Working Paper 9416

EXECUTIVE COMPENSATION: A CALIBRATION APPROACH

by Joseph G. Haubrich and Ivilina Popova

Joseph G. Haubrich is an economic advisor at the Federal Reserve Bank of Cleveland, and Ivilina Popova is a graduate student in the Department of Operations Research, Case Western Reserve University, Cleveland. The authors thank Ben Craig for helpful comments.

December 1994

Abstract

We use a version of the Grossman and Hart (1983) principal-agent model with 10 actions and 10 states to produce *quantitative* predictions for executive compensation. Performance incentives derived from the model are compared with the performance incentives of 350 firms from a survey by Michael Jensen and Kevin Murphy. The results suggest both that the model does a reasonable job of explaining the data and that actual incentives are close to the optimal incentives predicted by theory.

1. Introduction

Economists as far back as Adam Smith and Alfred Marshall have wondered about the incentives of top executives. The principal-agent model provides an elegant theory of incentives, but very little practical advice on how large those incentives should be. Because of this, much recent work aimed at reconciling theory with the data on executive pay has focused on the qualitative predictions of the theory. Papers such as Murdoch (1993), Habib (1993), and Kole (1993) document, for example, that high-growth firms provide more compensation via stock-option plans. Such work, though important, avoids key quantitative issues, such as whether compensation provides sufficient incentives to maximize firm value (see Jensen and Murphy [1990a,b] and Cowan [1992]). We directly address the quantitative issues by comparing actual CEO incentives with the predictions of a finite-state principal-agent model.

Haubrich (1994) took a preliminary step in this direction and showed how a simple parameterization of the Grossman and Hart (1983) principal-agent model produced performance incentives broadly in line with those documented by Jensen and Murphy. We generalize those results along two dimensions here. First, instead of a two-state model (with a closed-form solution), we consider a 10-state, 10-action model. This allows incentive pay to be nonlinear. Wang (1994) generalizes in a different direction, developing a multiperiod model with two states and two actions. Second, where Haubrich (and Wang) made a simple comparison between the model and the mean of Jensen and Murphy's performance/pay ratio, we take a calibration approach. Specifically, we choose parameters that minimize the distance between the data and the model's output for 350 firms, explicitly comparing the distributions. This calibration approach has an added benefit: It provides an estimate of CEO productivity, a central but difficult aspect of executive pay.

The calibration begins with 350 firms chosen from Jensen and Murphy's "New Survey of Executive Compensation" (1990b). These firms appeared in the dataset of the Center for Research in Security Prices (CRSP) long enough for us to calculate the standard deviation of shareholder value. This variance, along with some global parameters, pins down the principal-agent problem for each firm. The program then solves the problem for many values of risk aversion and for another parameter that measures CEO productivity. It conducts a grid search for the values that minimize the distance between the 350 predictions and the actual values calculated by Jensen and Murphy. We use several metrics, including the sum of squared errors and the difference of sample means.

Determining if the principal-agent model correctly describes executive incentives does matter. Competing models have very different implications. Jensen (1989) argues that political constraints keep firms from tying compensation closely enough to firm performance, and that as a consequence, leveraged buyouts will replace corporations. The underinvestment model of Myers and Majluf (1984), by contrast, argues that compensation is tied *too* closely to firm performance. The desirability of proposals pending before Congress -- and shareholders -- depends on the resolution of this issue.

2. The Model and Solution Technique

The key question in the modeling of executive compensation was aptly put by Marglin (1974): "What do bosses do?" Grossman and Hart's (1983) answer is that bosses raise the likelihood of good outcomes. In their model, increased effort by the agent increases the probability of good states occurring. The boss adds value to the firm, but observing output does not let you infer his actions. A good outcome may reflect luck as well as hard work.

2.1. A Discrete Principal-Agent Model

More formally, we assume that the firm has 10 profit levels: $q_1 < q_2 < \ldots < q_{10}$. The action set A consists of 10 possible actions: $a_1, a_2, \ldots, a_{10}$. The restrictions come from limitations of GAMS, the software we use to solve the nonlinear programming problem (see Brooke, Kendrick, and Meeraus [1992]). If we used, for example, the industrial version of GAMS, the number of profit levels and actions could be increased significantly. $\pi_i(a)$ denotes the probability of state (i.e., profit level) i given action a. To forestall some technical problems, $\pi_i(a)>0$ for all a and i.

The agent's utility depends on actions and income, expressed as U(a,I). Grossman and Hart consider a fairly general form, but for calibration purposes, we use constant absolute risk aversion (CARA), $(a,I) = -e^{-\gamma(I-a)}$, in which effort appears as negative income.[1] Choosing the correct functional form has its difficulties, but Grossman and Hart find this utility function particularly useful in principal-agent theory, in part because it has a mutiplicatively separable representation, U(a,I)=K(a)V(I). In addition, since compensation depends on the disutility of effort, treating effort as negative income makes the resulting contract easier to interpret. For a more extensive discussion of the choice, see Haubrich (1994). The agent also has a reservation utility $\overline{U}$, derived from alternative employment or a leisure-time activity.

Grossman and Hart concentrate on the cost of getting the agent to choose a particular action. When the principal observes the action, the cost is simply the agent's reservation price for action a, denoted CFB(a)=h[$\overline{U}$/K(a)], where h=V^{-1}.

The point of the principal-agent problem is that the principal cannot observe the action taken by the agent. She can only make payment dependent on the observed

[1] Wang (1994) uses the slightly more general specification $-e^{-\gamma(I-\varphi a)}$.

output state. This incentive scheme $\{I_1 \ldots I_{10}\}$, a set of payments contingent on the state, gives the agent utility levels $v_1=V(I_1) \ldots v_{10}=V(I_{10})$.

Although the principal cannot observe the action, she can design the incentive scheme so that the agent chooses a particular action. The expected value of payments to the agent defines the second-best cost of an action a, C(a). For a given action, the incentive scheme minimizes the principal's cost (the expected value of the incentive scheme) subject to three constraints.

The first is the incentive compatibility constraint, which states that the agent takes action a only if it gives a higher payoff than any other action. The second and third constraints are the participation constraints, which state that the agent must get a certain minimum utility and that some income level produces that utility.

Several incentive schemes (I or v sets) may induce the agent to choose action a (that is, to implement a). Define C(a) as the least costly of these (technically, the greatest lower bound [infinum or inf] of $\sum \pi_i[a]h[v_i]$) on the constraint set. If the principal cannot induce action a (an empty constraint set), set C(a) to infinity.

A little terminology about the principal completes the basic notation. The risk-neutral principal receives the gross profits q_i, so her expected gross benefit from the agent's action is $B(a)=\sum \pi_i(a)q_i$. Her expected net benefit, B(a)-C(a), subtracts the cost of the action, and the (second-best) optimal action maximizes her expected net benefit.

Grossman and Hart take a simple approach to solving the principal-agent problem. First, they compute the cost C(a) for each action a. Then, they optimize the net benefit, B(a)-C(a), over all actions a.

Of central concern here is the proper specification of the $\pi_i(a)$ function: Measuring the CEO's contribution to the firm is the most problematic aspect of calibrating the principal-agent problem. Perhaps the best evidence comes from studies of CEO turnover, where the effects, though at times substantial, are generally small but significant (Weisbach [1988]). At one extreme, when James Crosby, the controlling shareholder and chief executive of Resorts International, died, Resorts' stock increased by 37 percent in a single day (Holderness and Sheehan [1991]).

With 10 states and 10 actions, the specification problem becomes even more difficult, since there are many ways that an action can make good states more likely. Finding an intuitively appealing specification proved difficult. Even after restricting the search to probability structures that satisfy the "spanning condition" (SC), where better performance means higher pay, many structures had only degenerate feasible solutions or implied implausible CEO productivity.

This point -- how CEO effort benefits the firm -- is clearly the major difficulty in using the direct quantitative approach. Squarely confronting that problem gives us a better idea of what we lack, both in terms of the data we would like to have and in regard to the theoretical concepts that need clarification.

We generated the probabilities $\pi_i(a)$ that satisfy the SC in Grossman and Hart, namely, that there exist vectors $\hat{\pi}, \bar{\pi}$ such that for each action $a \in A$,

(1) $\pi(a) = \lambda(a)\hat{\pi} + [1 - \lambda(a)]\bar{\pi}$ for some $0 \leq \lambda\ (a) \leq 1$

and

(2) $\dfrac{\hat{\pi}_i}{\bar{\pi}_i}$ is nonincreasing in I.

This precise form is a technical condition to ensure that the incentive scheme increases in effort. The function $\lambda(a)$ measures the effects of CEO effort and describes how much better the probability distribution gets as the CEO expends more effort and takes increasingly difficult actions. We use a $\lambda(a)$ function that is decreasing in a. This means that increasing *a* moves the probability distribution away from the "bad" vector $\hat{\pi}$ and closer to the "good" vector $\bar{\pi}$, making good states more likely.

Clearly, a major parameter in the calibration exercise is the function λ(a). We did an extensive search in this direction, initially starting with the linear functions λ(a)=a and λ(a)=1-a. Unfortunately, in both cases the optimal solution was action 1 or action 10. For this function, the problem reduces to the two-state case.[2] Therefore, we use a nonlinear function to avoid the problem.

2.2. Solution Procedures

All the pieces are now in place to delineate the nonlinear programming problem. Given specific risk aversion γ and specific $\overline{U}$, we have:

$$\min \sum_{i=1}^{10} \pi_i(a^*)\left(-\frac{\ln(-v_i)}{\gamma}\right)$$

$$s.t.$$

$$e^{\gamma a^*}\sum_{i=1}^{10} \pi_i(a^*)v_i \geq e^{\gamma a}\sum_{i=1}^{10} \pi_i(a)v_i, \forall a \in A$$

$$e^{\gamma a^*}\sum_{i=1}^{10} \pi_i(a^*)v_i \geq e^{-\gamma\overline{U}}$$

For every action a^* from the action set **A**, GAMS produced the optimal solution whenever the problem was feasible. It did this by using one of its solvers, MINOS, which implements some of

---

2 We do not yet have analytical proof of this, but it holds true for every set of parameters we have checked. Let's take a closer look at the linear case:

$$\lambda(a) = 1 - a, \pi(a) = (1-a)\hat{\pi} + a\overline{\pi}.$$

The function B(a) has the following form:

$$B(a) = \sum_{i=1}^{10} \pi_i(a)q_i = \sum_{i=1}^{10} [(1-a)\hat{\pi}_i + a\overline{\pi}_i]q_i = \underbrace{\sum_{i=1}^{10} \pi_i q_i}_{K_1} + \underbrace{a\sum_{i=1}^{10} [\overline{\pi}_i - \hat{\pi}_i]q_i}_{K_2} = K_1 + K_2 a.$$

Here, K_1 and K_2 are constants not depending on action.

