Managerial Entrenchment and Firm Value: A Dynamic Perspective*

Xin Chang

Cambridge Judge Business School
The University of Cambridge

Division of Banking & Finance
Nanyang Business School
Nanyang Technological University

Hong Feng Zhang

School of Accounting Economics and Finance
Faculty of Business and Law
Deakin University, Australia

Keywords: Corporate Governance; Managerial Entrenchment; Long Difference Estimator; Reverse Causality; Panel VAR.

JEL classification: **G30; G34; L25**

* We are grateful to Paul Malatesta (the editor), an anonymous referee, Rob Brown, Greg Schwann, Qi Zeng, and workshop participants at the University of Melbourne and the 21st Australasian Banking and Finance Conference for helpful comments and suggestions. We thank Rongbing Huang for providing us his codes for the long-differencing technique. We are grateful to Yujun Lian for sharing with us the codes for estimating the PVAR model with exogenous variables. We thank Carmen Shih and Jiaquan Yao for excellent research assistance. All remaining errors are ours. Chang acknowledges financial support from *Academic Research Fund Tier 1 provided by* Ministry of Education (Singapore). Zhang acknowledges the financial supports for his PhD research from the University of Melbourne.

Managerial Entrenchment and Firm Value: A Dynamic Perspective

Abstract

We examine the impact of managerial entrenchment on firm value using a dynamic model with firm fixed effects. To estimate the model, we employ the long difference technique, which is shown by our simulation to deliver the least biased estimates. Based on a large sample of U.S. companies, we document a significantly negative and causal effect of managerial entrenchment on firm value after taking into account omitted variables, reverse causality, and highly persistent endogenous variables. Additional analysis suggests that the causality running from managerial entrenchment to firm value is more pronounced than reverse causality.

Keywords: Corporate Governance; Managerial Entrenchment; Long Difference Estimator; Reverse Causality; Panel VAR.

JEL classification: **G30; G34; L25**

I. Introduction

Endogeneity plagues empirical research on the relation between managerial entrenchment and firm value, often resulting in biased parameter estimates or disagreements on the direction of causality.[1] Based on a managerial entrenchment index (*E* index hereafter) constructed using six anti-takeover provisions, Bebchuk, Cohen, and Ferrell (2009) (BCF hereafter) document a negative effect of managerial entrenchment on firm value, implying that entrenched managers who experience less pressure from corporate governance mechanisms may adopt value-destroying corporate policies. In contrast, Lehn, Patro, and Zhao (2007) (LPZ hereafter) find that the negative relation between firm value and managerial entrenchment disappears after controlling for historical firm values, indicating that the entrenchment index is associated with the current firm value mainly through past firm valuation. They conclude that managers of firms with low historical valuation may adopt more anti-takeover provisions to further entrench themselves, rather than that the adoption of more anti-takeover provisions entrenches management and thus reduces firm value.

In this paper, we use Monte Carlo simulation to show that the conflicting results in previous studies are largely due to the inadequacy of the models used to address endogeneity issues, which include omitted variables and reverse causality that runs from *past* firm value to the *current* level of managerial entrenchment. Furthermore, we employ more appropriate econometric techniques to tackle endogeneity issues and identify the causal effect of managerial entrenchment on firm value.

---

[1] We follow Berger, Ofek, and Yermack (1997) and define entrenchment as the extent to which managers do not experience discipline from "*the full range of corporate governance and control mechanisms, including monitoring by the board, the threat of dismissal or takeover*". A higher degree of managerial entrenchment implies weaker corporate governance.

Specifically, our empirical analysis focuses on the estimation of a dynamic model with firm fixed effects. The model accounts for time-invariant omitted variables using fixed effects and controls for reverse causality using lagged firm values. The estimation of dynamic panel models with firm fixed effects is sensitive to the econometric procedure used. As discussed in Section II.A, the conventional mean-differencing estimates are biased in the presence of reverse causality. The system Generalized Method of Moments (GMM) approach suffers from finite sample biases, especially when endogenous variables are highly persistent (Alonso-Borrego and Arellano 1999). This limitation is particularly relevant to our analysis since firm value is highly autocorrelated and the measure of managerial entrenchment is nearly time-invariant. Instead, we employ Hahn, Hausman, and Kuersteiner's (2007) long difference technique, which involves taking multi-year rather than one-year differences, relies on a small set of moment conditions, and can be used to enhance the explanatory power of the instruments and reduce the finite sample bias even if the endogenous variables are highly persistent.

We generate simulated data from a dynamic model that consists of firm fixed effects, reverse causality, and serially correlated endogenous variables. We find that the long difference estimator indeed offers the least biased estimate, and that the advantage of the long difference estimator over other estimators increases with the autocorrelations of endogenous variables. In addition, our simulation demonstrates that adding distantly lagged values of endogenous variables to static models, as in LPZ and BCF, inadequately accounts for reverse causality and insufficiently draws causal inferences.

Using a large panel of the U.S. firms from 1990 to 2007, we empirically estimate the impact of managerial entrenchment on firm value with the long difference estimator. We measure firm value using the industry-adjusted Tobin's *Q*, and use BCF's *E* index to proxy for managerial entrenchment. The results show that the changes in the *E* index have a significant and negative effect on the changes in firm value after controlling for the influence of past changes in firm value on the changes in the *E* index, indicating that managerial entrenchment causally reduces firm value. In terms of economic significance, a one-standard deviation (1.3) increase in the *E* index is associated with an annual decrease in the industry-adjusted *Q* by 0.014, which amounts to 4.4% of the mean value of the industry-adjusted *Q*.

The causal relation between managerial entrenchment and firm value can be bidirectional, and the two directions of causality are not necessarily mutually exclusive. To evaluate the relative importance of forward causality (entrenchment affecting firm value) and reverse causality (firm value affecting entrenchment), we use a panel-data vector autoregression (PVAR) specification, which controls for autocorrelations and time trends and allows for firm-specific unobserved heterogeneity. In addition, we perform the bivariate Granger causality test for each firm in our sample and summarize the results using the Fisher method proposed by Maddala and Wu (1999). The results show that forward causality is statistically more significant and economically stronger than reverse causality.

Our paper contributes to the extant literature in two ways. First, our simulation analysis reveals that the empirical models of both LPZ and BCF can give rise to biased coefficients, and thus neither is adequate to address the inherent endogeneity problems and make causal inferences. Second, utilizing dynamic panel models that account for the estimation and inference problems in the existing literature, we document a significantly negative and causal effect of management entrenchment on firm value.

The rest of the paper is organized as follows. Section II outlines our dynamic panel model and compares different estimators. Section III describes the sample and variables. Regression results and causality tests are presented in Section IV. Section V concludes.

II. Empirical Methodology

In this section we first present our dynamic panel model and discuss the advantages of the long difference estimator over the mean-differencing and system GMM estimators in mitigating the biases. We then use simulations to run a horse race among different estimators.

A. Estimating Dynamic Models with Fixed Effects

In a seminal paper, Gompers, Ishii, and Metrick (2003) use the following static model to examine the impact of managerial entrenchment (or corporate governance) on firm value.

(1) $Q_{it} = \alpha + \beta E_{it} + \delta C_{it} + \varepsilon_{it}$, $\qquad$ (i = 1, …, N; t = 1, …, T),

where Q_{it} is the value of firm i at the end of year t, E_{it} is the measure of managerial entrenchment, C is a vector of firm-specific control variables whose effects on firm value are represented by δ, ε is the error term, α is the constant term, and β captures the impact of managerial entrenchment on firm value.

There are at least two sources of endogeneity commonly known to bias the β estimates: (1) omitted variables, such as CEO ability or corporate culture, and (2) reverse causality, which arises if past firm valuation affects the current level of managerial entrenchment and if firm valuation is highly persistent.[2] To mitigate the concern of omitted variables, the common practice is to add

[2] Both LPZ and BCF suggest that low-Q firms are more likely to become takeover targets, thus their managers may adopt more anti-takeover provisions to insulate themselves from corporate control markets. On the flip side, managers of high-Q firms may adopt fewer entrenchment provisions since the likelihood of being a target is low.

firm fixed effects to control for the effects of time-invariant unobserved firm heterogeneity.[3] To deal with reverse causality, it is important to control for the lagged dependent variable (Q_{t-1}). Thus, we focus on estimating the following dynamic model with firm fixed effects:

(2) $$Q_{it} = \alpha + \rho Q_{it-1} + \beta E_{it} + \delta C_{it} + f_i + \varepsilon_{it},$$

where ρ captures the serial correlation of Q and f_i is the firm fixed effects. In the model we allow E_{it} to be determined by its own lagged values, the lagged value of Q, and a set of control variables (D) that affect the extent of managerial entrenchment. That is,

(3) $$E_{it} = \gamma E_{it-1} + \lambda Q_{it-1} + \phi D_{it} + \mu_{it},$$

where γ measures the autocorrelation of the entrenchment measure, λ captures the extent of reverse causality, ϕ represents the effects of control variables, and μ is the error term.[4]

Both LPZ and BCF are aware of the problems associated with the static model and attempt to address endogeneity using dynamic specifications, which involve adding to equation (1) deeply lagged values of E and/or Q. In Section II.B we evaluate the validity of their approaches using simulation, alongside with other approaches outlined below.

How to estimate β in equation (2) in the presence of reverse causality specified in equation (3)? It is well known that the standard mean-differencing technique results in biased estimates of ρ (e.g., Nickell 1981, Huang and Ritter 2009), however, the bias in β estimates is less clear. We derive in the Appendix the bias in the mean-differencing estimate of β. Both the sign and

[3] Chi (2006) estimates equation (1) with firm fixed effects and documents that the negative impact of managerial entrenchment on firm value is statistically and economically significant.

