	Nominal, one-year, risk-free rate (One-Year Treasury Constant Maturity Rate)
$\pi_{s,t}^{e}$	=	Sector-specific capital goods price inflation rate from t to t+1. Sector-specific data was not yet available for 2005 at the time of data construction, so $\pi_{s,t}^{e}$ for 2003 was also used for 2004.
σ_s	=	Equity risk premium.
τ_t	=	Marginal rate of corporate income taxation.
λ_s	=	Sector-specific leverage ratio calculated as the mean of book debt for the sector divided by the mean of (book debt + book equity) for the sector. In two sectors (Fabricated Metal Product Manufacturing, NAICS industry 332, and Broadcasting and Telecommuniations, NAICS industries 515-517), book equity is negative, so we set λ_s to 1.

Under the CAPM,

$$\sigma_s = \beta_s (\mu^{EQUITY} - \mu^F), \qquad \text{(A7)}$$

where

β_s	=	CAPM β for sector s
μ^{EQUITY}	=	Total return on equities from 1950-2004. The source is the value-weighted CRSP index (including dividends).

μ^F	=	Total return on risk-free Treasury bills from 1950-2004. The source is the FRED database, specifically the series for 1-Year Treasury Constant Maturity Rate.

Appendix 2: Bootstrap

We consider four cases:

1. True model: fixed effects that are potentially correlated with R_{it}
   a. Estimated by DOLS-CI
   b. Estimated by DOLS-W
2. True model: common intercept
   a. Estimated by DOLS-CI
   b. Estimated by DOLS-W

The following notes describe the fully parametric bootstrap. As noted in the paper, we also conduct a semi-parametric bootstrap, which differs from the fully parametric bootstrap in two respects. First, in the semi-parametric bootstrap, we relax the assumption that the distribution of the shocks to the autoregressive processes for u and v follows a joint Gaussian distribution. Instead, we draw with replacement from the empirical distribution of the residuals to the estimated autoregressive processes for u and v. Second, instead of assuming starting values equal to zero in constructing the bootstrapped u and v series, we draw a block of starting values from the actual u and v series for each firm (a common procedure in time series bootstraps).

1. True Model: Fixed Effects

We begin by considering the first two cases, where the true model involves fixed effects; i.e., where the true model is

$$k_{it} = \alpha_i + \alpha_R R_{it} + u_{it} \qquad (1)$$

$$R_{it} = R_{i,t-1} + v_{it} \qquad (2)$$

Note that R_{it} can also be expressed as

$$R_{it} = R_{i0} + \sum_{j=1}^{t} v_{ij} \qquad (3)$$

Define $u_t = (u_{1t}, u_{2t}, ..., u_{Nt})'$ and $v_t = (v_{1t}, v_{2t}, ..., v_{Nt})'$. It is assumed that u_{it} and v_{it} follow a general linear processes specified as

$$u_{it} = \pi_u(L)\varepsilon_{u,it}\,, \qquad (4)$$

$$v_{it} = \pi_v(L)\varepsilon_{v,it} \qquad (5)$$

where L is the usual lag operator and $\pi_u(z) = \sum_{k=0}^{\infty} \pi_{u,k} z^k$, and $\pi_v(z) = \sum_{k=0}^{\infty} \pi_{v,k} z^k$. Therefore we can approximate the errors u_{it} and v_{it} by finite order AR processes, i.e.:

$$u_{it} = \alpha_{u,1}u_{i,t-1} + ... + \alpha_{u,qu}u_{it-qu} + \varepsilon_{u,it}^{qu} \qquad (6)$$

$$v_{it} = \alpha_{v,1}v_{i,t-1} + ... + \alpha_{v,qv}v_{it-qv} + \varepsilon_{v,it}^{qv} \qquad (7)$$

where

$$\varepsilon_{it}^{qu} = \sum_{k=qu+1}^{\infty} \alpha_{u,k}u_{it-k} \qquad (8)$$

$$\varepsilon_{it}^{qv} = \sum_{k=qv+1}^{\infty} \alpha_{v,k}v_{it-k} \qquad (9)$$