Hence, $\max_a$ {B(a) - C(a)}= $\max_a$ {$K_1 + K_2$a - C(a)}. In every case we checked, all of the problems except those for actions 1 and 10 are infeasible, and so the values of their cost functions are infinite. Therefore, the maximum will be obtained at either action 1 or action 10.

the most popular algorithms for solving nonlinear programming problems, i.e., problems in which at least the objective function or the constraint set is a nonlinear function. In our case, it is obvious that the objective function is nonlinear. GAMS specifies the above general model for every particular action from the action set. It then tries to find the optimum. As usual, the first step is to find a feasible solution. If successful, the next step is to find the optimum. In this specific example, because the objective function is convex $\left(-\frac{\ln(-v_i)}{\gamma}\right)$, the Kuhn-Tucker-Karush theorem guarantees that the optimal solution exists, and this is the result that GAMS produces. From this, GAMS generates the cost-function value C(a), where $C(a) = \min\sum_{i=1}^{10}\pi_i(a)\left(-\frac{\ln(-v_i)}{\gamma}\right)$ for every action from the action set, together with the vector $v_1, v_2, \ldots, v_{10}$.

For some configurations, however, there are no points $v_1, v_2, \ldots, v_{10}$ that simultaneously satisfy the constraints. Some actions "a" cannot be implemented by the principal at any cost. For those cases, we assign an infinitely large number to be the value of the objective function, i.e., $C(a)=\infty$.

The next step is to choose which action to implement, that is, to choose $a \in A$ so as to maximize B(a) - C(a), where $B(a) = \sum_{i=1}^{10}\pi_i(a)q_i$.

As mentioned before, in the case of linear $\lambda(a)$ functions, only actions 1 and 10 are feasible. A nonlinear function avoids the reduction to the two-state case. We use $\lambda(a) = e^{-\delta a}$, where δ is a parameter measuring how additional effort affects output, and $a \in A=\{0.1,0.2...0.9,1.0\}$. Figure 1 plots $\lambda(a)$ for three representative δ's.

In our empirical study, we chose the following values for the vectors $\hat{\pi}$ and $\bar{\pi}$:

$\hat{\pi}$=[0.145, 0.135, 0.125, 0.115, 0.105, 0.095, 0.085, 0.075, 0.065, 0.055]

$\bar{\pi}$=[0.055, 0.065, 0.075, 0.085, 0.095, 0.105, 0.115, 0.125, 0.135, 0.145].

The above briefly describes one cycle of our procedure. The actual steps are

Step 1 Choose some initial starting values for risk aversion γ and CEO productivity δ.

Step 2 Generate the probability distribution from equation (1) for every action a.

Step 3 Use GAMS to solve the nonlinear problem (NP) for every action a. Produce as an output C(a) and $v_1, v_2, \ldots, v_{10}$.

Step 4 Find $\max_{a \in A} \{B(a) - C(a)\}$. Obtain the second-best optimal action $\hat{a}$.

Step 5 Compute SSE, BAR, and DSSQ statistics (defined below) using data for 350 companies.

Step 6 Increment γ and δ.

Detailed description of Step 5:

The data comprised those 350 companies in Jensen and Murphy's (1990b) "New Survey of Executive Compensation" for which we could extract shareholder value from the CRSP database. We extracted the stock price and the number of shares outstanding for the last trading day of each quarter for the years 1982 - 1990, then used this information to generate the profit levels q_i based on the standard deviations computed for every company from the set. This meant rescaling the q_i's given $\hat{\pi}$. We next compared the profit shares produced by our procedure with the real profit shares obtained from Jensen and Murphy. Following their approach, we define profit share as the fraction of increased shareholder wealth that the CEO receives in total compensation. In our model, that translates to $\frac{I_{10} - I_1}{q_{10} - q_1}$. We arrive at this by using the following functions: If we denote by $x_1, x_2, \ldots, x_{350}$ the profit shares from Jensen/Murphy and by $y_1, y_2, \ldots, y_{350}$ the profit shares from our procedure, then

$$SSE = \sum_{i=1}^{350} (x_i - y_i)^2 \tag{1}$$

$$BAR = |\bar{y} - \bar{x}| \tag{2}$$

$$DSSQ = |S_y^2 - S_x^2| \tag{3}$$

These three statistics are actually functions of γ and δ. The procedure was to minimize them with respect to δ and γ.

The three metrics all have a natural interpretation. The first, the sum of squared errors, is the standard quadratic loss function. The others attempt to match specific moments. *BAR* matches the means, and *DSSQ* matches second moments.

3. Results

In looking at the results, three questions stand out: 1) What parameters does the calibration choose, 2) How closely do we match the data, and 3) What does the optimal compensation contract look like? Answering these questions resolves the deeper issue -- What have we learned about principal-agent theory and executive compensation? We see where the theory falters and what missing factors hold promise of better fits.

3.1. Basic Results

As described in section 2, the calibration approach searches across parameter combinations for the values that best match the observed profit shares. Figures 2, 3, and 4 illustrate the procedure by plotting the three different loss functions against risk aversion. Figure 2 plots the sum of squared errors, figure 3 plots the absolute difference in sample means, aiming at matching the first moment, and figure 4 plots the absolute difference in sample standard deviations, aiming to match the second moment. Table 1 reports the underlying numbers.

The functions differ, of course, but a common pattern emerges. The global minimum occurs at a low level of risk aversion. For the sum of squared errors, there is a global minimum at 0.125. The two moment-matching cases show lower risk aversion. The mean case selects 0.025, and the variance case selects 0.025, a boundary value, suggesting that the actual minimum may occur at even lower values. The results for δ show greater variability. The SSE and BAR metrics produce values of 10 and 13.5, but matching variance produces a lower value of 3.5.

What do these parameter values tell us? The risk aversion parameters may initially seem rather low, but they represent absolute risk aversion, not the relative risk aversion calculated in most consumption and asset-pricing studies. To convert absolute to relative risk aversion, we multiply by wealth. One measure of wealth, the median value of CEO stockholdings, is $3.5 million in the Jensen and Murphy sample. Since our paper works in million-dollar units, this suggests adjusting risk aversion by a factor between one and ten. With this in mind, the numbers look reasonable but still low.

The meaning of the parameter δ, labeled CEO productivity, is less obvious. It describes how increased effort heightens the probability of good states, moving away from probability vector $\hat{\pi}$ toward $\overline{\pi}$. For the SSE optimal value of δ=10, for example, with the lowest level of effort a_1=0.1, the probability of the best state is 0.11; for a_2=0.2, the probability is 0.13; and for a_{10}=1, the probability is 0.14. For the DSSQ optimal value of δ=3.5, the corresponding good-state probabilities are 0.08, 0.10, and 0.14.

How well does this calibration match the data? Table 1 and figures 2, 3, and 4 provide one set of answers (since they are explicit metrics), but these are hard to interpret. Another way to look at the match is as follows. The average actual performance/pay ratio for the 350 firms in the sample is 0.01003 (the CEO gets $10.03

for every \$1,000 increase in shareholder value), and the sample standard deviation is 0.032. For the mean (BAR) case, the corresponding figures are 0.01004 and 0.008. The calibration matches the mean quite well, to within 1 cent per \$1,000 of shareholder wealth. It seriously understates the standard deviation, however, a topic we pursue in the next section. The calibration designed to match standard deviations did better, of course.

Figure 5 plots the optimal incentive scheme (compensation contract) for each metric and lists the optimal action chosen under each scheme. The incentive schemes are monotonic, meaning that the agent gets paid more in good states. (This must happen because the probabilities satisfy the SC.) However, they are also nonlinear: A given increase in firm profits (a constant difference in gross profits from one state to the next) corresponds to a different change in the agent's income.

Figure 5 also indicates that the linear compensation scheme, implicitly assumed in Jensen and Murphy's empirical work and explicitly assumed in Rosen (1990), Haubrich (1994), and Wang (1994), is not the fully optimal contract. The CEO receives greater rewards for improving a bad state than for improving a good state. Kaplan (1994) finds evidence that incentives may differ across states in this manner. Figure 5 also suggests that the model takes this too far, overemphasizing the negative payments in bad states. Wang (1994) argues that a dynamic approach avoids this problem.

3.2. Comparing Distributions

Formal metrics have the advantage of being explicit, but they can also hide information about the distributions being compared. The problem boils down to comparing two distributions. We use a series of graphs developed by statisticians to examine the total distributions in more detail. Figure 6 shows a percentile plot of the actual profit shares and the profit shares generated by the model (SSE case). A percentile plot graphs the value against its percentile, allowing easy comparison between percentiles. Examining two together provides a picture of how the distributions differ, even if single numbers such as means match up closely. For the actual values, notice the small number of extreme values at the top. For the model, notice the absence of both negative and very high values. No predicted profit share exceeds 0.05, while 30 actual values do so, reaching as high as 0.43. In general, figure 6 shows that the model slightly overpredicts profit shares for most companies, but never produces the large profit shares found at the high end of the data. Our judgment is that in these extreme cases, such as An Wang, where the CEO is also a substantial stockholder of the firm, the distinction between principal and agent breaks down, making our model inappropriate. These major errors also explain why matching the standard deviation and the sum of squared errors is difficult.