[4] λ is referred to as dynamic endogeneity by Wintoki, Linck, and Netter (2012) who examine the causal effect of board structure on firm performance. They define dynamic endogeneity as the impact of past firm performance on current board structure.

magnitude of the bias are found to be determined by the autocorrelation of Q and the extent of reverse causality, that is, ρ and λ, respectively.

The system Generalized Method of Moments (GMM) approach developed by Arellano and Bover (1995) and Blundell and Bond (1998) has been increasingly used in recent studies to estimate dynamic models with fixed effects.[5] Blundell and Bond (1998) show that the system GMM estimator outperforms the mean-differencing estimator. In particular, the system GMM estimator takes the first difference of equation (2):

$$(4)\quad Q_{it}-Q_{it-1}=\rho(Q_{it-1}-Q_{it-2})+\beta(E_{it}-E_{it-1})+\delta(C_{it}-C_{it-1})+(\varepsilon_{it}-\varepsilon_{it-1}).$$

Equations (2) and (4) are then simultaneously estimated as a "system" using the lagged differences (Q_{it-2} - Q_{it-3}, . . . , Q_{i1} - Q_{i0}, and E_{it-1} - E_{it-2}, . . . , E_{i1} - E_{i0}) as instruments for equation (2) and the lagged levels (Q_{it-2}, . . . , Q_{i0}, and E_{it-1}, . . . , E_{i0} as instruments for equation (4). Wintoki, Linck, and Netter (2012) argue that, under reasonable assumptions, the system GMM procedure offers efficient estimates in the presence of omitted variables and reverse causality.[6] However, Alonso-Borrego and Arellano (1999) show that the system GMM estimator suffers from finite sample biases, especially when the dependent variable is highly persistent, as is the case with firm valuation (Q) in our analysis. In addition, since GMM exploits all the linear moment restrictions

---

5 Among others, Antoniou, Guney, and Paudyal (2008) and Lemmon, Roberts, and Zender (2008) investigate the dynamics of leverage ratio using the system-GMM procedure. Wintoki, Linck, and Netter (2012) use a dynamic GMM panel estimator to investigate the relation between board structure and firm performance.

6 Wintoki, Linck, and Netter (2012) propose that, for the system GMM estimator to produce efficient estimates, one needs to impose two orthogonality conditions. The first condition requires the explanatory variables beyond period t-p to be uncorrelated with the error term (ε_{it}), where p is the number of lags of the dependent variable included in the dynamic model. The second condition requires the correlation between explanatory variables and unobserved heterogeneity (omitted variables) to be constant over time.

specified by the model, the set of moment conditions for the system GMM estimator may explode as the time dimension increases. This can be a severe problem for finite samples containing no sufficient information for estimating a large instrument matrix.[7]

Hahn, Hausman, and Kuersteiner (2007) and Huang and Ritter (2009) show that the long difference estimator, which relies on a small set of moment conditions, is less biased than the mean-differencing and system GMM estimators when the dependent variable is highly persistent. The long difference approach involves taking a multi-year difference of equation (2). Specifically, firm valuation (Q) at the end of year $t-k$ is written as

$$(5)\quad Q_{it-k} = \alpha + \rho Q_{it-k-1} + \beta E_{it-k} + \delta C_{it-k} + f_i + \varepsilon_{it-k}.$$

Subtracting equation (5) from equation (2) yields

$$(6)\quad Q_{it} - Q_{it-k} = \rho(Q_{it-1} - Q_{it-k-1}) + \beta(E_{it} - E_{it-k}) + \delta(C_{it} - C_{it-k}) + (\varepsilon_{it} - \varepsilon_{it-k}),$$

or

$$(7)\quad \Delta Q_{i[t,t-k]} = \rho \Delta Q_{i[t-1,t-k-1]} + \beta \Delta E_{i[t,t-k]} + \delta \Delta C_{i[t,t-k]} + \eta_{i[t,t-k]}.$$

To estimate equation (7), we follow Hahn, Hausman, and Kuersteiner (2007) and Huang and Ritter (2009) by taking an iterated two-stage least squares (2SLS) approach.[8] To obtain the initial

---

[7] Hsiao (2003) shows that theoretically, using more moment conditions can improve the efficiency of system GMM estimator, however, the efficiency gain in a finite sample is very limited. Ziliak (1997) also demonstrates that the downward bias in GMM is becoming severe with the increasing use of more moment conditions. Wintoki, Linck, and Netter (2012) also point out that while the use of deeply lagged variables as instruments in the system GMM may increase their exogeneity, such use worsens the potential weak-instrument problem.

[8] Note that both Hahn, Hausman, and Kuersteiner (2007) and Huang and Ritter (2009) focus on the coefficient of the lagged dependent variable, ρ, while we are interested in β, the coefficient of the endogenous variable.

coefficient estimates of $\hat{\rho}$, $\hat{\beta}$, and $\hat{\delta}$, we use Q_{it-k-1} and E_{it-k} as valid instruments for $\Delta Q_{i[t-1,t-k-1]}$ and $\Delta E_{i[t,t-k]}$, respectively, and estimate equation (7) with 2SLS. Each iteration starts with computing the residuals, $Q_{it-1} - \hat{\rho} Q_{it-2} - \hat{\beta} E_{it-1} - \hat{\delta} C_{it-1}$,..., and $Q_{it-k} - \hat{\rho} Q_{it-k-1} - \hat{\beta} E_{it-k} - \hat{\delta} C_{it-k}$, which are shown by Hahn, Hausman, and Kuersteiner (2007) to be valid instruments. Each iteration ends with updating coefficient estimates ($\hat{\rho}$, $\hat{\beta}$, and $\hat{\delta}$) using the residuals as well as Q_{it-k-1} and E_{it-k} as instruments to estimate equation (7) with 2SLS. We iterate the process three times because Hahn, Hausman, and Kuersteiner (2007) suggest that three iterations are often sufficient.

The long difference estimator has another important advantage over other methods; it can better identify the causal relation when the key *explanatory* variable is highly persistent. In the managerial entrenchment/firm valuation context, if the *E* index changes very little from year to year, any mean-differencing or first differencing approaches force the identification of the causal relation between *Q* and *E* from only the very few firm-years that experience nonzero changes in the *E* index. In contrast, the long difference procedure takes a multi-year difference which is more likely than one-year differences to be nonzero, resulting in a higher signal-to-noise ratio and less biased coefficient estimates (Hsiao (2003)).[9]

B. Monte Carlo Simulations

In this subsection we use Monte Carlo simulations to illustrate the biases in the mean-

[9] Hahn, Hausman, and Kuersteiner (2007) argue that the long difference approach is essentially the traditional 2SLS approach albeit the iteration process. The optimal numbers of moment conditions and instrumental variables are minimized to make it comparable with the traditional 2SLS.

differencing and system GMM estimators, and the unbiasedness of the long difference estimator when estimating dynamic models with firm fixed effects and persistent dependent variables. Unlike most of the simulation exercises in previous studies (e.g., Hahn, Hausman, and Kuersteiner 2007, Phillips and Sul 2007, Huang and Ritter 2009), which primarily investigate the bias in the autoregressive parameter ($\hat{\rho}$) in equation (2), our simulations focus on the bias in $\hat{\beta}$.[10]

Specifically, we use the data generating process given by equations (2) and (3) to generate a panel of 1,500 hypothetical firms for 15 years, similar in terms of the average time and cross-sectional dimensions, to the actual data described in Section III.A. The initial values of Q and E are drawn from normal distributions that have the same means and standard deviations as those in the actual data (reported in Table 1), namely Q_{i0} ~ NORMAL(0.32, 1.32) and E_{i0} ~ NORMAL(2.5, 1.69). Error terms and firm fixed effects are drawn from normal distributions; ε_{it} ~ NORMAL(0, 0.1), μ_{it} ~ NORMAL(0, 0.1), and f_i ~ NORMAL(0, 0.1).[11] For simplicity, we assume in simulations that there are no other explanatory variables (C and D) affecting Q and E in equations (2) and (3) and $\alpha = 0$.

We examine various values of β, but for the sake of brevity, we report the case of $\beta = 0.1$, which is representative. We vary the value of λ, which measures the extent of reverse causality, from -1 to 1 with a step of 0.1 so that we can examine how the biases in $\hat{\beta}$ change in response to

[10] Wintoki, Linck, and Netter (2012) also focus on the bias in the coefficient of endogenous variable and use simulation to show that the system GMM estimator outperforms the OLS and fixed effects estimators. However, the key difference between their simulation and ours is that they assume $\rho = 0$, while we do not impose any assumptions on ρ in the data generating process.

[11] Following Huang and Ritter (2009) and Wintoki, Linck, and Netter (2012), we set the variance of error terms equal to 0.1. Similar results (available upon request) are obtained if we use alternative values, such as 0.2. 0.3, or 0.5.

the change in the magnitude of reverse causality. The bias in $\hat{\beta}$ is defined as ($\hat{\beta}$ - β). For each value of λ, we repeat the simulation 500 times and estimate equation (2) using different estimators. The average biases in $\hat{\beta}$ across 500 simulations are plotted against λ in Graphs A-C of Figure 1 for the mean-differencing, system GMM and long difference estimators (differencing length = 8), respectively.[12]

Furthermore, to examine whether the OLS models, augmented by including distantly lagged values of *E* and *Q*, are adequate to address reverse causality, we estimate the following model using the simulated data.

(8) $$Q_{it} = \alpha + \rho \overline{Q}_{i[1,5]} + \beta E_{it} + \phi E_{i5} + \varepsilon_{it}. \qquad (i = 1, \ldots, N;\ t = 6, \ldots, T),$$

LPZ augment the OLS model (equation (1)) by including the historical average value of *Q* measured over the period 1980-1985. BCF regress *Q* in the 1998-2002 period on the value of *E* in 1990 while controlling for *Q* as of 1990, the beginning of their sample period.[13]

We mimic their specifications by including the average simulated value of *Q* in the first five years ($\overline{Q}_{i[1,5]}$) and the value of simulated *E* in year 5, and estimate equation (8) using observations from year 6 onwards. The average bias in $\hat{\beta}$ across 500 simulations is plotted against λ in Graph D of Figure 1.