To obtain an estimate of the coefficients α_i and α_R, we estimate (1) using DOLS; i.e., we run the regression

$$k_{it} = \alpha_i + \alpha_R R_{it} + \sum_{s=-p}^{p} \beta_s \Delta R_{it-s} + \eta_{it}^{SOLS-FE} \qquad (10)$$

We refer to the resulting estimates as $\hat{\alpha}_i^{DOLS-FE}$ and $\hat{\alpha}_R^{DOLS-FE}$. We define:

$$\hat{u}_{it} = k_{it} - \hat{\alpha}_i^{DOLS-FE} - \hat{\alpha}_R^{DOLS-FE} R_{it} \qquad (11)$$

We estimate the α_u by regressing the $\hat{u}_{it}$ on *qu* lags of $\hat{u}_{it}$.[16] We denote the residual from this expression $\hat{\varepsilon}_{u,it}^{qu}$.

To estimate the α_v, we let $\hat{v}_{it} = \Delta R_{it}$ and run the following regression

$$\hat{v}_{it} = \alpha_{v,0,i} + \alpha_{v,1}\hat{v}_{i,t-1} + ... + \alpha_{v,qv}\hat{v}_{it-qv} + \varepsilon_{v,it}^{qv} \qquad (12)$$

where $\alpha_{v,0,i}$ is a firm-specific intercept in the auto-regressive specification for v_{it}.[17] We denote the residuals from this regression $\hat{\varepsilon}_{v,it}^{qv}$.

We define:

$$\varepsilon_t = (\varepsilon_{u,1t}^{qu}, ..., \varepsilon_{u,Nt}^{qu}; \varepsilon_{v,1t}^{qv}, ..., \varepsilon_{v,Nt}^{qv}) \qquad (13)$$

---

[16] We choose *qu* to minimize the mean (over firms) of the BIC, subject to $\max qu \le T^{1/3}$. The resulting value of qu is XX.

[17] Including $\alpha_{v,0,i}$ means that $\hat{\varepsilon}_{v,it}^{qv}$ will have zero mean for each i. Including a firm-specific intercept is not necessary in the autoregression for $\hat{u}_{it}$, since there is a firm-specific intercept in equation (1). We choose *qv* to minimize the mean (over firms) of the BIC, subject to $\max qv \le T^{1/3}$. The resulting value of qv is XX.

$$\alpha_u = (\alpha_{u,1},...,\alpha_{u,qu}) \qquad (14)$$

$$\alpha_v = (\alpha_{v,0,1},...,\alpha_{v,0,N},\alpha_{v,1},...,\alpha_{v,qv}) \qquad (15)$$

We assume that $\varepsilon_{u,it}^{qu}$ and $\varepsilon_{v,it}^{qv}$ have a joint Gaussian distribution with means 0, variances σ_u and σ_v, and correlation ρ_{uv}. The parameters σ_u, σ_v, and ρ_{uv} are set equal to the variances of $\varepsilon_{u,it}^{qu}$ and $\varepsilon_{v,it}^{qv}$ and their correlation, respectively.

The simulation begins by drawing a particular realization ε_t^S from the distribution of ε_t.[3] We construct the simulated counterparts to u_t and v_t (which we denote u_t^S and v_t^S) using ε_t^S, equations (6) and (7), $\hat{\alpha}_u$, and $\hat{\alpha}_v$. The initial values of u_{it} and v_{it} are set equal to zero. To eliminate any potential sensitivity to the choice of initial values, we discard the first 20 realizations of u_{it} and v_{it}.