Figures 7 and 8 take the comparison one step further. Figure 7 shows a *percentile comparison graph* (see Cleveland [1985]), which plots the ordered values of one dataset against the ordered values of another. Identical distributions result in a perfect x = y line, while a small amount of noise results in random deviations around that line. One defect of the graph is that the human eye is a poor judge of distance from a slanted line. The Tukey Sum-Difference graph (figure 8) resolves the problem, plotting y_i-x_i against y_i+x_i and in effect rotating the 45° line to the horizontal. Notice that for most values, the model predicts a profit share that is a little too high. For larger values, the model underpredicts profit share. This problem gets worse for larger values.

3.3. Truncated Sample Results

The comparisons in section 3.2 indicate that the model fails in cases of high performance pay. In these cases, the CEO is also a (often *the*) major stockholder in the firm, a point emphasized by looking at the names: Barron Hilton of Hilton Hotels, An

Wang of Wang Laboratories, and Richard Timken of Timken Industries. It no longer seems clear that the CEO is the agent of the stockholders, and it is not surprising that the model breaks down.

To account for this, we truncated our sample by removing all executives (12) with a performance/pay ratio above 0.05 ($50/$1,000). After checking each to make certain that the high ratio was due to large shareholdings, we recalibrated the model. The results are presented in table 2 and figures 9 and 10.

Table 2 shows that in addition to matching the mean, the model can also match the standard deviation of the truncated sample very closely. The percentile plots show an even closer match, but a similar pattern to before: overprediction of profit shares for most firms, underprediction for the highest. Figures 9 and 10 compare the actual and predicted distributions for the truncated sample. The extreme outliers are gone, though the model again does worse at high levels. Note the relative paucity of profit shares above 0.01.

The model can clearly generate a distribution of profit shares that closely matches the actual distribution. This is not the same as accurately predicting each firm's profit share, however. Figure 11 illustrates this, plotting the predicted profit share for each firm against its actual profit share. While the model produces a distribution similar to that found in the data, a firm with a high predicted profit share may or may not have a high actual profit share.

3.4. Parameter Uncertainty

Calibration chooses parameters, but some degree of uncertainty necessarily surrounds the parameters chosen. More important, because we care little about the uncertainty in γ and δ *per se*, there is uncertainty in the predicted profit shares.

Cecchetti, Lam, and Mark (1993) correctly emphasize that this uncertainty has two parts. One part arises because the input data, the variance of shareholder value, is only an estimate of the true variance and thus has its own uncertainty, i.e., its own finite sample distribution. The other part exists because of the error in the final estimate; we estimate γ and δ with error, so the predicted distribution of profit shares, which depends crucially on these parameters, also has an associated error.[3]

Simulations can address the first source of uncertainty. For a normal population, the sample variance has a χ^2 distribution, or more precisely, $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2$ with (n-1) degrees of freedom. We took 100 draws from a $\chi_{349}{}^2$ distribution and rescaled the sample variances to produce new input data. This new data, in conjunction with the old optimal contract, yielded new predictions of profit shares and a corresponding value for the distance between those predictions and the actual profit shares. Table 3 reports the results -- how the distribution varies when the underlying variance changes, given particular values for γ and δ. The first panel reports the findings for the γ and δ that minimize the SSE in the actual data, while the second and third panels report the combinations that minimize differences in means and variances.

We find that the uncertainty does matter: The variation around the optimal is nontrivial. This is particularly noticeable in the mean and variance case, which matched the original data most closely. For example, originally, mean-predicted pay matched the actual mean to within 1 cent in \$1,000; changing the variances dropped the match to between \$3 and \$5 per \$1,000.

[3]A procedure such as the Generalized Method of Moments would explicitly introduce these two types of uncertainty. Unfortunately, our model, both because of its particular form and because it has no closed-form solution, makes it difficult (if not impossible) to apply the required orthogonality conditions.

The other source of uncertainty, the estimated values of γ and δ, already showed up in table 1, which revealed how the difference between predicted and actual values changes with shifts in γ and δ -- and how the distribution depends on the estimated parameters, which is exactly what we wish to know.

4. Conclusion

No one would deny the insights gained from looking at the qualitative correspondence between economic theory and executive compensation. Still, as recognized by Jensen and Murphy (1990a), exercises such as correlating CEO compensation with firm risk can run afoul of the biblical injunction about straining gnats and swallowing camels. The finer nuances may not matter if the CEO has inadequate incentives. Our results show the feasibility of using calibration to undertake a direct, quantitative approach.

Beyond demonstrating feasibility, calibration produces some useful information by forcing us to look at questions that would not come up in most purely econometric settings. In so doing, we get an estimate of CEO productivity: By taking the best action rather than the worst, the CEO increases the probability of the most profitable outcome from 0.08 to 0.14. We also find that theory predicts a decidedly nonlinear pay schedule for top executives, one that rewards improvements from bad outcomes more than improvements from good outcomes.

Our results suggest that standard principal-agent theory predicts low profit shares for CEOs. Results such as those of Jensen and Murphy should not be taken as strong evidence that CEO compensation schemes are seriously out of line with proper incentives.

Quantitatively, the theory can be said to match the data successfully by two criteria. First, by matching moments, the mean of the predicted values differs from the mean of the actual values by only 1 cent in $1,000. Next, on a more subjective level, the percentile comparison plots show broad coherence between actual and predicted distributions. This occurs despite ignoring differences known to affect CEO pay, such as company size (Rosen [1990]) and CEO tenure (Gibbons and Murphy [1992]).

Calibration has contributed substantially to our understanding of asset pricing and business cycles. We believe that taking the quantitative predictions of theoretical models seriously can also contribute to the study of executive compensation and, more broadly, to corporate finance as well.

REFERENCES

Brooke, Anthony, David Kendrick, and Alexander Meeraus, *GAMS: A User's Guide*, release 2.25, Scientific Press, San Francisco, 1992.

Cecchetti, Stephen G., Pok-sang Lam, and Nelson C. Mark, "The Equity Premium and the Risk-Free Rate: Matching the Moments," *Journal of Monetary Economics*, vol. 31, no. 1, February 1993, pp. 21-45.

Cleveland, William S., *The Elements of Graphing Data,* Wadsworth Advanced Books and Software, Monterey, Calif., 1985.

Cowan, Alison Leigh, "The Gadfly C.E.O.'s Want to Swat," *New York Times*, February 2, 1992.

Gibbons, Robert, and Kevin J. Murphy, "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, vol. 100, no. 3, June 1992, pp. 468-505.

Grossman, Sanford J., and Oliver D. Hart, "An Analysis of the Principal-Agent Problem," *Econometrica*, vol. 51, no. 1, January 1983, pp. 7-45.

Habib, Michel, "Performance Payment and Firm Characteristics," London Business School, unpublished manuscript, 1993.

Haubrich, Joseph G., "Risk Aversion, Performance Pay, and the Principal-Agent Problem," *Journal of Political Economy,* vol. 102, no. 2, April 1994, pp. 258-276.

Holderness, Clifford G., and Dennis P. Sheehan, "Monitoring an Owner: The Case of Turner Broadcasting," *Journal of Financial Economics*, vol. 30, 1991, pp. 325-346.

Jensen, Michael C., "Eclipse of the Public Corporation," *Harvard Business Review*, September/October 1989, pp. 61-74.

_______, and Kevin J. Murphy, "Performance Pay and Top-Management Incentives," *Journal of Political Economy,* vol. 98, no. 2, April 1990a, pp. 225-264.

_______, and _______, "A New Survey of Executive Compensation: Full Survey and Technical Appendix," Harvard Business School, Division of Research, Working Paper 90-067, 1990b.

Kaplan, Steven N., "Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States," *Journal of Political Economy*, vol. 102, no. 3, June 1994, pp. 510-546.

Kole, Stacey R., "The Bundling of Compensation Plans," University of Rochester, Simon School, unpublished manuscript, June 1993.

Marglin, Stephen A., "What Do Bosses Do?: The Origins and Functions of Hierarchy in Capitalist Production," *Review of Radical Political Economy,* vol. 6, Summer 1974, pp. 60-112.

Murdoch, Jane, "Factors Explaining the Variation in the Use of Executive Incentive Contracts," Charles River Associates, Boston, unpublished manuscript, June 1993.

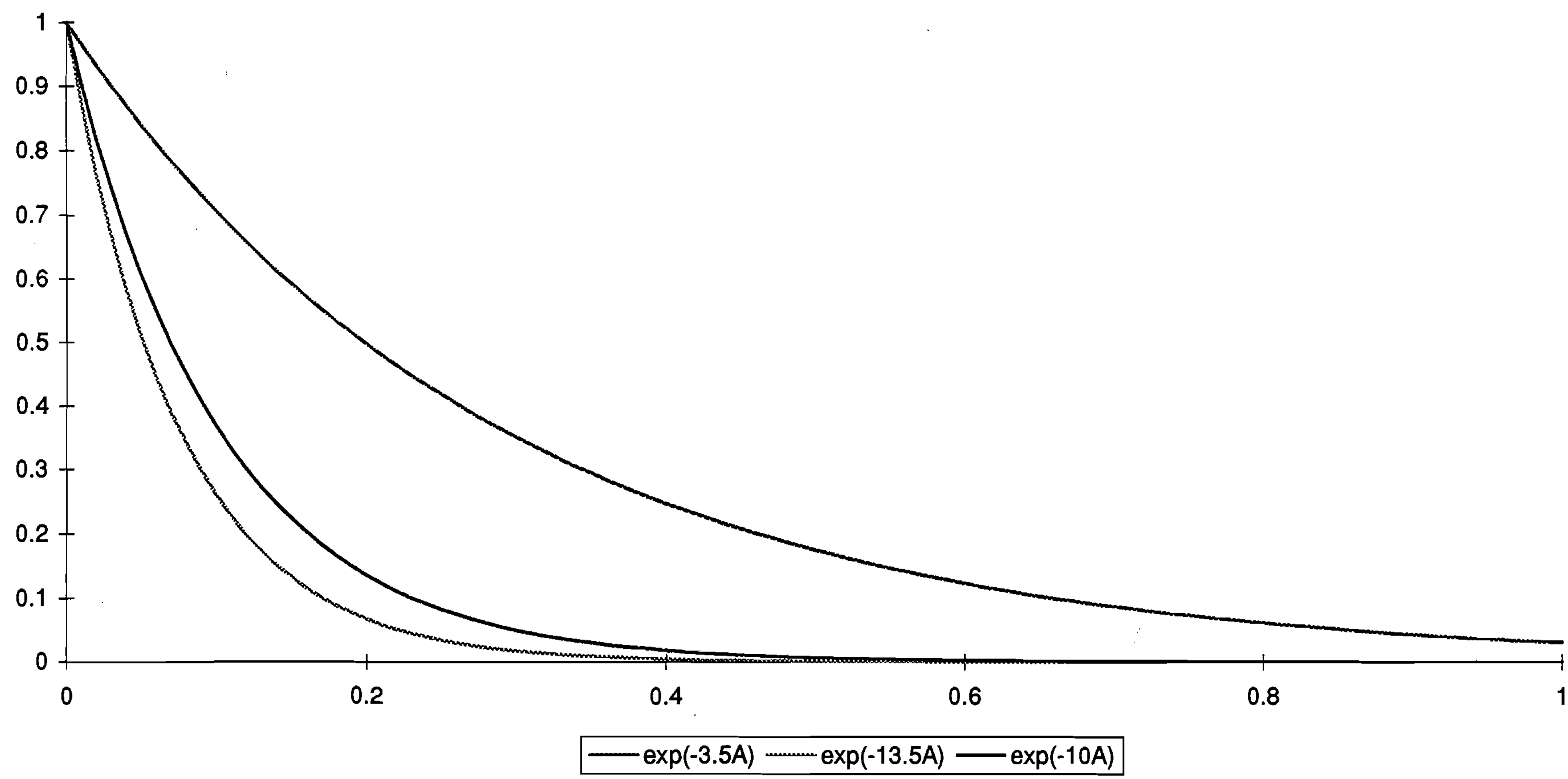
Myers, Stuart C., and Nicholas S. Majluf, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics*, vol. 13, 1984, pp. 187-221.