---

[12] For the long difference estimator, we follow Huang and Ritter (2009) and set differencing length = 8 in simulation. Unreported results show that similar simulation results are obtained for differencing length = 6 or 10.

[13] The intuition behind this specification is that while the value of *E* in 1990 is correlated with *E* during the period 1998-2002 because managerial entrenchment is highly persistent over time, the extent of managerial entrenchment (*E*) in 1990 should not be determined by low firm valuation (*Q*) during the 1998-2002 period. *Q* as of 1990 is added to control for the correlation between *E* and *Q* in 1990.

To illustrate the impact of autocorrelations of Q and E on the bias in $\hat{\beta}$ for all four estimators, we report the simulation results for two scenarios with high ($\rho = \gamma = 0.9$) and low ($\rho = \gamma = 0.1$) autoregressive parameters, respectively.

[Inset Figure 1 here]

Graph A of Figure 1 reveals that both autocorrelation and reverse causality contribute to the bias of the mean-differencing estimator, consistent with the discussion in Section II.A. In particular, when autocorrelations of Q and X are low, the bias of the mean-differencing estimator is negatively associated with the extent of reverse causality. As autocorrelations increase, the absolute value of the bias in $\hat{\beta}$ increases and the relation between the bias in $\hat{\beta}$ and the extent of reverse causality becomes non-monotonic. Graph B shows that the system GMM estimator yields small biases when the autoregressive parameter is low. However, consistent with Alonso-Borrego and Arellano (1999) and Hahn, Hausman, and Kuersteiner (2007), we find that the system GMM estimator produces significantly biased coefficients when the dependent variable is highly persistent ($\rho = 0.9$).

In contrast to the mean-differencing and GMM estimators, Graph C of Figure 1 demonstrates that the long difference estimator produces the least biased estimates of β regardless of the magnitude of reverse causality and the degree of autocorrelation, confirming the advantages of the long difference approach as discussed in Section II.A. Finally, Graph D shows that the model augmentations employed by LPZ and BCF are inadequate in addressing endogeneity issues. Adding distant lags of Q and E to the static model results in severely biased estimates of β, especially when autoregressive parameters are high.[14]

[14] Note that the vertical scale of Graph D ranges from -0.5 to 1, while those of Graphs A-C range from -0.1 to 0.1.

Taken together, our simulations illustrate the relative unbiasedness of the long difference estimator when the data generating process consists of reverse causality, unobserved heterogeneity, and highly persistent endogenous variables.

III. Data, Variables, and Summary Statistics

A. Data and Variables

Our empirical analysis focuses on all the U.S. public companies covered in at least one of the volumes published by the Investor Responsibility Research Center (IRRC) between 1990 and 2007. The IRRC volumes contain information on anti-takeover provisions for firms in the Standard & Poor's 1500 and other major U.S. companies, which collectively represent over 90% of the total market capitalization of the combined NYSE, AMEX, and NASDAQ markets (Gompers, Ishii, and Metrick 2003). The IRRC has published eight volumes for the years 1990, 1993, 1995, 1998, 2000, 2002, 2004, and 2006.[15] Following Gompers, Ishii, and Metrick (2003) and BCF, we fill in for the missing years by assuming that the anti-takeover provisions reported in any given year are also in place in years prior to the next volume's publication.[16] Because our regression analysis

---

Therefore, the biases in Graph D are of a larger scale than those in Graphs A-C.

[15] The 2008 volume has been published but is not included in our analysis because there are significant changes in the method of data collection for anti-takeover provisions after 2007 This makes the volumes before and after 2006 incomparable. Wharton Research Data Services (WRDS) offers a detailed comparison between different data formats.

[16] For instance, in the case of 2001, for which there was no IRRC volume in the subsequent year, we assume that the anti-takeover provisions were the same as those reported in the IRRC volume published in 2000. We obtain very

involves a number of control variables, we match the IRRC data with data from Compustat and CRSP for each firm year. To examine the dynamic relation between managerial entrenchment and firm value, we further require firms to have at least four IRRC survey data available in our sample. We also exclude firms with dual-class shares because their governance structures are very different from those of single-class firms. This leaves an unbalanced panel consisting of 1,070 firms and 13,735 firm-year observations.

Gompers, Ishii, and Metrick (2003) employ 24 anti-takeover provisions to construct their governance index (*G* index). BCF point out that some of the 24 provisions might matter more than others and that some may be redundant. Accordingly, they identify six provisions as effective entrenchment devices adopted by a firm, namely staggered boards, bylaw and charter amendment limitations, supermajority requirements for mergers and charter amendments, poison pills, and golden parachute arrangements. They construct the entrenchment index (*E* index) using those six provisions and show that the index is significantly and negatively associated with firm valuation, while the other eighteen provisions not in the *E* index are uncorrelated with firm value.[17]

We thus use BCF's *E* index as our measure of managerial entrenchment. The *E* index is constructed as follows. For every firm-year, we add one point for every provision that increases

---

similar results (unreported, but available upon request) if we assume that firms have the same anti-takeover provisions as in the next publication year.

[17] Robustness checks (untabulated but available upon request) reveal that our main results hold if we use the *G* index in regressions. The *G* index generally has coefficients that are statistically consistent with those obtained using the *E* index but exhibits weaker economic effects on firm value. The weaker effect of the *G* index on firm value is consistent with BCF’s argument that six anti-takeover provisions included in the *E* index, as a subset of the 24 provisions included in the *G* index, drive the economic effect of the *G* index on firm value.

managerial entrenchment (weakens shareholder rights). The *E* index is then the sum of one point for the existence (or absence) of each provision. The value of the *E* index for a particular firm-year ranges from 0 to 6. An increase in the *E* index implies a higher level of takeover defenses, a lower level of shareholder rights protection, and a higher level of managerial entrenchment.

Following BCF, we use Tobin's *Q* as the proxy for firm value. Tobin's *Q* is defined as the market value of assets divided by the book value of assets. The market value of assets is equal to the market value of equity plus the book value of assets minus the book value of common equity and the deferred taxes. The industry-adjusted Tobin's *Q*, $\hat{Q}$, as the dependent variable in our regressions, is computed by subtracting the median industry *Q* from a firm's *Q*. Industry is defined using the Fama-French (1997) classification of forty-eight industry groups. Similar results (untabulated) are obtained by using unadjusted Tobin's *Q* or adjusting *Q* using the two-digit SIC code. The autocorrelation of $\hat{Q}$ in our sample is high but decays rapidly, suggesting that it is sufficient to control for the persistence of firm valuation using one-year lag of firm value in equation (2).[18]

We use in our regression analysis the standard control variables that have been shown by previous studies to be associated with firm value. Following BCF, we include the "other provisions index" (*O* index) obtained by allotting one point to each of the eighteen anti-takeover provisions that are not included in the *E* index. We include the log of the book value of assets, *Ln*(*Assets*), as a proxy for firm size, the log of the firm's age, *Ln*(*Age*), which is defined as the

[18] The untabulated results show that the averages of the first-order autocorrelation coefficients and partial autocorrelation coefficients of the industry-adjusted *Q* in our sample are 0.32 and 0.42, respectively. The second and third-order partial autocorrelation coefficients are -0.14 and -0.03, respectively.

number of years since the firm entered CRSP, a dummy variable for Delaware incorporation (*Delaware*), earnings before extraordinary items scaled by the book value of total assets (*ROA*), capital expenditures scaled by total book value of assets (*Capex/Assets*), leverage ratio (*Leverage*) which is defined as total debt (the sum of short-term and long-term debt) divided by the book value of total assets, and research and development expenses scaled by net sales (*R&D/Sales*). We also include an *R&D* indicator variable (*R&D Dummy*) that equals one if *R&D* expenses are missing, and zero otherwise. To mitigate the impact of outliers or misrecorded data, all explanatory variables are winsorized at the 1% level at both tails of the distribution. All dollar values are converted into year 2000 constant dollars using the GDP deflator.

B. Summary Statistics

Table 1 reports the summary statistics for the overall sample. The firms included in the sample are generally larger and older than a typical firm in Compustat. The mean (median) book value of total assets is $5,333.0 ($1,529.7) million. The mean (median) number of years included in CRSP is 31.1 (27.0). 50% of the firms in our sample are incorporated in Delaware. The last column reports the Pearson correlation coefficients between the *E* index and other variables. All the correlation coefficients are statistically significant at the 1% level, except for *R&D Dummy*. Both the unadjusted Tobin's *Q* and industry-adjusted Tobin's *Q* are statistically and negatively correlated with the *E* index.

[Insert Table 1 here]

We then partition firm-years into seven groups according to the level of the *E* index, which ranges from 0 to 6. Panel A of Table 2 reports the incidence of the index levels and the mean values of variables for each group. There are only 4.1% of the observations with the *E* index greater than 4. Panel B presents mean values of firm characteristics for each group. Firms in the low *E* index groups are larger in size and have higher values of unadjusted and industry-adjusted Tobin's *Q*.

[Insert Table 2 here]

The *E* index is highly persistent, confirming the need for the long difference technique to estimate dynamic panel data models with highly persistent dependent and explanatory variables. By examining the distribution of the *change* in the *E* index between IRRC publication dates (1990, 1993, 1995, 1998, 2000, 2002, 2004, and 2006), we find that 77.8% of the firm-years have no change in the *E* index between the publication dates (untabulated). The mean and median absolute changes in the *E* index between the publication dates are 0.1 and 0, respectively.