We obtain the simulated R_{it} as

$$R_{it}^S = R_{io}^S + \sum_{j=1}^{t} v_{ij}^S \qquad (16)$$

$$R_{i0}^S = \frac{I}{T}\sum_{j=1}^{T} R_{ij} \qquad (17)$$

The simulated k_{it} are obtained from

$$k_{it}^S = \hat{\alpha}_i^{DOLS-FE} + \hat{\alpha}_R^{DOLS-FE} R_{it}^S + u_{it}^S \qquad (18)$$

We follow the procedure described above to generate a sample of *NT* observations of k_{it}^S and R_{it}^S.

a. Common Intercept Estimator

In the first case, we consider the DOLS-CI estimator, which imposes the assumption that $\alpha_i = \alpha_0$ for all i. To assess the properties of the DOLS-CI estimator, we estimate the regression

$$k_{it} = \alpha_0 + \alpha_R R_{it} + \sum_{s=-p}^{p} \beta_s \Delta R_{it-s} + \eta_{it}^{DOLS-CI} \qquad (19)$$

using the simulated data. We denote the resulting estimates of α_0 and α_R as $\hat{\alpha}_0^{S,FE,CI}$ and $\hat{\alpha}_R^{S,FE,CI}$ where S in the superscript indicates that the estimate is based on simulated data, FE indicates that the true model involves fixed effects, and CI indicates that the estimator assumes a common intercept. We iterate 1,000 times. The bias estimate from the simulation is the mean (over the 1,000 iterations) of $\hat{\alpha}_R^{S,FE,CI} - \hat{\alpha}_R^{DOLS-FE}$.

b. Within Estimator

In the second case, we consider the DOLS-W estimator. We begin by applying the within transformation to equation (1) to obtain

$$\tilde{k}_{it} = \alpha_R \tilde{R}_{it} + \tilde{u}_{it} \qquad (20)$$

where the within transformation (denoted with a tilde) is:

$$\tilde{x}_{it} = x_{it} - \frac{1}{T}\sum_{t=1}^{T} x_{it} \qquad (21)$$

Equation (19) differs from the DOLS Within specification because it omits the demeaned leads and lags of first differences of R_{it}. In order to implement the DOLS Within estimator of α_R on the simulated data, we therefore include these in the specification:

$$\tilde{k}_{it} = \alpha_R \tilde{R}_{it} + \sum_{s=-p}^{p} \beta_s \tilde{\Delta} R_{it-s} + \tilde{\eta}_{it}^{DOLS-FE} \qquad (22)$$

where

$$\tilde{\Delta} x_{it} = \Delta x_{it} - \frac{1}{T_p^{\#} - 1}\sum_{t \in T_p} \Delta x_{it} \qquad (23)$$

$$T_p = \{p+2, p+3, ..., T-p\} \qquad (24)$$

$$T_p^{\#} = \#\{p+2, p+3, ..., T-p\} \qquad (25)$$

and #A denotes the number of elements in the set A. Again, we iterate 1,000 times and construct the bias estimate as the mean of $\hat{\alpha}_R^{S,FE,W} - \hat{\alpha}_R^{DOLS-FE}$, where the W superscript indicates that the within estimator is used on the simulated data.

2. True Model: Common Intercept

So far, we have considered the two cases where the true model involves fixed effects. We turn next to the case where the true model involves a common intercept. The procedure is broadly similar, except that the true model is now

$$k_{it} = \alpha_0 + \alpha_R R_{it} + u_{it} \qquad (26)$$

To obtain the values of α_0 and α_R that we use in the simulations, we estimate DOLS on (26); i.e., we estimate the following equation on the actual data.

$$k_{it} = \alpha_0 + \alpha_R R_{it} + \sum_{s=-p}^{p} \beta_s \Delta R_{it-s} + u_{it} \qquad (27)$$

We refer to the resulting estimates as $\hat{\alpha}_0^{DOLS-CI}$ and $\hat{\alpha}_R^{DOLS-CI}$. We then define the residuals in a way comparable to the way they are defined in equation (11), except using $\hat{\alpha}_0^{DOLS-CI}$ and $\hat{\alpha}_R^{DOLS-CI}$. We estimate the α_u and α_v in the same way as in the fixed effects case and define ε_t, α_u, and α_v as in equations (13)-(15). The parameters σ_u, σ_v, and ρ_{uv} are set equal to the variances of $\varepsilon_{u,it}^{qu}$ and $\varepsilon_{v,it}^{qv}$ and their correlation, respectively.

