Online Appendix

Appendix A. Treatment Regression Method and Alternative Linear 2SLS Approach

In this appendix, we discuss two identification approaches we adopt in this paper, to address the potential endogeneity of the $Hedging$ indicators. The first one is the treatment regression maximum likelihood estimator; the second one is two-stage least squares.

The treatment regression used in the paper approaches the problem of estimating a regression equation, when one of the regressors is an indicator variable, and potentially endogenous. In the context of our paper, the regression of interest is:

$$Uncertainty_i = \alpha + \beta Hedging_i + \gamma' x_i + \varepsilon_i \qquad \text{(A.1)}$$

Our hedging indicators are assumed to stem from a latent variable H^*, the actual hedging amount, unobservable to the econometrician, and following:

$$H_i^* = \rho + \mu Convexity_i + \nu' x_i + \eta_i \qquad \text{(A.2)}$$

The treatment $Hedging$ is determined according to:

$$Hedging_i = \begin{cases} 1 & \text{if} \quad H_i^* > 0 \\ 0 & \text{if} \quad H_i^* \leq 0 \end{cases} \qquad \text{(A.3)}$$

Under the assumption that (ε_i, η_i) are jointly normally distributed, Maddala (1983, pp. 117-122) derives a maximum likelihood estimator that estimates jointly (A.1) and (A.2), which we implement in the paper.

The treatment regression method can achieve more efficient estimates, at the potential cost of making an explicit distributional assumption. An alternative which does not rely on distributional assumptions, but is, as a result, a less efficient estimation method, can be the conventional two-stage least squares method. Given that the potentially endogenous variable $Hedging$ is an indicator variable, this method also needs to be treated carefully, as we discuss below.

Two-stage least squares (2SLS)

The potentially endogenous variable in (A.1), $Hedging$, is a binary variable. As is well known, in this situation one cannot simply run a probit regression in the first stage – in fact, a nonlinear first stage is a case of "forbidden regression."

Following the suggestion of Angrist and Pischke (2009, p. 191), however, it is possible to proceed as follows. First, estimate a probit model:

$$\Pr(Hedging_i = 1) = \Phi(Convexity_i, x_i) \qquad \text{(A.4)}$$

where $\Phi(\cdot)$ denotes the normal cdf. From the above model, obtain predicted probabilities $\hat{\pi}$. Next, use the predicted probabilities as excluded instruments for the $Hedging$ indicators, in a conventional 2SLS procedure. That is, in a first stage estimate:

$$Hedging_i = \rho + \mu\hat{\pi}_i + \nu' x_i + \eta_i \tag{A.5}$$

Finally, obtain predicted values $\widehat{H}$, and plug them into the second stage regression:

$$Uncertainty_i = \alpha + \beta\widehat{H}_i + \gamma' x_i + \varepsilon_i \tag{A.6}$$

as in the standard 2SLS approach. This approach can also address cases in which there is more than one endogenous regressor, as the treatment regression method is designed to handle only one endogenous regressor. We resort to it, therefore, in the tests on the market reaction to earnings announcements reported in Table IX.

In addition, this approach provides two natural checks for the strength of the identification. First, we can examine the strength of $Convexity$ as a determinant of $Hedging$. The t-statistics on $Convexity$ reported in Tables II, reporting the estimates of (A.4), are all above 3, indicating that there is a highly significant relationship between $Convexity$ and the decision to hedge corporate risk. In addition, $Convexity$ is an economically important predictor of the hedging decision: one-standard deviation increase in $Convexity$ is associated with a 1-3 percentage points increase in the propensity to hedge, corresponding to a 6-7% increase relative to the baseline.

Second, one can also examine the first-stage F test statistic on $\hat{\pi}$ in (A.5). Across all specifications summarized in Table A.1, the is in nearly all cases well above the conventional value of 10 that suggests no concern about weak instruments (Staiger and Stock (1997)).

Table A.1 Alternative Estimates with Two-Stage Least Squares

The table reports the estimates of regressions of the main specifications, estimated using two-stage least squares. In each regression specification, the hedging indicator (IR/FX, FX, or IR) is instrumented using the predicted values from the probit regressions of Table II. For brevity, the table only reports the coefficient on the hedging indicator, the associated t-statistic, and the first-stage F test. Each specification is otherwise identical to the ones reported in the corresponding tables. The t-statistics and first-stage F tests are based on standard errors clustered around size and book-to-market portfolios in a given year. The symbols *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	IR/FX		**FX**		**IR**	
Dependent variable	Coeff. / t-stat	1st stage F	Coeff. / t-stat	1st stage F	Coeff. / t-stat	1st stage F
Implied Volatility	-0.344***	92.19	-0.101***	43.45	-0.384***	90.40
	-7.72		*-2.85*		*-6.94*	
Analyst Forecast Disp.	-0.003	121.44	-0.002	49.09	-0.007**	91.49
	-0.93		*-0.52*		*-2.04*	
Breadth of Own.	0.943***	153.81	1.117***	83.01	1.437***	158.73
	6.29		*4.97*		*7.99*	
Breadth of Own. - All Inv.	2.690***	155.46	2.078***	86.82	4.493***	164.33
	6.42		*3.74*		*9.61*	
PIN	-3.968***	111.9	-0.219	60.11	-6.895***	111.36
	-3.17		*-0.16*		*-4.73*	
Roll	-0.277***	162.05	-0.173**	87.88	-0.215***	166.61
	-4.7		*-2.53*		*-3.29*	
Short Selling	-0.033***	31.60	-0.069	1.47	-0.037**	21.41
	-2.78		*-1.24*		*-2.16*	
Utilization Ratio	-0.257***	31.60	-0.44	1.47	-0.323***	21.41
	-3.68		*-1.19*		*-3.11*	