IV. Empirical Results

In this section we first present the empirical results obtained using different estimation approaches outlined in Section II, and then explore the direction of causality between managerial entrenchment and firm value.

A. The Impact of Managerial Entrenchment on Firm Value

We first briefly replicate previous studies (e.g., Gompers, Ishii, and Metrick 2003, LPZ, BCF) using our sample. In column (1) of Table 3, we apply the Fama-MacBeth (1973) procedure to the static model (equation (1)) with no firm fixed effects. The estimated coefficient of the *E* index (-0.059) is similar in magnitude to those reported by LPZ and BCF.

[Insert Table 3 here]

To illustrate the possibility of reverse causality, we regress the contemporaneous *E* on the lagged value of $\hat{Q}$ and other control variables. The coefficient of $\hat{Q}_{t-1}$ is negative (-0.130) and significant at the 1% level (t-statistic = -9.85), indicating that the current value of E is strongly associated with the lagged value of Q.

We also replicate, with our sample, the tests of LPZ and BCF by adding to equation (1) the average $\hat{Q}$ during 1980-1985 ($\hat{Q}_{[1980,1985]}$) and the E index in 1990 (E_{1990}). The results in column (3) of Table 3 show that the coefficients of both E_{1990} and the concurrent value of E are negative and statistically significant at the 5% level even after controlling for the historical average value of Q. This finding appears to indicate that the result of LPZ do not hold up outside their sample period (1990-2002). When we use a larger sample over the period 1990-2007, the result of LPZ disappears, while that of BCF remains. Nevertheless, we take care in interpreting the results in columns (1) and (3) as evidence of a causal effect of the E index on firm value, because the simulation exercise in Section II.B illustrates that using static models and/or adding historical values of E and Q to the static model do not eliminate estimation biases, especially when E and Q are highly persistent.[19]

We now move to our main specification, the dynamic model with firm fixed effects (equation (2)), which controls for both omitted variables and allows for reverse causality specified in equation (3). Column (4) of Table 3 reports results obtained using the mean-differencing estimator. The estimated coefficient of the E index is positive and statistically insignificant. As discussed in Section II.A, this estimate, however, is biased due to the autocorrelation of Q and reverse causality.[20] Column (5) reports the results produced by the system GMM estimator. We

[19] BCF also suggest that one should interpret their results with caution due to the limitations of static models in revealing the direction of causality.

[20] In addition, Gompers, Ishii, and Metrick (2003) argue that the coefficient is statistically insignificant possibly

employ the two-step system GMM estimator that is robust to heteroskedasticity and serial correlations in the residuals.[21] The coefficient of the *E* index is negative (-0.090) and significant at the 5% level (*z*-statistic = -2.33). The coefficient of the lagged $\hat{Q}$ is statistically significant and positive (0.475), consistent with the high autocorrelation of $\hat{Q}$ discussed in Section III.A. A highly persistent dependent variable, however, exacerbates the finite sample bias of the GMM estimator.

To circumvent the problems associated with the mean-differencing and system GMM estimators, we employ the long difference approach outlined in Section II.A.

[Insert Table 4 here]

We view both $\hat{Q}$ and the *E* index as endogeneous variables and use the iterated two-stage least squares (2SLS) approach. Table 4 reports the results obtained using differencing lengths of k = 6, 8, and 10 years. The changes in the *E* index have a consistently negative impact on the changes in firm value after controlling for the influence of past changes in firm value on the changes in the *E* index. The statistical significance of the estimated coefficients increases with differencing length. The coefficients are -0.104 for k = 6 (*t*-statistic = -1.23), -0.086 for k = 8 (*t*-statistic = -2.38), and -0.113 for k = 10 (*t*-statistic = -2.95). Economically, a one-standard-

because the conventional fixed effect estimators force the identification of the governance index coefficient only from observations with nonzero changes, which are scarce by nature.

[21] To mitigate the finite sample bias, we employ the Windmeijer (2005) correction for standard errors. The resulting *z*-statistics are reported in parentheses. The two-step GMM estimator uses one-step GMM residuals to construct asymptotically optimal weighting matrices, and is thus more efficient than the one-step estimator if the error term is subject to heteroskedasticity in the large panel data with a relative long time span. See Blundell and Bond (1998) for further discussions on the one-step and two-step GMM estimators.

deviation (1.3) increase in E index is associated with a decrease in $\hat{Q}$ by 0.11 (= -0.086×1.3) over a period of eight years (for k = 8), or 0.014 annually. Given that the mean value of $\hat{Q}$ is 0.32, these magnitudes represent economically meaningful entrenchment effects. The coefficients of control variables are generally consistent with those reported in Table 3.

The estimated coefficients are sensitive to the differencing length partly because the firm-year observations decrease with differencing length from 6,797 in Column (1) to 3,166 in Column (3). In addition, to the extent that it takes time for the causal relation between managerial entrenchment and firm value to take effect, the change in managerial entrenchment over a longer time period should have statistically stronger influences on the change in firm value.

Taken together, based on the long difference approach that is able to uncover the dynamic relationship while controlling for possible omitted variable bias, reverse causality, and autocorrelations, we complement Gompers, Ishii, and Metrick (2003) and BCF by revealing a negative and causal effect of managerial entrenchment on firm value. The robust causal effect of managerial entrenchment on firm value is consistent with a recent paper by Bebchuk, Cohen, and Wang (2011) who find that although the G index and the E index no longer generate abnormal returns in the 2000s due to investors gradually learning to appreciate the difference between well-governed and poorly governed companies, their relevance for firm value and operating performance persists.

B. Causality between Firm Value and Managerial Entrenchment

In this subsection, we examine the casual relation between firm value and managerial entrenchment, i.e., whether managerial entrenchment (as proxied by the E index) lowers firm value, or managers of low value firms adopt more anti-takeover provisions to entrench themselves, or both. To this end, we first employ a panel-data vector autoregression (PVAR) methodology, proposed by Holtz-Eakin, Newey, and Rosen (1988).[22] This approach combines the conventional vector autoregression technique, which allows a vector of variables to be endogenously determined in the system, with the panel-data approach, which controls for unobserved heterogeneity. Specifically, our two-equation reduced-form PVAR model can be written as follows.

$$(9) \quad \hat{Q}_{it} = a_{0t} + \sum_{k=1}^{m} a_k \hat{Q}_{it-k} + \sum_{k=1}^{m} b_k E_{it-k} + \delta C_{it} + f_i + x_t + \varepsilon_{it},$$

$$(10) \quad E_{it} = c_{0t} + \sum_{k=1}^{m} c_k \hat{Q}_{it-k} + \sum_{k=1}^{m} d_k E_{it-k} + \phi C_{it} + g_i + y_t + \omega_{it},$$

where m is the number of time lags that is sufficiently large to ensure that ε_{it} and ω_{it} are white noise error terms, C is a vector of exogenous control variables, x_t and y_t are year fixed effects, and f_i and g_i are unobserved firm fixed effects for the industry-adjusted Tobin's Q and the E index, respectively.

To address the issue that firm fixed effects are correlated with the regressors due to lags of the dependent variables, we follow Love and Zicchino (2006) and take the forward mean-differencing approach (the Helmert procedure), which removes the fixed effects by transforming all variables in the model to deviations from forward means, i.e., the mean values of all future observations for each firm in a given year. This transformation preserves homoscedasticity and the orthogonality

[22] The PVAR approach has become increasingly popular as a tool for disentangling the causality effect and investigating intertemporal interactions between variables. For example, Grinstein and Michaely (2005) apply the methodology to investigating the causality effect between institutional holdings and payout policy. Love and Zicchino (2006) use the PVAR technique to examine the dynamic relationship between firms' financial conditions and investment. Goto, Watanabe, and Xu (2009) employ the technique to study the interaction between strategic disclosure and stock returns.

between transformed variables and lagged regressors (Arellano and Bover 1995), enabling us to use the lagged values of regressors as instruments to estimate the coefficients with the GMM approach. Year fixed effects are removed by subtracting the mean value of each variable computed for each year.

In the empirical estimation, we set the number of lags equal to one (m = 1) on the basis of three commonly used order-selection criteria, namely the Akaike information criterion (AIC), the Schwarz Bayesian information criterion (SBIC), and the Hannan and Quinn information criterion (HQIC).[23] Because we need one-period-lagged values of E and $\hat{Q}$ as regressors and two-period-lagged values as instruments, the number of firms is reduced to 679 (11,438 firm-years) for this test.

We estimate the coefficients of the system given in equations (9) and (10) and report the results in Table 5. Column (1) suggests a strong negative causal effect of the E index on the industry-adjusted Tobin's Q. The effect is statistically significant at the 1% level (z-statistic = -2.59). Column (2) presents the effect of firm value on the E index. The coefficient of $\hat{Q}_{t-1}$ is negative and statistically significant at the 10% level (z-statistic = -1.82), suggesting that past firm value causes the current extent of managerial entrenchment.

Based on the coefficients reported in Table 5, we construct the impulse response functions, which trace the impact of a one-unit positive shock (or innovation) to one endogenous variable on the current and future values of other endogenous variables in the system, assuming that the shock reverts to zero in subsequent periods and that all other shocks are equal to zero. Since the shocks

[23] Specifically, when m = 1, all three information criteria are minimized (AIC = 2.17, BIC = 2.98, and HQIC = 4.04). Untabulated robustness checks suggest that setting m = 2 or m = 3 does not qualitatively change the results that follow.

to different endogenous variables are allowed to be correlated, we use the inverse of the Cholesky decomposition of the residual covariance matrix to orthogonalize the impulses.[24] The standard errors and confidence intervals of the impulse-response functions are calculated using Monte Carlo simulations with 500 repetitions to gauge the statistical significance of the responses.