The simulation proceeds in the same way as in the case where the true model involves fixed effects, again drawing a particular realization ε_t^S from the distribution of ε_t, constructing the simulated counterparts to u_t and v_t, and obtaining the simulated R_{it} from equations (16) and (17). The simulated k_{it} are obtained from

$$k_{it}^S = \hat{\alpha}_0^{DOLS-CI} + \hat{\alpha}_R^{DOLS-CI} R_{it}^S + u_{it}^S \qquad (28)$$

We follow the procedure described above to generate a sample of *NT* observations of k_{it}^S and R_{it}^S.

a. Common Intercept Estimator

In the first case where the true model involves a common intercept, we consider the DOLS-CI estimator. To assess the properties of the DOLS-CI estimator, we estimate the regression (19) using the simulated data. We iterate 1,000 times. The bias estimate from the simulation is the mean (over the 1,000 iterations) of $\hat{\alpha}_R^{S,CI,CI} - \hat{\alpha}_R^{DOLS-CI}$, where the first CI in the superscript indicates that the true model has a common intercept and the second CI in the superscript indicates that the estimator assumes a common intercept.

b. Within Estimator

In the second case where the true model involves a common intercept, we consider the DOLS-W estimator. We estimate the specification in equation (22) on the simulated data. In this case, we construct the bias estimate as the mean of $\hat{\alpha}_R^{S,CI,W} - \hat{\alpha}_R^{DOLS-CE}$, where the W in the superscript indicates that the within estimator is used on the simulated data.

References

Bernanke, Ben, and Mark Gertler, “Agency Costs, Net Worth, and Business Fluctuations, *American Economic Review*, March 1989, 79(1), 14-31.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist, “The Financial Accelerator and the Flight to Quality,” *The Review of Economics and Statistics*, February 1996, 78(1), 1-15.

Bertola, Guiseppe, and Ricardo J. Caballero, “Irreversibility and Aggregate Investment,” *Review of Economic Studies*, April 1994, 61(2), 223-46.

Caballero, Ricardo J., “Small Sample Bias and Adjustment Costs,” *The Review of Economics and Statistics*, February 1994, 76(1), 52-58.

Caballero, Ricardo J., “Aggregate Investment,” in John B. Taylor and Michael Woodford (eds.), *Handbook Of Macroeconomics*, vol. 1B, Amsterdam: Elsevier North Holland, 1999, 813-862.

Caballero, Ricardo J., E.M.R.A. Engel, and J.C. Haltiwanger, “Plant-level Adjustment and Aggregate Investment Dynamics,” *Brookings Papers on Economic Activity* 2, 1995, 1-54.

Chang, Yoosoon, “Bootstrap Unit Root Tests in Panels with Cross-Sectional Dependency,” *Journal of Econometrics*, 2004, 120(2), 263-293.

Chirinko, Robert, “Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications,” *Journal of Economic Literature*, December 1993, 31(4), 1875-1911.

Chirinko, Robert, Steven M. Fazzari, and Andrew P. Meyer, “How Responsive is Business Capital Formation to its User Cost? An Exploration with Micro Data,” *Journal of Public Economics*, 1999, 74, 53-80.

Chirinko, Robert S., Steven M. Fazzari, and Andrew P. Meyer, “That Elusive Elasticity: A Long-Panel Approach To Estimating The Price Sensitivity Of Business Capital,” *10th International Conference on Panel Data, Berlin*, July 5-6, 2002 B3-1, International Conferences on Panel Data.

Chirinko, Robert S., and Schaller, Huntley, "A Revealed Preference Approach to Understanding Corporate Governance Problems: Evidence from Canada," *Journal of Financial Economics*, October 2004, 74(1), 181-206.