Rosen, Sherwin, "Contracts and the Market for Executives," NBER Working Paper No. 3542, December 1990.

Wang, Cheng, "Incentives, CEO Compensation, and Shareholder Wealth in a Dynamic Agency Model," University of Iowa, Department of Economics, unpublished manuscript, April 1994.

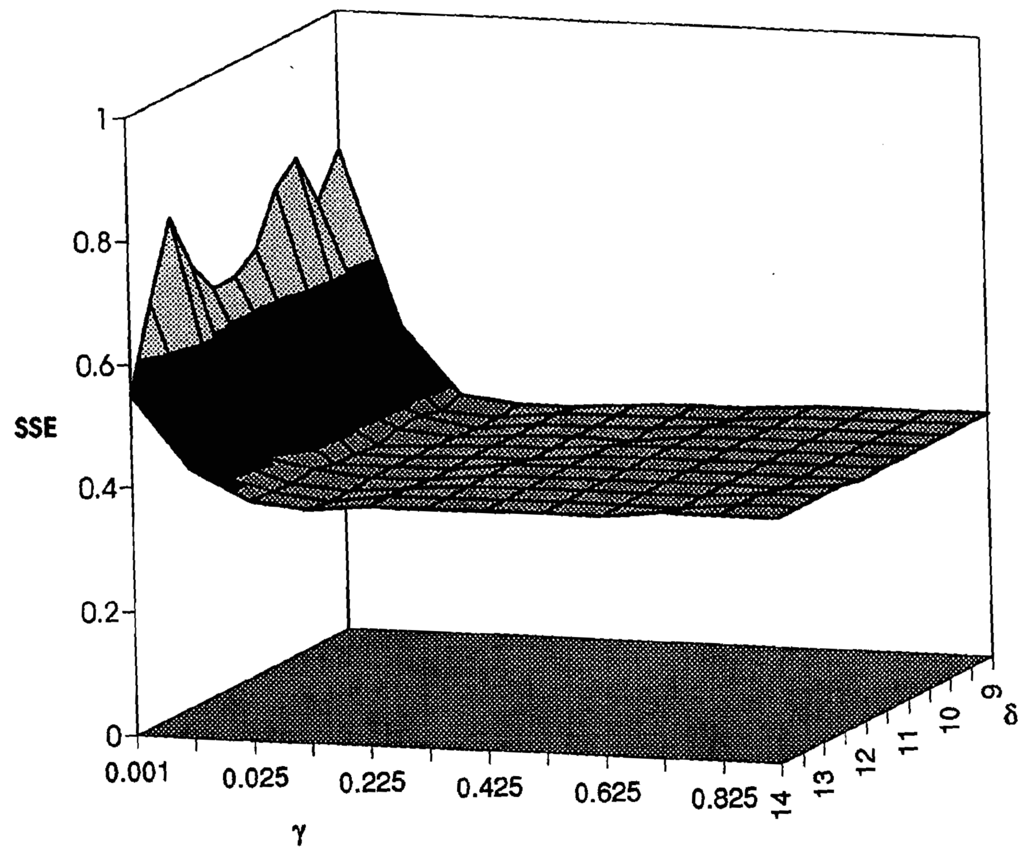
Weisbach, Michael S., "Outside Directors and CEO Turnover," *Journal of Financial Economics,* vol. 20, 1988, pp. 431-460.

Figure 1: Different Lambda Functions



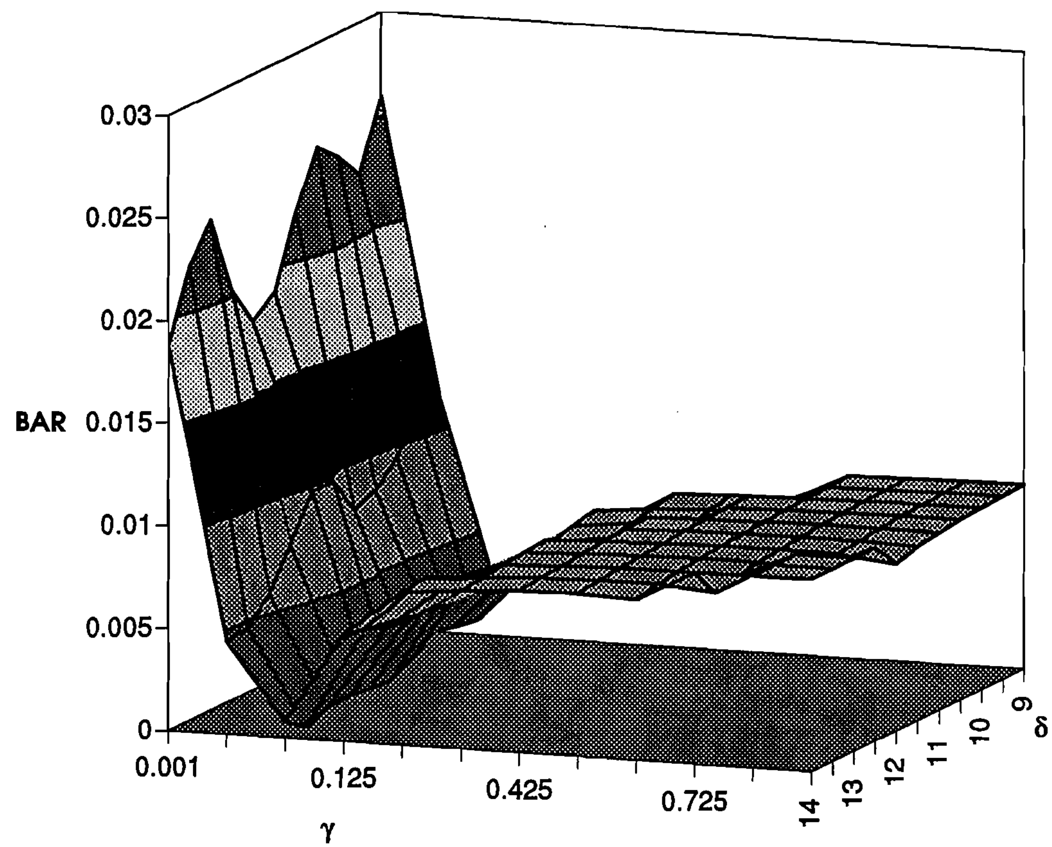
Source: Authors' calculations.

Figure 2: 3D Plot SSE



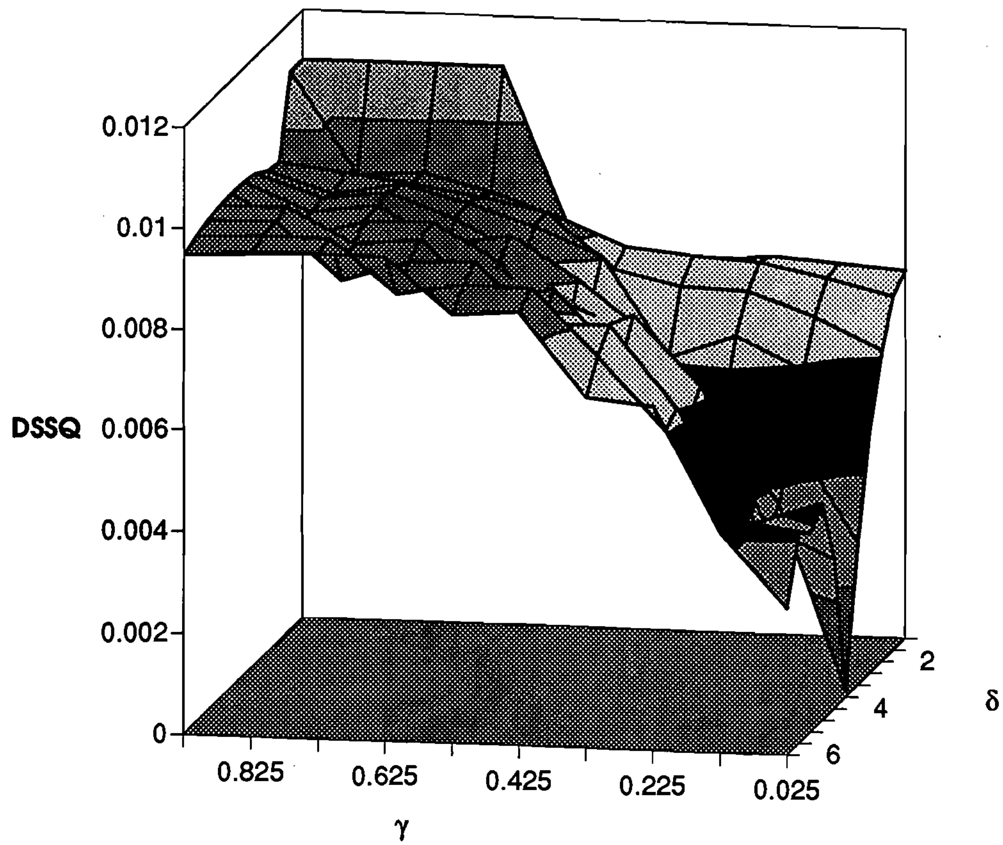
Source: Authors' calculations.