Figure 2 presents two charts of the impulse-response functions (bold lines) and the 5% error bands (dotted lines), which correspond respectively to the 5^{th} and 95^{th} percentiles of the 500 bootstraps. Graph A shows that a one-unit increase in the current value of the *E* index results in a reduction in firm value by approximately 2% for a period of six years. The response of firm value to the *E* index is statistically significant at better than the 5% level because the 95% error band is below the zero line.

Graph B of Figure 2 presents the response of managerial entrenchment to a one-unit increase in current firm value. It shows that the extent of managerial entrenchment responds negatively to a shock in firm value, consistent with our conjecture that low valued firms adopt more anti-takeover provisions to entrench themselves. However, compared with the response of firm value to the *E* index, the response of the *E* index to firm value is smaller in magnitude and less statistically significant (i.e., a part of the 95% error band is above the zero line).

Finally, we perform the Granger causality test at firm level to examine the direction of causality between managerial entrenchment and firm value. Using the 679 firms in the PVAR

[24] Results from the impulse-response functions are generally sensitive to the specific ordering of the endogenous variables. More specifically, placing a variable earlier in the ordering tends to increase its impact on the variables that follow it. While we do not have a strong prior preference for the relative ordering of firm value and the *E* index in the PVAR system, untabulated robustness checks reveal that our impulse response results are robust to the ordering of the two variables.

estimation, we conduct the following bivariate Granger causality test for each individual firm and summarize the results in Table 6.[25]

(11) $\hat{Q}_t = a_0 + a_1\hat{Q}_{t-1} + a_2E_{t-1} + \delta C_t + \varepsilon_t$

(12) $E_i = b_0 + b_1\hat{Q}_{t-1} + b_2E_{t-1} + \phi C_t + \omega_t$

Column (1) of Table 6 describes the result of the null hypothesis that the *E* index does not Granger-cause $\hat{Q}$. We find that the null hypotheses can be rejected at the 10% level of significance for 52.4% of the firms (356 firms). 22.7% of the firms exhibit Granger causality running from *E* to $\hat{Q}$ at the 1% significance level. We then swap the *E* index and $\hat{Q}$ to test the null hypothesis that $\hat{Q}$ does not Granger-cause the *E* index. We find that the hypothesis can be rejected at the 10% (1%) level of significance for 42.2% (15.3%) of the firms. We use the Fisher method of Maddala and Wu (1999) to combine *p*-values from independent tests of significance for each firm. The test statistics, which follow a χ^2 distribution with 2×679 degrees of freedom, indicate that both hypotheses, namely "the *E* index does not Granger-cause $\hat{Q}$" and "$\hat{Q}$ does not Granger-cause the *E* index, can be rejected at the 1% level. However, the χ^2 statistic of the former hypothesis (χ^2 = 5270.3) is larger than that of the latter (χ^2 = 4307.7), suggesting that the causality running from *E* to $\hat{Q}$ is statistically stronger than that running from $\hat{Q}$ to *E*.

Overall, both the PVAR analysis using panel data and the Granger causality tests at the firm level indicate that the causality between managerial entrenchment and firm run in both directions, confirming the necessity and importance of controlling for reverse causality when investigating

[25] Based on the analysis using the Panel VAR model, we set the lag length (*m*) equal to one in equations (11) and (12). Untabulated results show that similar inferences can be drawn if we increase the lag length to two or three.

the causal effect of managerial entrenchment on firm value. In addition, the tests suggest that the causality running from the *E* index to Tobin's *Q* is statistically more significant and economically stronger than the reverse causation running from *Q* to the *E* index, consistent with the findings of Cremers and Ferrell (2011) who hand-collect data of shareholder rights over the period 1978-1989 and use almost thirty years (1978-2006) of data to investigate the relationship between shareholder rights and firm value. They also find a robust and significant negative effect of the *E* index on firm value, but limited evidence for reverse causation.

VI. Conclusion

Self-interested managers of low value firms may fear the loss of their private benefits of control and have incentives to entrench themselves against pressures from internal and external corporate governance mechanisms. On the other hand, entrenched managers who are not pressured by corporate governance may adopt corporate policies that reduce firm value. In this study, we empirically evaluate the causal relation between managerial entrenchment and firm value.

We use BCF's entrenchment index to measure managerial entrenchment and employ the long difference technique, which is shown by our simulation to be appropriate for panel data with a highly persistent endogenous variables, to address two specific threats to causal inference, namely omitted variables and reverse causality. Our results reveal that the changes in the *E* index have a significant and negative impact on the changes in firm value after controlling for the influence of past changes in firm value on the changes in the *E* index, implying that managerial entrenchment causally reduces firm value.

We also evaluate the relative importance of forward causality (entrenchment affecting firm value) and reverse causality (firm value affecting entrenchment) using the PVAR model and the Granger causality test. Apart from identifying the bidirectional causality between managerial entrenchment and firm value, the results suggest that forward causality is statistically more significant and economically stronger than reverse causality.

References

Alonso-Borrego, C., and M. Arellano. "Symmetrically Normalized Instrumental-variable Estimation Using Panel Data." *Journal of Business and Economic Statistics*, 17 (1999), 36–49.

Antoniou, A.; Y. Guney; and K. Paudyal. "The Determinants of Capital Structure: Capital Market-oriented versus Bank-oriented Institutions." *Journal of Financial and Quantitative Analysis*, 43 (2008), 59–92.

Arellano, M., and O. Bover. "Another Look at the Instrumental Variable Estimation of Error-components Models." *Journal of Econometrics*, 68 (1995), 29-51.

Bebchuk, L.; A. Cohen; and A. Ferrell. "What Matters in Corporate Governance?" *Review of Financial Studies*, 22 (2009), 783-827.

Bebchuk, L.; A. Cohen; and C. Wang. "Learning and the Disappearing Association between Governance and Returns." *Journal of Financial Economics*, forthcoming (2011).

Berger, P. G.; E. Ofek; and D. L. Yermack. "Managerial Entrenchment and Capital Structure Decisions." *Journal of Finance*, 52 (1997), 1411-1438.

Blundell, R., and S. Bond. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics*, 87 (1998), 115-43.

Chi, J. D. "Understanding the Endogeneity between Firm Value and Shareholder Rights." *Financial Management*, 34(2006), 65-76.

Cremers, M., and A. Ferrell. "Thirty Years of Shareholder Rights and Firm Valuation." Working paper, Yale School of Management (2011).

Fama, E., and K. French. "Industry Costs of Equity." *Journal of Financial Economics*, 43(1997), 153-193.

Fama, E., and J. MacBeth. "Risk, Return and Equilibrium: Empirical Tests." *Journal of Political Economy*, 81 (1973), 607-636.

Gompers, P.; J. Ishii; and A. Metrick. "Corporate Governance and Equity Prices." *Quarterly Journal of Economics*, 118 (2003), 107-155.

Goto, S.; M. Watanabe; and Y. Xu. "Strategic Disclosure and Stock Returns: Theory and Evidence from US Cross-Listing." *Review of Financial Studies*, 22 (2009), 1585-1620.

Grinstein, Y., and R. Michaely. "Institutional Holdings and Payout Policy." *Journal of Finance*, 60 (2005), 1389-1425.

Hahn, J.; J. Hausman; and G. Kuersteiner. "Long Difference Instrumental Variables Estimation for

Dynamic Panel Models with Fixed Effects." *Journal of Econometrics*, 140 (2007), 574–617.

Holtz-Eakin, D.; W. Newey; and H. S. Rosen. "Estimating Vector Autoregressions with Panel Data." *Econometrica*, 56 (1988), 1371–1395.

Hsiao, C. *Analysis of Panel Data*, 2nd ed. Cambridge: Cambridge University Press (2003).

Huang, R., and J. R. Ritter. "Testing Theories of Capital Structure and Estimating the Speed of Adjustment." *Journal of Financial and Quantitative Analysis*, 44 (2009), 237-271.

Lehn, K.; S. Patro; and M. Zhao. "Governance Indices and Valuation Multiples: Which Causes Which?" *Journal of Corporate Finance*, 13 (2007), 907-928.

Lemmon, M. L.; M. R. Roberts; and J. F. Zender. "Back to the Beginning: Persistence and the Cross-section of Corporate Capital Structure." *Journal of Finance*, 63 (2008), 1575–1608.

Love, I., and L. Zicchino. "Financial Development and Dynamic Investment Behavior: Evidence from Panel VAR." *The Quarterly Review of Economics and Finance*, 46 (2006), 190–210.

Maddala, G.S., and S. Wu. "A Comparative Study of Unit Root Tests With Panel Data and A New Simple Test." *Oxford Bulletin of Economics and Statistics*, 61 (1999), 631-652.

Nickell, S. “Biases in Dynamic Models with Fixed Effects.” *Econometrica*, 49 (1981), 1417–1426.

Phillips, P.C.B., and D. Sul. “Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence.” *Journal of Econometrics*, 137(2007), 162-188.

Windmeijer, F. “A Finite Sample Correction for the Variance of Linear Efficient Two-step GMM Estimators.” *Journal of Econometrics*, 126 (2005), 25-51.

Wintoki, M. B.; J. S. Linck; and J. M. Netter. “Endogeneity and the Dynamics of Internal Corporate Governance.” *Journal of Financial Economics*,105 (2012), 581-606.

Ziliak, J.P. “Efficient Estimation with Panel Data When Instruments Are Predetermined: An Empirical Comparison of Moment-Condition Estimators.” *Journal of Business and Economic Statistics*, 15(1997), 419–431.