Clark, Peter K., “Tax Incentives and Equipment Investment,” *Brookings Papers on Economic Activity*, 1993, 1, 317-347.

Cummins, J.G., and Hassett, K.A., "The Effects of Taxation on Investment: New Evidence from Firm Level Panel Data," *National Tax Journal*, 1992, 45, 243-251.

Erickson, Timothy, and Toni M. Whited, "Measurement Error and the Relationship between Investment and q," *Journal of Political Economy*, October 2000, 108(5), 1027-1057.

Fazzari, Steven M., R. Glenn Hubbard, and Bruce C. Petersen, "Investment, Financing Decisions, and Tax Policy," *American Economic Review*, May 1988, 78(2), 200-205.

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

Goolsbee, A., "The Importance of Measurement Error in the Cost of Capital," *National Tax Journal*, 2000, 53, 215-228

Hassett, Kevin A., and R. Glenn Hubbard, "Tax Policy and Business Investment," in Auerbach, A.J., Feldstein, M. (eds.), *Handbook Of Public Economics*, vol. 3, Amsterdam: North Holland, 2002, 1293-1343.

Jensen, Michael C., "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review*, May 1986, 76(2), 323-329.

Jermann, Urban, and Vincenzo Quadrini, "Stock Market Boom and the Productivity Gains of the 1990's," *University of Pennsylvania*, March 2003.

Kao, Chihwa, "Spurious Regression and Residual-based Tests for Cointegration in Panel Data," *Journal of Econometrics*, May 1999, 90(1), 1-44.

Kao, Chihwa, and M.H. Chiang, "On the Estimation and Inference of a Cointegrated Regression in Panel Data," *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, 2000, 15, 179-222.

Kaplan, Steven, and Luigi Zingales, "Do Investment-Cashflow Sensitivities Provide Useful Measures of Financing Constraints?" *Quarterly Journal of Economics*, February 1997, 112, 169-215.

Kiyotaki, N., and Kenneth D. West, "Business Fixed Investment and the Recent Business Cycle in Japan," *NBER Macroeconomics Annual*, 1996, 277-323.

Levin, A., C.F. Lin, and C.J. Chu, "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics*, 2002, 108, 1-24.

Ng, Serena, and Huntley Schaller, "The Risky Spread, Investment, and Monetary Policy Transmission: Evidence on the Role of Asymmetric Information," *The Review of Economics and Statistics*, August 1996, 78(3), 375-83.

Schaller, Huntley, "Estimating the long-run user cost elasticity," *Journal of Monetary Economics*, May 2006, 53(4), 725-736.

Stock, James H., and Mark W. Watson, "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, July 1993, 61(4), 783-820.

Tevlin, Stacey, and Karl Whelan, "Explaining the Investment Boom of the 1990s," *Journal of Money, Credit and Banking*, February 2003, 35(1), 1-22.

Whited, Toni, "Debt, Liquidity Constraints and Corporate Investment: Evidence from Panel Data," *Journal of Finance*, 1992, 47, 1425-1460.

Table 1
Summary Statistics – Full Sample

	Mean	Median	Standard Deviation	Skewness	Kurtosis
kr	1138.041	57.872	5290.592	14.509	321.373
yr	1398.936	159.867	6344.198	15.197	340.104
kryr	1.208	0.401	6.537	333.133	128443.312
$\tilde{R}$	0.074	0.067	0.049	2.450	18.604
ct	0.421	0.421	0.109	0.733	2.197
cp	1.159	1.068	0.609	3.348	38.389
cr	0.163	0.159	0.071	0.531	0.806
yrg	0.114	0.063	0.314	3.006	17.815
krg	0.100	0.044	0.196	3.336	16.106

The variable kr is the replacement value of the capital stock, measured in millions of 1996 dollars, yr is output measured in millions of 1996 dollars, kryr is the ratio kr/yr, $\tilde{R}$ is user cost, ct is the tax component of user cost (the second term in parentheses in equation (1)), cp is the price of capital goods component (the third term in parentheses in equation (1)), cr is the interest rate component (the first term in parentheses in equation (1)), yrg is the growth rate of yr, and krg is the growth rate of kr.