Figure 3: 3D Plot BAR



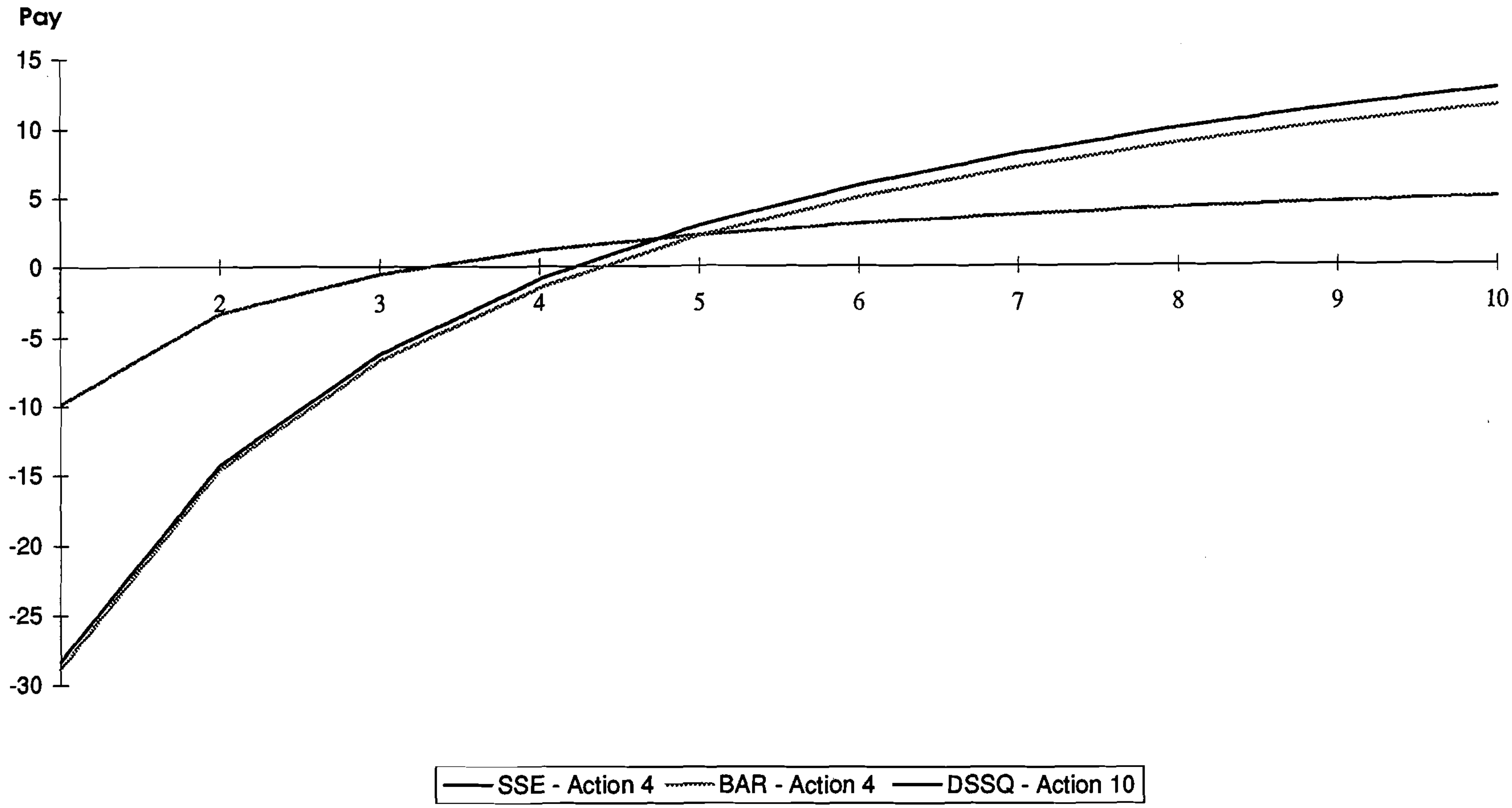
Source: Authors' calculations.

Figure 4: 3D Plot DSSQ



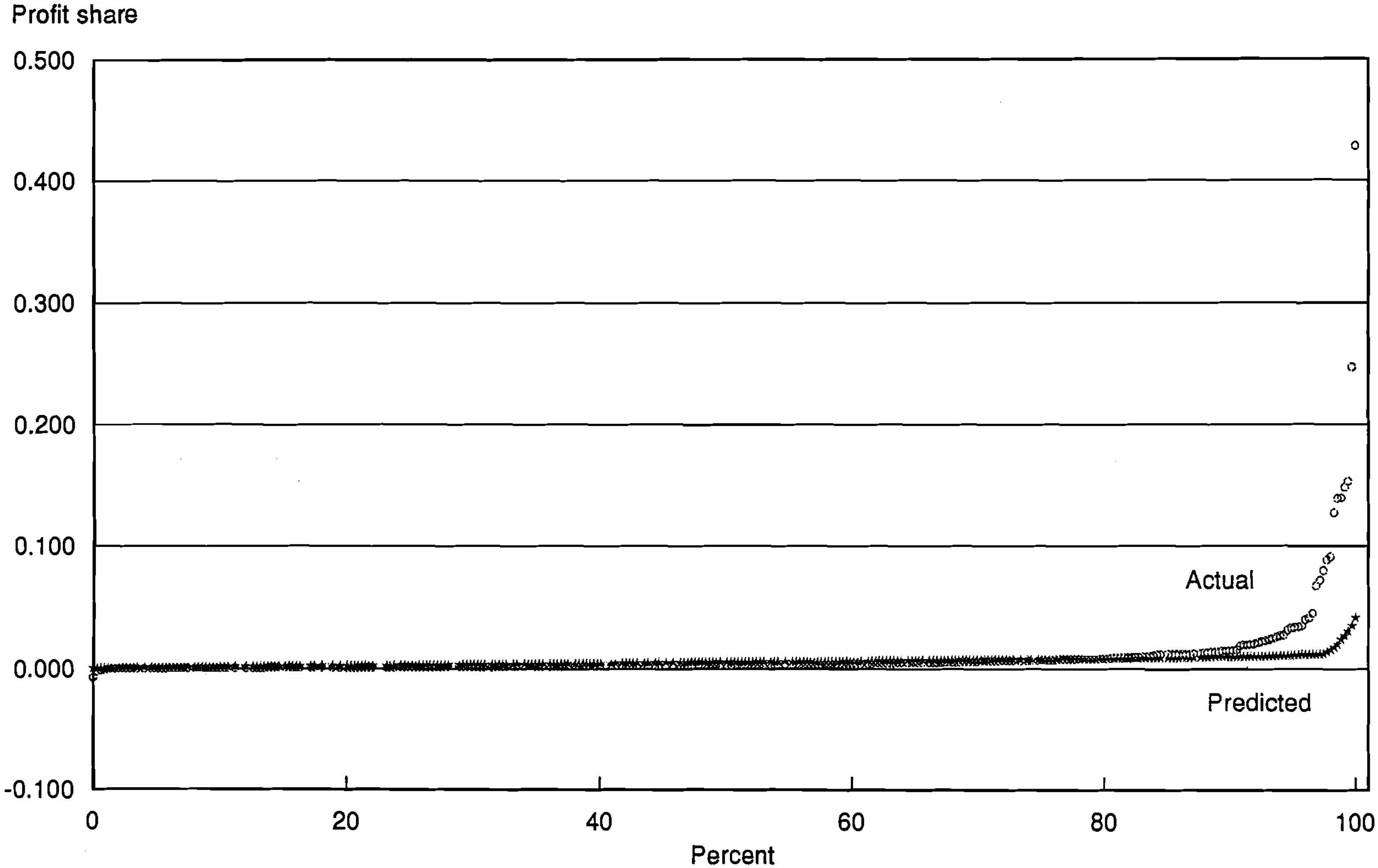
Source: Authors' calculations.