Appendix: The Bias in the Mean-differencing Estimator

In this appendix, we derive the bias in the mean-differencing estimator. To estimate equation (2), a dynamic panel model with firm fixed effects, the standard procedure is to start by removing the fixed effects f_i, which can be achieved by subtracting the time mean of (2) from (2) itself to yield

(A1) $Q_{it} - Q_{i\cdot} = \rho(Q_{it-1} - Q_{i\cdot\,-1}) + \beta(E_{it} - E_{i\cdot}) + \delta(C_{it} - C_{i\cdot}) + (\varepsilon_{it} - \varepsilon_{i\cdot}),$

where for any variable z_{it}, $z_{i\cdot} = (1/T)\sum_{t=1}^{T} z_{it}$ and $z_{i\cdot\,\text{-}1} = (1/T)\sum_{t=0}^{T-1} z_{it}$.

By first stacking cross section data and then time series observations, equation (A1) can be written as

$$\tilde{Q} = \rho\tilde{Q}_{-1} + \beta\tilde{E} + \tilde{C}\delta + \varepsilon,$$

where Q, E, and C are vectors, and the tilde affix signifies that the variables are demeaned. Using the standard technique, we can obtain from equation (A1) the OLS estimate, $\hat{\beta}$, which is a biased estimate of β. The bias in $\hat{\beta}$ can be written as follows.

(A2) $\hat{\beta} - \beta = A(\tilde{Q}_{-1}^{'}\tilde{\varepsilon}) + B(\tilde{E}^{'}\tilde{\varepsilon}),$

where A and B are scalars. Specifically,

$$A = \frac{-(\tilde{E}^{'}\tilde{E})^{-1}(\tilde{E}^{'}\tilde{Q}_{-1})}{\tilde{Q}_{-1}^{'}M\tilde{Q}_{-1}}, B = \left[\frac{(\tilde{E}^{'}\tilde{E})^{-1}(\tilde{E}^{'}\tilde{Q}_{-1})(\tilde{Q}_{-1}^{'}\tilde{E}(\tilde{E}^{'}\tilde{E})^{-1})}{\tilde{Q}_{-1}^{'}M\tilde{Q}_{-1}} + (\tilde{E}^{'}\tilde{E})^{-1}\right], \text{and } M = I - \tilde{E}(\tilde{E}^{'}\tilde{E})^{-1}\tilde{E}^{'}.$$

The derivation extends the analysis of Nickell (1981) and Phillips and Sul (2007) who investigate the biases analytically under the assumption that E is strictly exogenous with respect to the dependent variable. Our equation (A2), however, is derived based on the assumption that there is reverse causality, i.e., $\lambda \neq 0$ in equation (3). A detailed derivation of equation (A2) is available upon request from the authors.

Equation (A2) suggests that the bias in $\hat{\beta}$ comes from two sources. The first source is driven by the correlation between the transformed lagged dependent variable ($\tilde{Q}_{-1}$) and the transformed error term ($\tilde{\varepsilon}$), which has been shown by Nickell (1981) to be dependent on ρ and be negative if $\rho > 0$. The second source of bias depends on the correlation between the transformed explanatory variable ($\tilde{E}$) and $\tilde{\varepsilon}$, which is driven by the extent of reverse causality (λ). Wintoki, Linck, and Netter (2012) show that the correlation between $\tilde{E}$ and $\tilde{\varepsilon}$ is nonzero if $\lambda \neq 0$ in equation (3). More specifically, $\mathrm{E}(\tilde{X}'\tilde{\varepsilon}) = -\lambda\sigma_{\varepsilon}^{2}[(T-1) - T\pi + \pi^{T}] / T^{2}(1-\pi)^{2}$, where $\mathrm{E}(\cdot)$ is the expectation operator, σ_{ε}^{2} is the variance of the error term in equation (3), and $\pi = \beta\,\lambda$.

Taken together, the direction of the bias in $\hat{\beta}$ is ambiguous since it is determined by the signs and magnitudes of autocorrelation and reverse causality, that is, ρ and λ, respectively.

Figure 1: Measuring Biases for Different Estimators

We use the data generating process given by equations (2) and (3) to generate a panel of 1,500 hypothetical firms for 15 years. The initial values of Q and X, error terms, and firm fixed effects are drawn from normal distributions: Q_{i0} ~ NORMAL(0.32, 1.32), E_{i0} ~ NORMAL(2.5, 1.69), ε_{it} ~ NORMAL(0, 0.1), μ_{it} ~ NORMAL(0, 0.1), and f_i ~ NORMAL(0, 0.1). We set in simulations $\beta = 0.1$. The value of λ varies from -1 to 1 with a step of 0.1. The y-axis of all charts measures the bias in $\hat{\beta}$, which is defined as ($\hat{\beta}$ - β). For each value of λ, we repeat the simulation 500 times and estimate equation (2) using different estimators. The average biases in $\hat{\beta}$ across 500 simulations are plotted against λ in Graphs A-C for the mean-differencing, system GMM and long difference estimators, respectively. Graph D reports the results obtained by estimating equation (8) which involves adding distantly lags of Q and E to the static model. For all four estimators we report the simulation results for two scenarios with high ($\rho = \gamma = 0.9$) and low ($\rho = \gamma = 0.1$) autoregressive parameters, respectively.

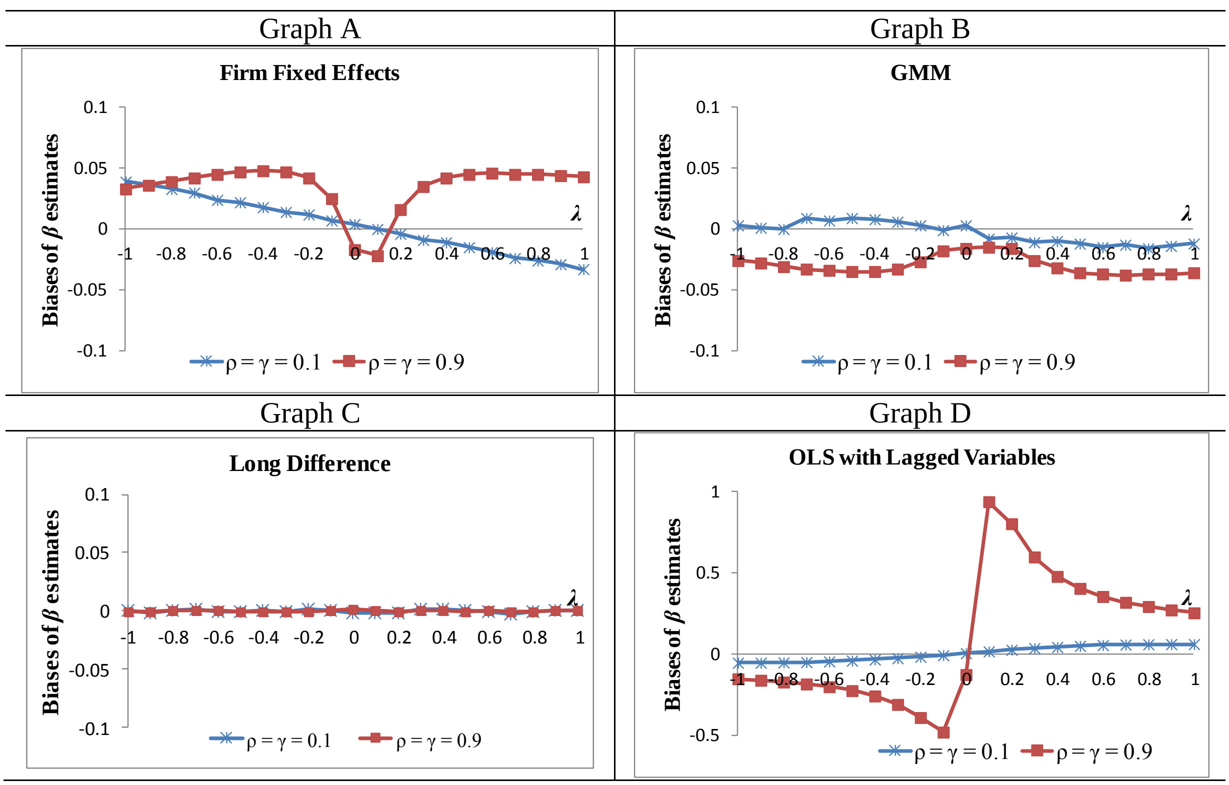
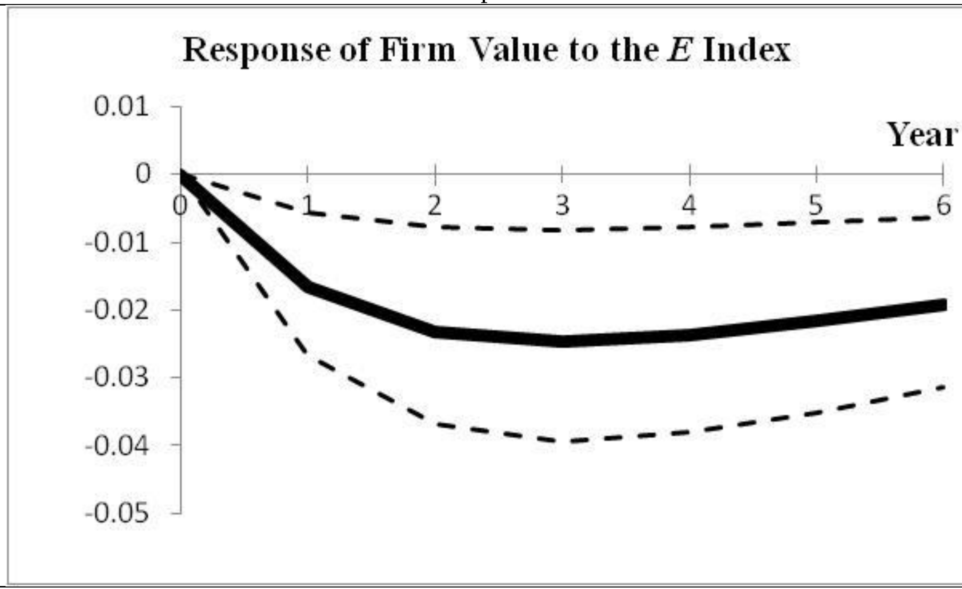


Figure 2: Impulse Responses between Firm Value and Managerial Entrenchment

This figure presents the impulse-response functions and the 5% error bands generated by Monte Carlo simulations with 500 repetitions. Firms are required to have data for at least eight consecutive years between 1990 and 2007. The impulse functions are constructed based on the coefficients estimated using the panel vector autoregression (PVAR) model reported in Table 5. Firm value is measured by the industry-adjusted Tobin's *Q* obtained by subtracting the median industry *Q* from a firm's *Q*. Tobin's *Q* is defined as the market value of assets divided by the book value of assets. Industry is defined using the Fama-French (1997) classification of 48 industry groups. The *E* index is the entrenchment index of Bebchuk, Cohen, and Ferrell (2009). Graph A shows the responses of firm value for six years to a one-unit increase in the current value of the *E* index. Graph B presents the responses of managerial entrenchment for six year to a one-unit increase in the current firm value. Bold lines show the impulse responses. The dotted lines represent the 5% error bands, which correspond respectively to the 5^{th} and 95^{th} percentiles of the 500 bootstraps.