Table 2
Levin-Lin-Chu Test Results for Unit Root

Variable	Parameter	t*	P>t
k	-0.0834	95.607	1.000
R	-0.6213	47.587	1.000

Table 3
Chang Test Results for Unit Root

Variable	Test	P>t
k	-1.201	1.000
R	0.736	1.000

Table 4
Tests for Cointegration

DF Test	P>t	PP Test	P>t	ADF Test	P>t
-35.263	0.000	-3.345	0.000	21.088	0.000

DF Test is the Kao Dickey-Fuller test, PP Test is the Kao Phillips-Perron test, and ADF Test is the Kao Augmented Dickey-Fuller test.

Table 5
Fixed Effect or Common Intercept?

	Panel A: Fully Parametric Bootstrap	
	True Model	
Estimator	Fixed Effect	Common Intercept
DOLS-CI	-0.342	-0.001
DOLS-W	-0.003	-0.004

	Panel B: Semi-Parametric Bootstrap	
	True Model	
Estimator	Fixed Effect	Common Intercept
DOLS-CI	-0.369	-0.007
DOLS-W	-0.005	-0.028

Each cell presents the bias of the estimator, conditional on the given true model, based on the bootstrap procedure described in Appendix 2.

Table 6
SOLS and DOLS Estimates of User Cost Elasticity

SOLS Estimate	DOLS Estimate
-0.476 (0.005)	-0.654 (0.025)

The standard error is in parentheses under the elasticity estimate. The estimates are for T=20, N=2356. The DOLS estimate is for p=1.

Table 7
DOLS Estimates of User Cost Elasticity for Different Values of p
Subsample with 20 years of continuous data

p	DOLS Estimate
1	-0.654 (0.025)
2	-0.708 (0.025)
3	-0.758 (0.025)
4	-0.774 (0.025)

The first column reports p, the number of leads and lags of first differences of the right-hand-side variable (user cost) used in DOLS estimation. The main entries in the cells of the second column are the DOLS estimate of user cost elasticity. The standard error is in parentheses under the elasticity estimate.

Table 8
Finance Constraints and User Cost Elasticity

Class	Elasticity Estimate			
	Class		Class Complement	
	N	DOLS-W	N	DOLS-W
Small and Good Investment Opportunities	47	-0.392 (0.174)	2165	-0.773 (0.025)
	Young		Mature	
	N	DOLS-W	N	DOLS-W
Young	308	-0.234 (0.042)	308	-0.858 (0.069)

The standard error is in parentheses under the elasticity estimate. The DOLS estimate is for p=1.

Table 9
Corporate Governance and User Cost Elasticity

Class	Elasticity Estimate			
	Class		Class Complement	
	N	DOLS-W	N	DOLS-W
High Free Cash Flow and Poor Investment Opportunities	174	-0.329 (0.106)	2028	-0.783 (0.026)
Old and Poor Investment Opportunities	137	-0.413 (0.112)	2075	0.769 (0.025)
Large and Poor Investment Opportunities	195	-0.233 (0.088)	2017	-0.786 (0.025)

The standard error is in parentheses under the elasticity estimate. The DOLS estimate is for p=1.

Table 10
Non-convex Adjustment Costs and User Cost Elasticity

Class	Elasticity Estimate			
	Class		Class Complement	
	N	DOLS-W	N	DOLS-W
Low Drift Parameter	1052	-0.029 (0.041)	1304	-0.862 (0.024)
Low Depreciation Rate	1047	-0.025 (0.036)	1309	0.899 (0.025)

The standard error is in parentheses under the elasticity estimate. The DOLS estimate is for p=1.

Figure 1
Finance Constraints and the Lemons Premium