Figure 5: Incentive Pay in Each State for the Different Optimal Values of Gamma and Delta



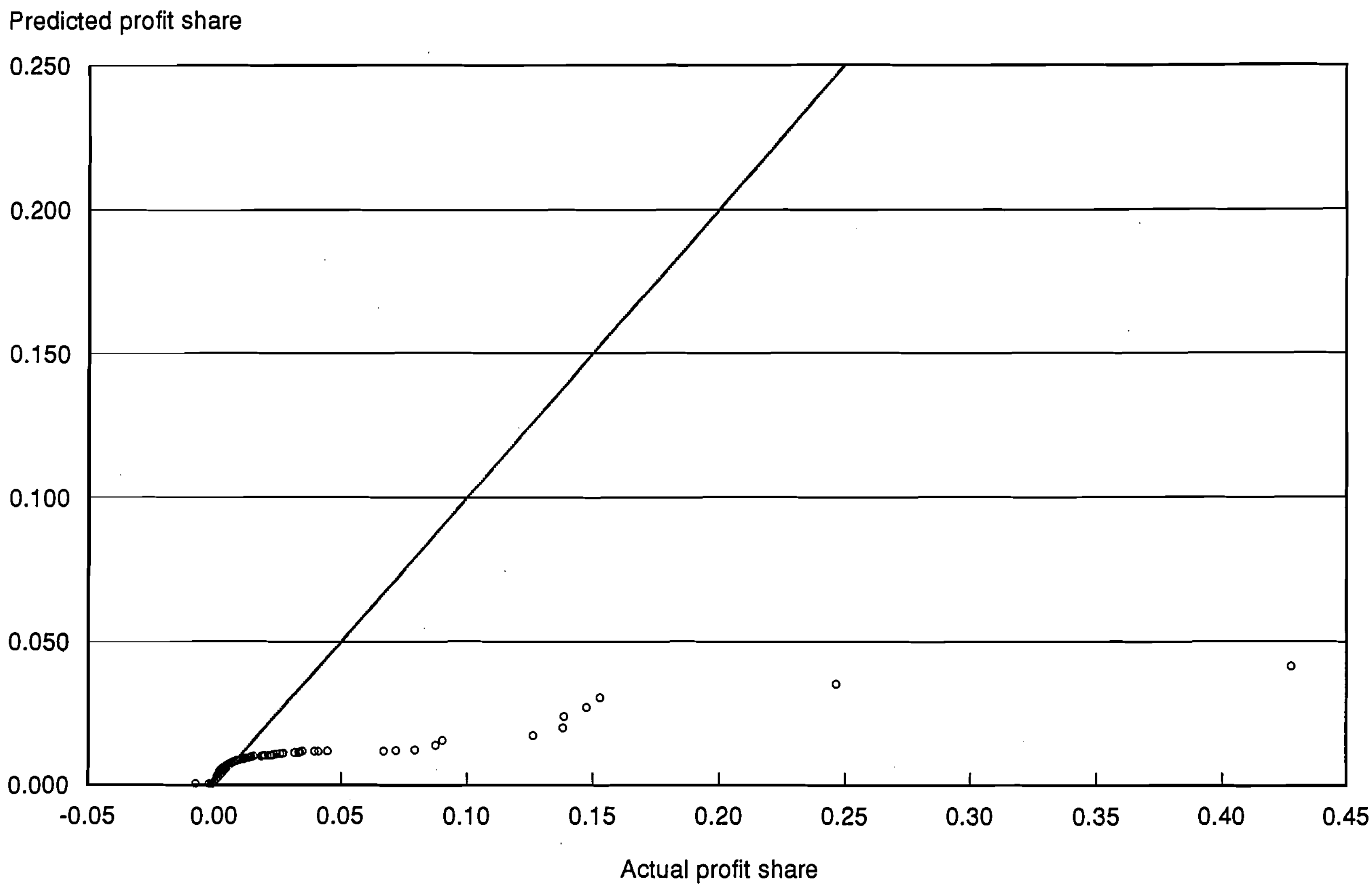
Source: Authors' calculations.

FIGURE 6: PERCENTILE PLOT, ACTUAL AND PREDICTED PROFIT SHARES, 350 FIRMS



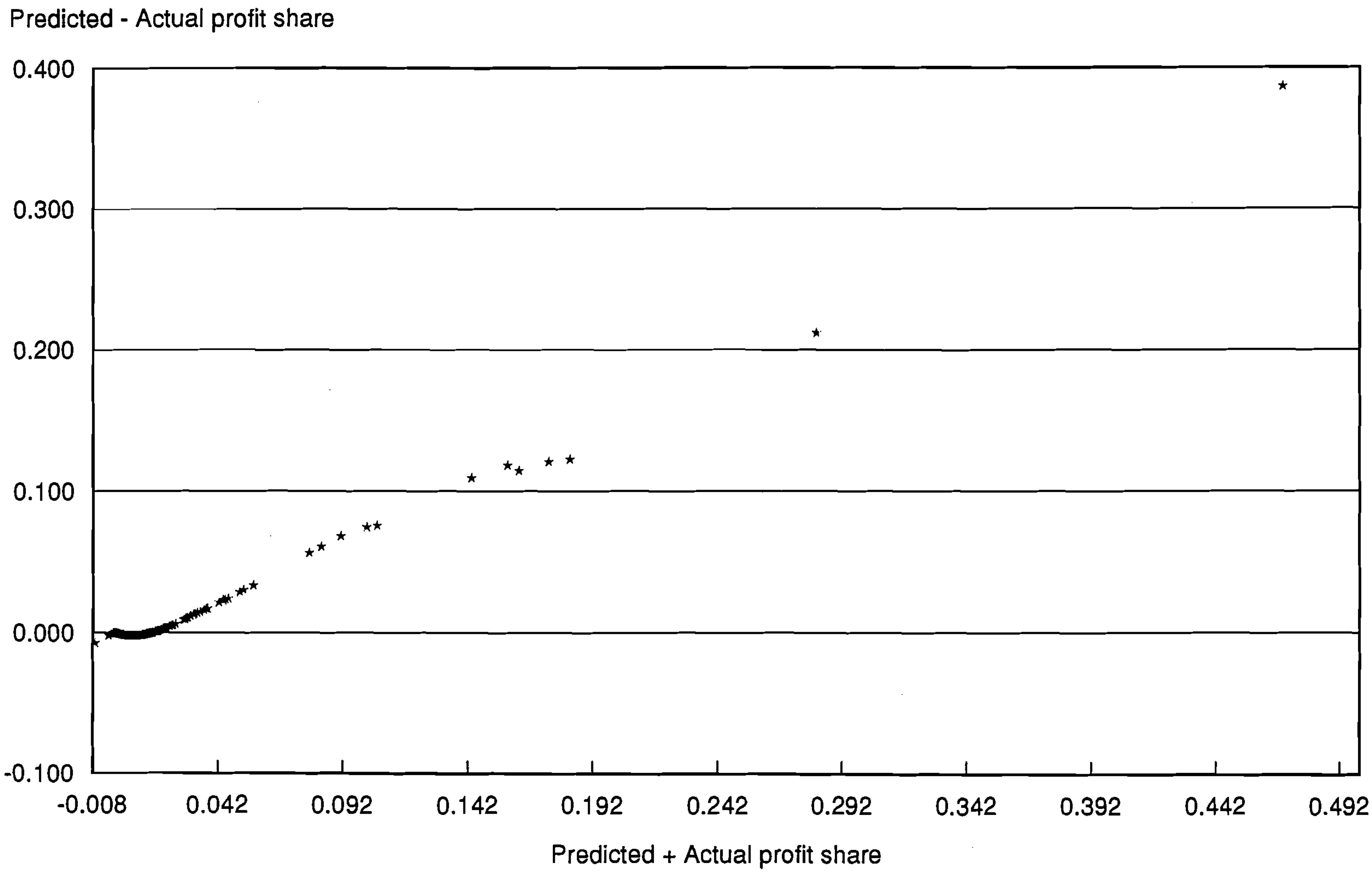
SOURCES: Michael C. Jensen and Kevin J. Murphy, "A New Survey of Executive Compensation" (1990); and authors' calculations.

FIGURE 7: PERCENTILE COMPARISON PLOT, ACTUAL AND PREDICTED PROFIT SHARES, 350 FIRMS



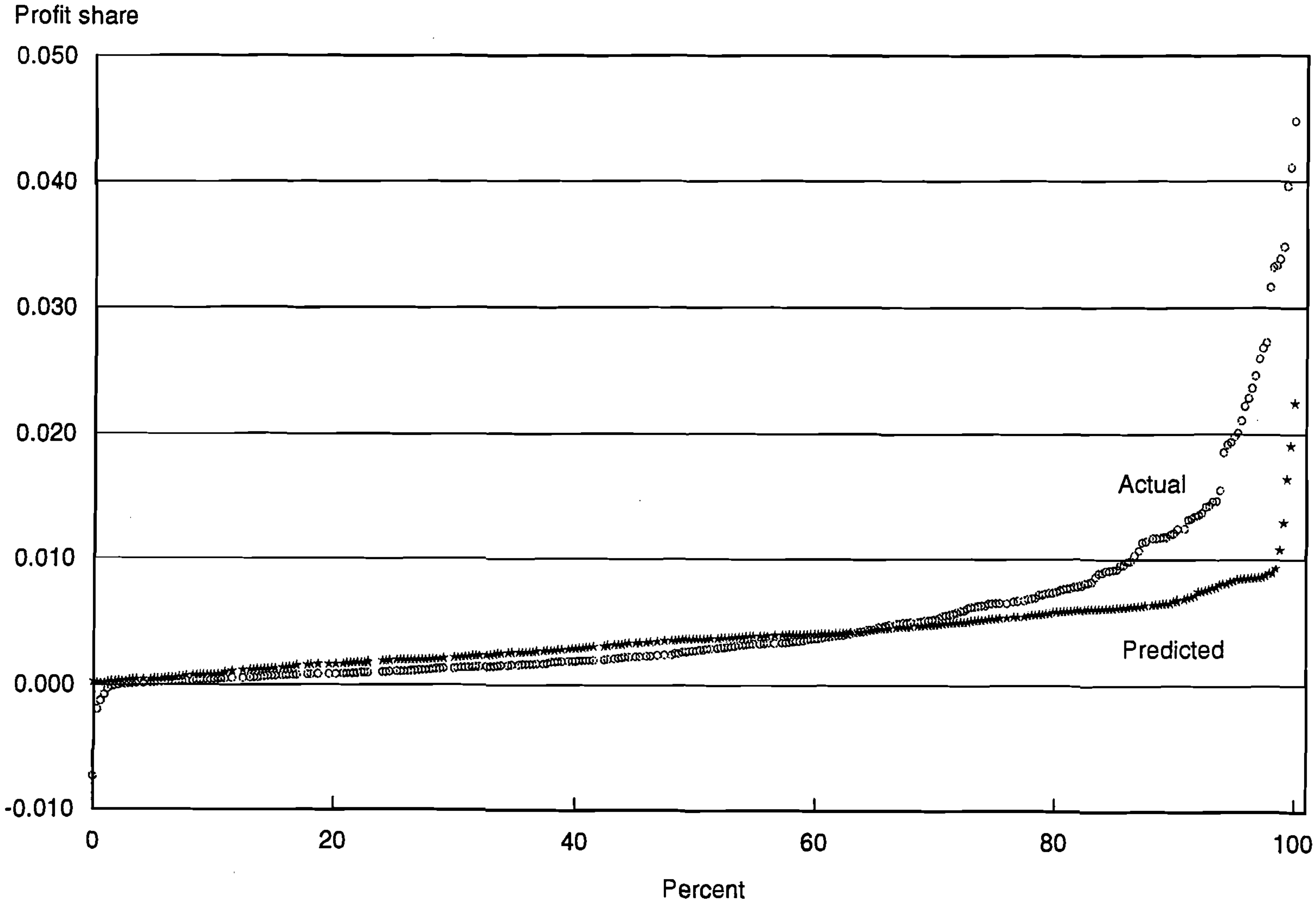
SOURCES: Michael C. Jensen and Kevin J. Murphy, "A New Survey of Executive Compensation" (1990); and authors' calculations.