Graph A



Graph B

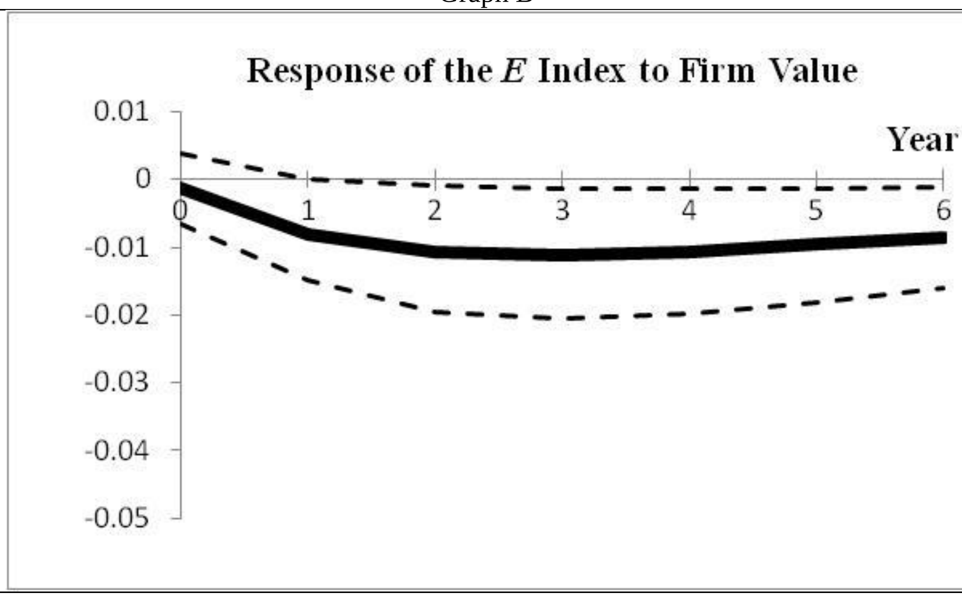


Table 1: Descriptive Statistics

Firms are required to have at least four IRRC survey data between 1990 and 2007. Financial data are retrieved from Compustat and CRSP. The *E* index is the entrenchment index of Bebchuk, Cohen, and Ferrell (2009). The *O* index is constructed using all the other 18 anti-takeover provisions that are tracked by the IRRC but not included in the *E* index. Tobin's *Q* is defined as the market value of assets divided by the book value of assets. Industry-adjusted Tobin's *Q* is obtained by subtracting the median industry *Q* from a firm's *Q*. Industry is defined using the Fama-French (1997) classification of 48 industry groups. *Assets* is the total book value of assets. *Age* is defined as the number of years since the firm entered CRSP. *Leverage* is defined as total debt (the sum of short-term and long-term debt) divided by total assets. *ROA* is the operating income before depreciation and amortization over total assets. *R&D/Sales* is research and development expenses scaled by net sales. *R&D Dummy* is an *R&D* indicator variable that equals one if *R&D* expenses are missing, zero otherwise. *CAPEX/Assets* is capital expenditures scaled by total assets. *Delaware* is a dummy variable that equals 1 if a firm is incorporated in Delaware, and 0 otherwise. All dollar values (e.g., total assets) are converted into year 2000 constant dollars using the GDP deflator. All variables except the governance indices and dummy variables are winsorized at the 1% level at both tails of the distribution. Reported descriptive statistics of each variable include the number of observations (N), mean, median, standard deviation, minimum, maximum, and the Pearson correlation coefficients between the *E* index and other variables. Correlation coefficients that are significant at the 1% level are marked with *** in superscripts.

Variables	N	Mean	Median	Standard Deviation	Minimum	Maximum	Correlation with the *E* index
E index	13,735	2.5	3.0	1.3	0.0	6.0	
O index	13,735	7.0	7.0	2.0	2.0	13.0	0.32^{***}
Tobin's *Q* (unadjusted)	13,735	1.80	1.39	1.27	0.44	29.50	-0.13^{***}
Industry-adjusted Tobin's *Q*	13,735	0.32	0.03	1.15	-1.52	25.98	-0.12^{***}
Assets	13,735	5,333.0	1,529.7	12,516.4	12.9	275,644.0	-0.12^{***}
Age	13,735	31.1	27.0	19.4	1.0	83.0	0.05^{***}
Leverage	13,735	0.24	0.25	0.16	0.00	0.85	0.10^{***}
ROA	13,735	0.14	0.14	0.09	-1.37	0.97	-0.04^{***}
R&D/Sales	13,735	0.03	0.00	0.13	0.00	2.99	-0.07^{***}
R&D Dummy	13,735	0.47	0.00	0.50	0.00	1.00	0.01
CAPEX/Assets	13,735	0.07	0.06	0.06	0.00	0.92	-0.03^{***}
Delaware	13,735	0.50	0.00	0.50	0.00	1.00	-0.14^{***}

Table 2: Firm Characteristics and the *E* Index

Firms are required to have at least four IRRC survey data between 1990 and 2007. Financial data are retrieved from Compustat and CRSP. The *E* index is the entrenchment index of Bebchuk, Cohen, and Ferrell (2008). The *O* index is constructed using all the other 18 anti-takeover provisions are tracked by the IRRC but not included in the *E* index. Firm-years are partitioned into 7 groups according to the level of the *E* index. Panel A reports the distribution of the *E* index. Panel B reports the mean values of variables for each E index group. Tobin's *Q* is defined as the market value of assets divided by the book value of assets. Industry-adjusted Tobin's *Q* is obtained by subtracting the median industry *Q* from a firm's *Q*. Industry is defined using the Fama-French (1997) classification of 48 industry groups. *Assets* is the total book value of assets. *Age* is defined as the number of years since the firm entered CRSP. *Leverage* is defined as total debt (the sum of short-term and long-term debt) divided by total assets. *ROA* is the operating income before depreciation and amortization over total assets. *R&D/Sales* is research and development expenses scaled by net sales. *R&D Dummy* is an *R&D* indicator variable that equals one if *R&D* expenses are missing, zero otherwise. *CAPEX/Assets* is capital expenditures scaled by total assets. *Delaware* is a dummy variable that equals 1 if a firm is incorporated in Delaware, and 0 otherwise. All dollar values (e.g.. total assets) are converted into year 2000 constant dollars using the GDP deflator. All variables except the governance indices and dummy variables are winsorized at the 1% level at both tails of the distribution.

Variables	*E* = 0	*E* = 1	*E* = 2	*E* = 3	*E* = 4	*E* = 5	*E* = 6

Panel A: Incidence of the *E* index							
Frequency	1,065	2,287	3,364	3,724	2,721	527	47
Percent (%)	7.8	16.7	24.5	27.1	19.8	3.8	0.3
Cumulative Percent (%)	7.8	24.4	48.9	76.0	95.8	99.7	100.0
Panel B: Mean values of variables							
O index	5.8	6.3	6.8	7.4	7.8	8.0	7.3
Tobin's *Q* (unadjusted)	2.18	1.96	1.82	1.77	1.61	1.43	1.47
Industry-adjusted Tobin's *Q*	0.67	0.47	0.33	0.27	0.17	0.06	0.06
Assets	10,192.0	6,659.2	5,330.6	4,273.9	4,144.0	3,674.2	1,144.3
Age	29.8	29.0	31.6	30.6	33.5	30.5	29.3
Leverage	0.22	0.23	0.24	0.25	0.27	0.28	0.18
ROA	0.16	0.14	0.14	0.14	0.14	0.14	0.15
R&D/Sales	0.04	0.05	0.03	0.03	0.02	0.01	0.01
R&D Dummy	0.43	0.48	0.51	0.45	0.46	0.52	0.47
CAPEX/Assets	0.08	0.07	0.07	0.07	0.07	0.08	0.08
Delaware	0.59	0.59	0.54	0.49	0.38	0.40	0.36

Table 3: Firm Value and Managerial Entrenchment

Firms are required to have at least four IRRC survey data between 1990 and 2007. Financial data are retrieved from Compustat and CRSP. The E index is the entrenchment index of Bebchuk, Cohen, and Ferrell (2009). $\hat{Q}$ is the industry-adjusted Tobin's Q obtained by subtracting the industry median Q from a firm's Q. $\hat{Q}_{[1980,1985]}$ is the average $\hat{Q}$ for the period 1980-1985. E_{1990} is the E index in 1990. Other control variables are defined in Table 1. Columns (1)-(3) report results of the Fama-MacBeth regressions. Column (4) reports results obtained using the mean-differencing estimator for the dynamic model with firm fixed effects. Year dummy variables are included to account for time specific effect. t-statistics are in parentheses in columns (1)-(4). Column (5) presents results produced by the two-step system GMM estimator. z-statistics in parentheses are computed using standard errors adjusted for finite-sample bias. Coefficients significant at the 10%, 5% and 1% levels are marked with *, **, *** respectively in superscripts.

	(1)	(2)	(3)	(4)	(5)
Dependent variables:	$\hat{Q}$	E	$\hat{Q}$	$\hat{Q}$	$\hat{Q}$
Estimators:	FM	FM	FM	Mean Differencing	GMM
E	-0.059***		-0.013**	0.007	-0.090**
	(-10.48)		(-2.36)	(0.68)	(-2.33)
$\hat{Q}_{t-1}$		-0.130***		0.398***	0.475***
		(-9.85)		(8.98)	(9.34)
E_{1990}			-0.016**		
			(-2.43)		
$\hat{Q}_{[1980,1985]}$			0.234***		
			(15.52)		
O index	0.002	0.221***	-0.014	0.010	0.004
	(0.73)	(57.86)	(-1.49)	(1.38)	(0.19)
Ln (*Assets*)	0.020	-0.080***	0.042**	-0.264***	-0.126**
	(1.04)	(-7.80)	(2.40)	(-10.64)	(-1.98)
Ln (*Age*)	-0.033	-0.066***	0.099***	0.026	0.189**
	(-1.70)	(-5.95)	(6.21)	(1.04)	(2.28)
Leverage	-1.013***	0.644***	-0.502***	-0.696***	-2.403***
	(-8.68)	(12.25)	(-7.15)	(-6.15)	(-4.41)
ROA	6.573***	0.359*	6.667***	2.963***	0.812*
	(20.55)	(1.96)	(34.56)	(10.69)	(1.93)
R&D/Sales	1.668***	-0.724***	0.811	0.567	-0.172
	(6.18)	(-3.98)	(1.40)	(1.35)	(-0.31)

R&D Dummy	-0.009	-0.080***	0.038**	0.030	-0.438*
	(-0.42)	(-6.48)	(2.31)	(0.60)	(-1.91)
CAPEX/Assets	0.283	0.567***	-0.398	0.317	1.319**
	(1.59)	(5.18)	(-1.48)	(1.59)	(2.05)
Delaware	-0.032*	-0.378***	0.024	0.110**	-0.018
	(-2.01)	(-18.64)	(1.45)	(1.98)	(-0.13)
N	13,735	13,735	9,710	13,735	12,493
R^2	0.37	0.16	0.48	0.71	-

Table 4: Firm Value and Managerial Entrenchment – Long Difference Approach

Firms are required to have at least four IRRC survey data between 1990 and 2007. Financial data are retrieved from Compustat and CRSP. The *E* index is the entrenchment index of Bebchuk, Cohen, and Ferrell (2009). $\hat{Q}$ is the industry-adjusted Tobin's *Q* obtained by subtracting the industry median *Q* from a firm's *Q*. Tobin's *Q* is defined as the market value of assets divided by the book value of assets. Other control variables are defined in Table 1. The model is $\Delta\hat{Q}_{i[t,t-k]} = \rho\Delta Q_{i[t-1,t-k-1]} + \beta\Delta E_{i[t,t-k]} + \gamma\Delta C_{i[t,t-k]} + \eta_{i[t,t-k]}$. All dependent and explanatory variables are the changes in the respective variables between the end of fiscal year *t* and the end of fiscal year *t-k* of firm *i* (*k* = 6, 8,or 10). We follow Hahn, Hausman, and Kuersteiner (2007) by estimating the equation with iterated two-stage least squares (2SLS). Year dummies are included to account for time specific effect. *t*-statistics are in parentheses. Coefficients significant at the 10%, 5% and 1% levels are marked with *, **, *** respectively in superscripts.

	(1)	(2)	(3)
Differencing length (*k*)	*k* = 6	*k* = 8	*k* = 10
$\Delta Q_{[t-1,t-k-1]}$	0.528***	0.684***	0.663***
	(41.02)	(45.50)	(27.39)
$\Delta E_{[t,t-k]}$	-0.104	-0.086**	-0.113***
	(-1.23)	(-2.38)	(-2.95)
ΔO index $_{[t-1, t-k-1]}$	0.027*	0.026	0.022*
	(1.66)	(1.49)	(1.79)
$\Delta Ln\ (Assets)_{[t-1, t-k-1]}$	-0.181***	-0.182***	-0.135***
	(-7.60)	(-7.91)	(-6.75)
$\Delta Ln\ (Age)_{[t-1, t-k-1]}$	0.139***	0.063	0.042
	(3.78)	(1.60)	(1.39)
$\Delta Leverage_{[t-1, t-k-1]}$	-0.549***	-0.584***	-0.409***
	(-6.72)	(-5.67)	(-4.23)
$\Delta ROA_{[t-1, t-k-1]}$	2.525***	2.235***	2.191***
	(15.17)	(12.18)	(8.99)
$\Delta R\&D/Sales_{[t-1, t-k-1]}$	0.884***	0.224	-0.059
	(6.31)	(1.04)	(-0.20)
$\Delta R\&D\ Dummy_{[t-1, t-k-1]}$	0.008	0.106*	-0.013
	(0.03)	(1.70)	(-0.23)
$\Delta CAPEX/Assets_{[t-1, t-k-1]}$	-0.015	-0.472**	-0.161

	(-0.08)	(-2.17)	(-0.72)
$\Delta Delaware_{[t-1,\ t-k-1]}$	0.010	0.073	0.006
	(0.14)	(0.91)	(0.08)
N	6,797	4,755	3,166

Table 5: Panel Vector Autoregressive (PVAR) Model

Firms are required to have at least four IRRC survey data between 1990 and 2007. Financial data are retrieved from Compustat and CRSP. The dependent variable $\hat{Q}$ in Column (1) is the industry-adjusted Tobin's Q obtained by subtracting the median industry Q from a firm's Q. Tobin's Q is defined as the market value of assets divided by the book value of assets. Industry is defined using the Fama-French (1997) classification of 48 industry groups. The E index is the entrenchment index of Bebchuk, Cohen, and Ferrell (2009). The following two-equation reduced-form PVAR model is estimated using the Holtz-Eakin, Newey, and Rosen (1988) methodology.

$$\begin{cases} \hat{Q}_{it} = a_{0t} + a_1\hat{Q}_{it-1} + b_1 E_{it-1} + \delta C_{it} + f_i + x_t + \varepsilon_{it} \\ E_{it} = c_{0t} + c_1\hat{Q}_{it-1} + d_1 E_{it-1} + \phi C_{it} + g_i + y_t + \omega_{it}, \end{cases}$$

where C are control variables defined in Table 1, x_t and y_t are year fixed effects, and f_i and g_i are unobserved firm fixed effects for $\hat{Q}$ and E, respectively. Firm fixed effects are removed by transforming all variables in the model in deviations from forward means. The lagged values of regressors are used as instruments to estimate the coefficients with the generalized method of moment (GMM). Year fixed effects are removed by subtracting the mean value of each variable computed for each year. z-statistics are in parentheses. Coefficients significant at the 10%, 5% and 1% levels are marked with *, **, *** respectively in superscripts.

	(1) Effect of Managerial Entrenchment on Firm Value (Dependent Variable: $\hat{Q}$)	(2) Effect of Firm Value on Managerial Entrenchment (Dependent Variable: E)
$\hat{Q}_{t-1}$	0.541***	-0.014*
	(20.45)	(-1.82)
E_{t-1}	-0.054***	0.851***
	(-2.59)	(57.36)
O index	0.010	0.108***
	(1.39)	(11.06)
Ln (*Assets*)	-0.225***	0.043***
	(-11.32)	(2.98)
Ln (*Age*)	-0.004	0.097***
	(-0.17)	(3.49)
Leverage	-0.706***	-0.027

	(-9.01)	(-0.46)
ROA	4.342***	-0.136
	(22.00)	(-1.35)
R&D/Sales	0.632**	-0.089
	(2.10)	(-1.62)
R&D Dummy	-0.015	0.047
	(-0.41)	(1.27)
CAPEX/Assets	1.434***	0.065
	(8.99)	(0.56)
Delaware	0.003	-0.225***
	(0.06)	(-3.66)
N	11,438	11,438

Table 6: Firm-level Granger Causality Tests

Firms are required to have at least four IRRC survey data between 1990 and 2007. Financial data are retrieved from Compustat and CRSP. We conduct the following bivariate Granger causality test for each individual firm.

$$\hat{Q}_t = a_0 + a_1\hat{Q}_{t-1} + a_2E_{t-1} + \delta C_t + \varepsilon_t \text{ ; } E_i = b_0 + b_1\hat{Q}_{t-1} + b_2E_{t-1} + \phi C_t + \omega_t$$

For the null hypothesis that the *E* index does not Granger-cause $\hat{Q}$, the percentages of firms with *p*-value significant at the 1% and 10% levels are reported in Column (1). Column (2) presents the summary for the reverse causality under the null hypothesis that $\hat{Q}$ does not Granger-cause the *E* index. The χ^2 statistics are obtained using Maddala and Wu's (1999) Fisher test that combines the *p*-values from all independent Granger causality tests at the firm level. χ^2 statistics significant at the 10%, 5% and 1% levels are marked with *, **, *** respectively in superscripts.

	(1)	(2)
Null hypotheses:	The *E* index does not Granger-cause $\hat{Q}$	$\hat{Q}$ does not Granger-cause the *E* index
Percentage of firms with *p*-value significant at the 1% level	22.7%	15.3%
Percentage of firms with *p*-value significant at the 10% level	52.4%	42.3%
X^2 statistics for combined *p*-values	5270.3***	4307.7***