FIGURE 8: TUKEY SUM-DIFFERENCE GRAPH OF PERCENTILES, 350 FIRMS



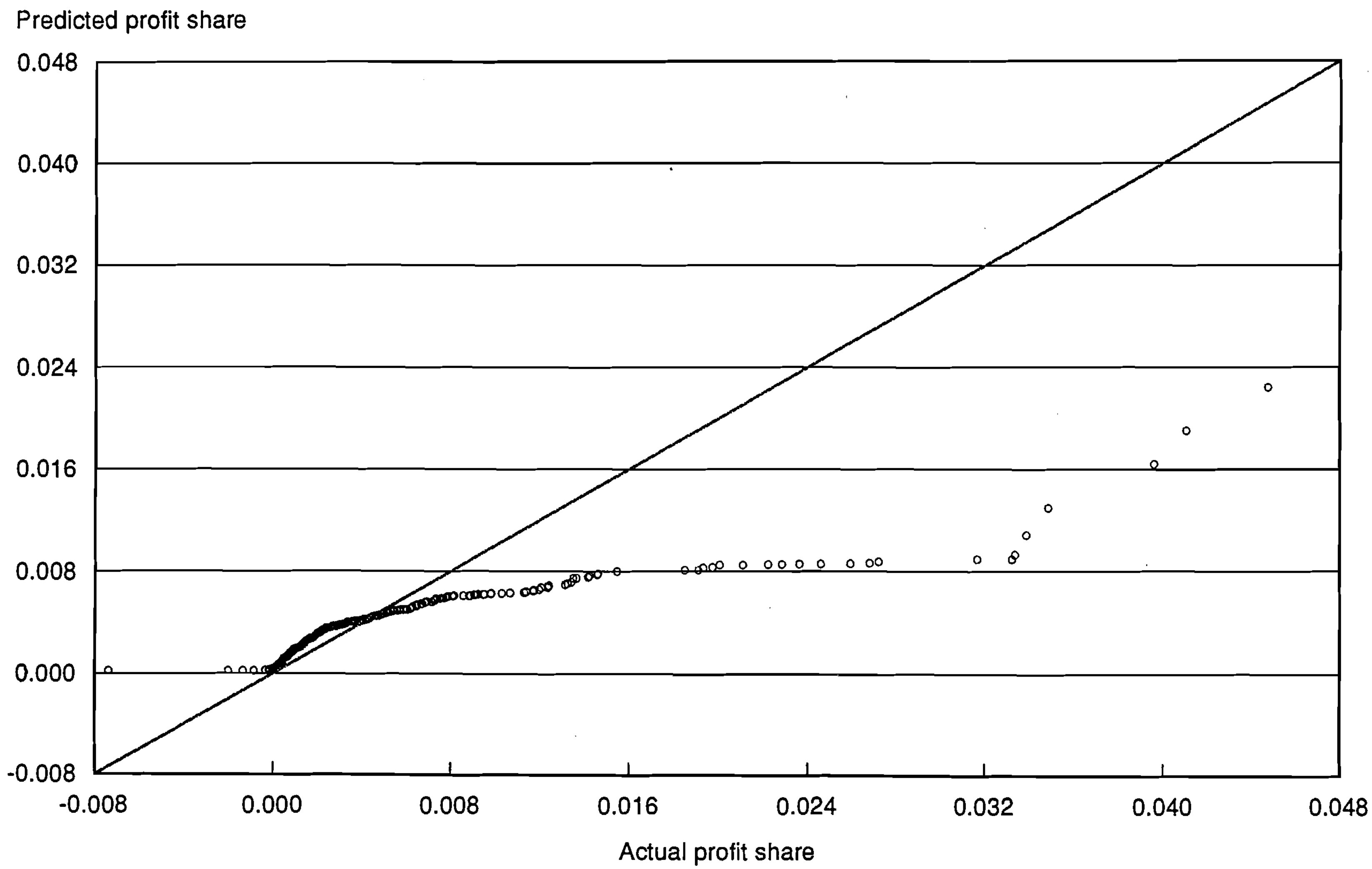
SOURCES: Michael C. Jensen and Kevin J. Murphy, "A New Survey of Executive Compensation" (1990); and authors' calculations.

FIGURE 9: PERCENTILE PLOT, ACTUAL AND PREDICTED PROFIT SHARES, 338 FIRMS



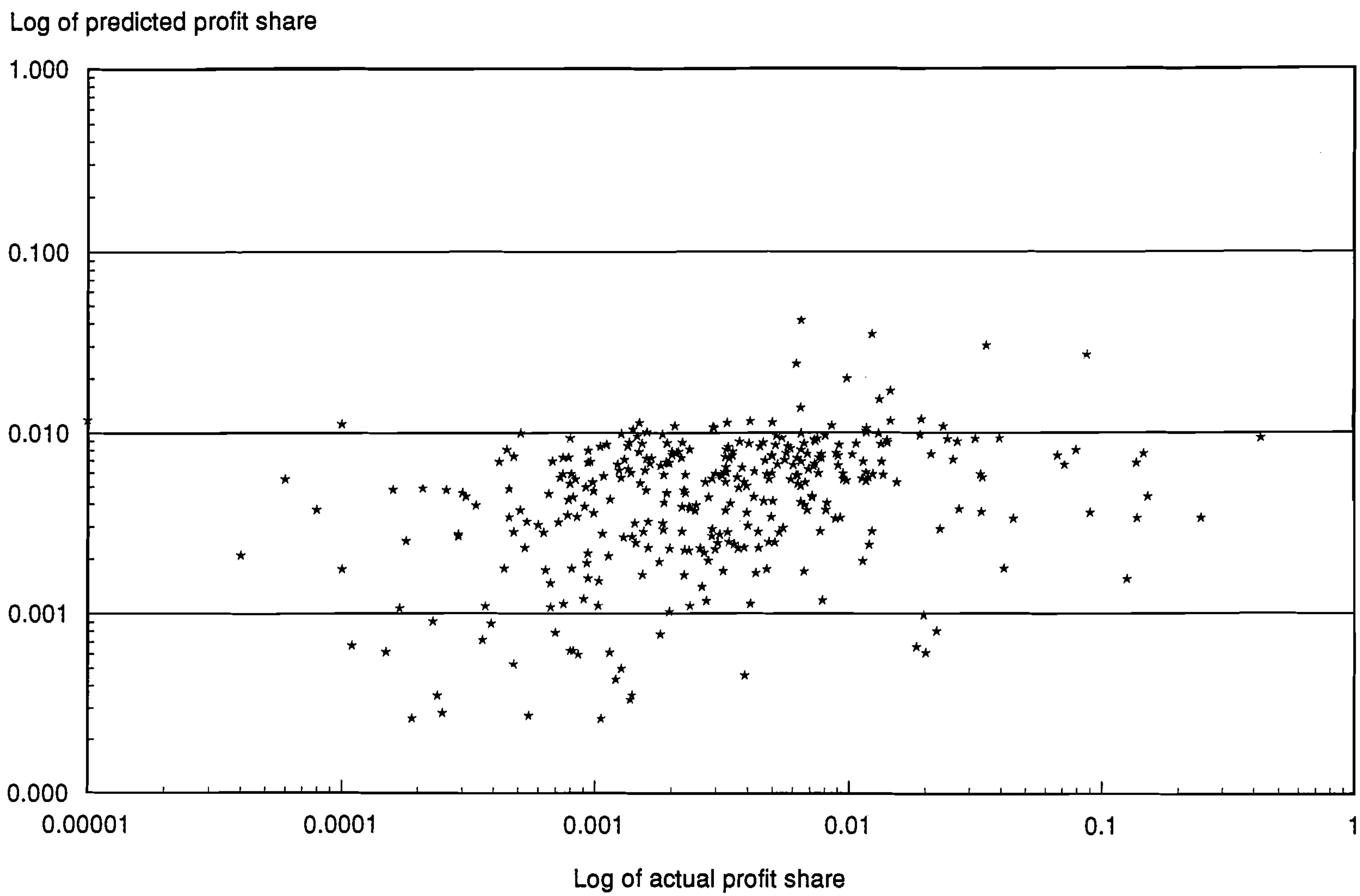
SOURCES: Michael C. Jensen and Kevin J. Murphy, "A New Survey of Executive Compensation" (1990); and authors' calculations.

FIGURE 10: PERCENTILE COMPARISON PLOT, ACTUAL AND PREDICTED PROFIT SHARES, 338 FIRMS



SOURCES: Michael C. Jensen and Kevin J. Murphy, "A New Survey of Executive Compensation" (1990); and authors' calculations.

FIGURE 11: PROFIT SHARE, ACTUAL VS. PREDICTED, 350 FIRMS



SOURCES: Michael C. Jensen and Kevin J. Murphy, "A New Survey of Executive Compensation" (1990); and authors' calculations.

Table 1: MINIMIZING THE LOSS FUNCTIONS

SSE

γ/δ	14	13.5	13	12.5	12	11.5	11	10.5	10	9.5	9
0.001	0.55111	0.68391	0.80316	0.70816	0.65400	0.65583	0.68357	0.76505	0.79702	0.70971	0.77705
0.005	0.43298	0.40798	0.41208	0.40503	0.44856	0.47376	0.44765	0.44017	0.45097	0.46733	0.49516
0.025	0.38553	0.38417	0.37865	0.38635	0.37702	0.37645	0.37248	0.38336	0.38728	0.37681	0.38660
0.125	0.37591	0.37225	0.37209	0.37545	0.37659	0.37517	0.37583	0.37484	0.37002	0.37185	0.37505
0.225	0.38400	0.38470	0.38092	0.37635	0.37672	0.37861	0.38013	0.37918	0.38095	0.38185	0.37464
0.325	0.38516	0.38580	0.38644	0.38705	0.38765	0.37910	0.38003	0.38193	0.38134	0.38296	0.38393
0.425	0.38594	0.38661	0.38723	0.38782	0.38838	0.38892	0.38943	0.37903	0.38271	0.38288	0.38454
0.525	0.38628	0.38707	0.38775	0.38837	0.38893	0.38946	0.38994	0.39040	0.39083	0.38262	0.38387
0.625	0.38580	0.38703	0.38793	0.38866	0.38928	0.38984	0.39033	0.39079	0.39120	0.39159	0.39195
0.725	0.39535	0.38531	0.38749	0.38858	0.38938	0.39003	0.39058	0.39106	0.39148	0.39187	0.39221
0.825	0.39547	0.39559	0.39571	0.38761	0.38906	0.38996	0.39064	0.39119	0.39166	0.39206	0.39242
0.925	0.39559	0.39571	0.39581	0.39591	0.38719	0.38935	0.39042	0.39114	0.39170	0.39215	0.39254

BAR

γ/δ	14	13.5	13	12.5	12	11.5	11	10.5	10	9.5	9
0.001	0.01880	0.02211	0.02387	0.01999	0.01796	0.01889	0.02230	0.02494	0.02397	0.02262	0.02592
0.005	0.00453	0.00469	0.00586	0.00773	0.00981	0.00979	0.00792	0.00834	0.00945	0.01109	0.01145
0.025	0.00078	0.00001	0.00048	0.00054	0.00038	0.00030	0.00088	0.00170	0.00142	0.00123	0.00190
0.125	0.00524	0.00497	0.00455	0.00443	0.00477	0.00511	0.00522	0.00494	0.00439	0.00408	0.00415
0.225	0.00746	0.00761	0.00671	0.00617	0.00568	0.00594	0.00624	0.00642	0.00666	0.00693	0.00580
0.325	0.00771	0.00784	0.00796	0.00807	0.00818	0.00656	0.00641	0.00679	0.00700	0.00722	0.00744
0.425	0.00786	0.00799	0.00810	0.00821	0.00831	0.00840	0.00848	0.00662	0.00705	0.00734	0.00758
0.525	0.00793	0.00807	0.00820	0.00830	0.00840	0.00848	0.00856	0.00863	0.00870	0.00703	0.00755
0.625	0.00784	0.00807	0.00823	0.00835	0.00846	0.00854	0.00862	0.00869	0.00876	0.00881	0.00887
0.725	0.00934	0.00784	0.00815	0.00834	0.00847	0.00858	0.00866	0.00873	0.00880	0.00886	0.00891
0.825	0.00936	0.00937	0.00939	0.00817	0.00842	0.00857	0.00867	0.00876	0.00882	0.00888	0.00894
0.925	0.00937	0.00939	0.00940	0.00941	0.00810	0.00847	0.00864	0.00875	0.00883	0.00890	0.00895

DSSQ

γ/δ	6	5.5	5	4.5	4	3.5	3	2.5	2	1.5	1
0.025	0.00289	0.00413	0.00390	0.00426	0.00293	0.00011	0.00279	0.00477	0.00613	0.00696	0.00724
0.125	0.00433	0.00429	0.00409	0.00382	0.00351	0.00371	0.00487	0.00592	0.00675	0.00728	0.00735
0.225	0.00678	0.00606	0.00628	0.00581	0.00627	0.00566	0.00543	0.00656	0.00722	0.00757	0.00746
0.325	0.00690	0.00748	0.00788	0.00693	0.00762	0.00703	0.00628	0.00610	0.00726	0.00764	0.00732
0.425	0.00854	0.00878	0.00784	0.00832	0.00704	0.00805	0.00725	0.00805	0.00756	0.00777	0.00699
0.525	0.00846	0.00881	0.00903	0.00797	0.00857	0.00886	0.00819	0.00857	0.00864	0.00832	0.00781
0.625	0.00954	0.00858	0.00899	0.00919	0.00932	0.00865	0.00893	0.00901	0.00890	0.00877	0.01101
0.725	0.00957	0.00965	0.00859	0.00907	0.00928	0.00938	0.00941	0.00936	0.00912	0.00911	0.01101
0.825	0.00955	0.00965	0.00972	0.00976	0.00895	0.00921	0.00926	0.00911	0.00930	0.00903	0.01101
0.925	0.00944	0.00960	0.00969	0.00975	0.00977	0.00976	0.00969	0.00951	0.00946	0.01101	0.01101

Source: Authors' calculations.

Table 2: RESULTS FOR 338 FIRMS

SSE										
γ/δ	14.5	14	13.5	13	12.5	12	11.5	11	10.5	10
0.025	0.04331	0.04694	0.04186	0.03799	0.04235	0.03700	0.03772	0.03901	0.05058	0.04985
0.125	0.01787	0.01796	0.01734	0.01869	0.02026	0.01978	0.01957	0.01905	0.01944	0.01863
0.225	0.01888	0.01897	0.01909	0.01776	0.01703	0.01798	0.01807	0.01836	0.01824	0.01891
0.325	0.01905	0.01919	0.01934	0.01951	0.01969	0.01987	0.01744	0.01803	0.01830	0.01823
0.425	0.01920	0.01938	0.01956	0.01974	0.01993	0.02011	0.02030	0.02048	0.01776	0.01823
0.525	0.01923	0.01947	0.01969	0.01990	0.02011	0.02030	0.02049	0.02068	0.02085	0.02102
0.625	0.01901	0.01934	0.01968	0.01996	0.02021	0.02043	0.02063	0.02082	0.02100	0.02117
0.725	0.02294	0.02300	0.01901	0.01982	0.02018	0.02047	0.02071	0.02092	0.02111	0.02129
0.825	0.02300	0.02306	0.02312	0.02317	0.01986	0.02035	0.02068	0.02095	0.02117	0.02136
0.925	0.02306	0.02312	0.02317	0.02322	0.02327	0.01973	0.02045	0.02086	0.02115	0.02138
BAR										
γ/δ	14.5	14	13.5	13	12.5	12	11.5	11	10.5	10
0.025	0.00561	0.00565	0.00490	0.00428	0.00427	0.00441	0.00509	0.00571	0.00652	0.00629
0.125	0.00114	0.00039	0.00010	0.00030	0.00044	0.00010	0.00029	0.00039	0.00010	0.00044
0.225	0.00244	0.00260	0.00275	0.00184	0.00130	0.00081	0.00106	0.00136	0.00158	0.00181
0.325	0.00270	0.00284	0.00297	0.00309	0.00320	0.00331	0.00169	0.00153	0.00191	0.00214
0.425	0.00286	0.00300	0.00312	0.00323	0.00334	0.00343	0.00352	0.00360	0.00176	0.00216
0.525	0.00288	0.00306	0.00321	0.00333	0.00343	0.00352	0.00361	0.00369	0.00376	0.00382
0.625	0.00266	0.00297	0.00320	0.00336	0.00348	0.00358	0.00367	0.00375	0.00382	0.00388
0.725	0.00444	0.00445	0.00298	0.00328	0.00347	0.00360	0.00370	0.00378	0.00386	0.00392
0.825	0.00445	0.00447	0.00449	0.00450	0.00330	0.00354	0.00369	0.00379	0.00388	0.00395
0.925	0.00447	0.00449	0.00450	0.00451	0.00453	0.00323	0.00359	0.00376	0.00387	0.00395
DSSQ										
γ/δ	14.5	14	13.5	13	12.5	12	11.5	11	10.5	10
0.025	0.00555	0.00561	0.00487	0.00426	0.00426	0.00438	0.00503	0.00564	0.00647	0.00624
0.125	0.00117	0.00042	0.00014	0.00027	0.00041	0.00008	0.00031	0.00041	0.00013	0.00040
0.225	0.00247	0.00263	0.00278	0.00187	0.00133	0.00084	0.00109	0.00138	0.00160	0.00183
0.325	0.00273	0.00287	0.00300	0.00312	0.00323	0.00334	0.00172	0.00156	0.00194	0.00217
0.425	0.00289	0.00303	0.00315	0.00327	0.00337	0.00347	0.00356	0.00364	0.00179	0.00220
0.525	0.00291	0.00309	0.00324	0.00336	0.00346	0.00356	0.00364	0.00372	0.00379	0.00386
0.625	0.00270	0.00300	0.00323	0.00339	0.00351	0.00362	0.00370	0.00378	0.00385	0.00391
0.725	0.00447	0.00449	0.00301	0.00331	0.00350	0.00363	0.00373	0.00382	0.00389	0.00395
0.825	0.00449	0.00451	0.00452	0.00454	0.00333	0.00358	0.00372	0.00383	0.00391	0.00398
0.925	0.00451	0.00452	0.00454	0.00455	0.00456	0.00326	0.00363	0.00379	0.00390	0.00399

Source: Authors' calculations.

Table 3 : EFFECT OF INPUT VALUE UNCERTAINTY ON LOSS FUNCTIONS

		SSE	BAR	DSSQ
Optimal results from the model		0.370024	0.004395	0.005366
	Average	0.382051	0.006860	0.007824
Simulation results	StDev	0.003645	0.000149	0.000151
100 draws	High	0.389342	0.007248	0.008215
	Low	0.371578	0.006441	0.007396
Optimal results from the model		0.384171	0.000013	0.000962
	Average	0.381940	0.004497	0.005443
Simulation results	StDev	0.006583	0.000261	0.000265
100 draws	High	0.395234	0.005066	0.006026
	Low	0.364810	0.003571	0.004498
Optimal results from the model		0.398710	0.000810	0.000110
	Average	0.390155	0.004549	0.005464
Simulation results	StDev	0.007941	0.000302	0.000312
100 draws	High	0.411602	0.005087	0.006011
	Low	0.367236	0.003725	0.004620

Source: Authors' calculations.