Appendix B. Corporate Hedging and Informed Trading

This Appendix illustrates the intuition at the basis of our tests, with the aid of a simplified version of the Kyle (1985) model that incorporates corporate hedging and taxation. A convex tax schedule creates an incentive for the firm to hedge. At the same time, hedging reduces the uncertainty around firm value, and thus the ability of informed traders (short sellers) to trade profitably. Thus, in the presence of corporate hedging we should expect less short selling, and less negative stock returns conditional on short selling.

Timing, cash flows, and corporate taxation

There are three dates (0, 1, 2), a firm, a market maker, and investors, who are universally risk neutral and have discount rate equal to 0. At $t = 2$, the firm generates before-tax cash flows $X \in \{X_L, X_H\}$, $X_H > X_L$, with $\Pr(X = X_H) = \vartheta$.

The corporate income tax schedule is given by:

$$T(X) = \max\{0, \tau(X - \bar{X})\} \tag{B.1}$$

with $X_L < \bar{X} < X_H$. This is the simplest convex tax schedule. Smith and Stulz (1985), Graham and Smith (1999), and Campello, Lin, Ma, and Zou (2011) argue that tax convexity favors hedging – the same is true here, as becomes clear below.

At $t = 0$, the expected after-tax terminal cash flow X' is:

$$E_0(X') = X_L + \vartheta \Delta_X \tag{B.2}$$

where $\Delta_X \equiv (1 - \tau)(X_H - \bar{X}) + \tau\bar{X} - X_L$ is the difference in the *after-tax* terminal cash flows in the two possible states of the world at $t = 2$.

Investors, stock prices, and expected returns

There are two classes of investors:

(i) Informed trader: At a cost $k > 0$, she can observe X at $t = 1$. She then conditions her trade on this information, making an order d_S equal to $-d$ if $X = X_L$, and $+d$ if $X = X_H$. The informed trader's cost k is drawn from a uniform distribution over $[0, k_{\max}]$, independent of all other random quantities in the model.

(ii) Liquidity traders: For exogenous reasons, they submit an aggregate order flow $d_L \in \{-d, +d\}$, with $\Pr(d_L = +d) = \frac{1}{2}$, independent of all other random quantities in the model.

We proceed as follows. First, we obtain market prices at $t = 1$. Second, we compute the ex-ante expected profits to the informed trader, denoted by π. Third, we find the market price at $t = 0$, which, given universal risk neutrality, must be $E_0(X') - \pi$.

Stock price at $t = 1$. Assume for now that $\pi > k$ and the informed trader trades. As in Kyle (1985), the market maker can observe the aggregate order flow $\{d_S, d_L\}$, but *cannot observe the origin of either component of the order flow* – i.e. she does not know which order is from the informed trader (but she can still infer information about the realization of X). The market maker sets the price as $P_1 = E_1(X'|\{d_S, d_L\})$. Depending on the order flow $\{d_S, d_L\}$, there are three possible cases:

(A) $\{+d, +d\}$. This can only happen if $d_S = +d$, i.e. the order flow reveals that $X = X_H$. The probability of this outcome is: $\vartheta/2$. The market maker sets $P_1 = (1-\tau)X_H + \tau\bar{X}$.

(B) $\{-d, -d\}$. This can only happen if $d_S = -d$, i.e. the order flow reveals that $X = X_L$. The probability of this outcome is: $(1-\vartheta)/2$. The market maker sets $P_1 = X_L$.

(C) $\{-d, +d\}$. Here the order flow reveals no information to the market maker: it could be $\{d_S = -d, d_L = +d\}$ or $\{d_S = +d, d_L = -d\}$. These outcomes occur with probabilities $(1-\vartheta)/2$ and $\vartheta/2$. In either case, the market maker sets $P_1 = X_L + \vartheta\Delta_X$.

Profits of the informed trader. The informed trader profits from deviations between the market price and the actual terminal after-tax cash flows. Her ex-ante expected profit is thus:[1]

$$\pi = E_0|P_1 - X'| = \vartheta(1-\vartheta)\Delta_X \tag{B.3}$$

Thus, the informed trader's profits are a function of the uncertainty around firm value $\vartheta(1-\vartheta)\Delta_X$. The informed trader will only trade if $k < k^* \equiv \pi$.

Stock price at $t = 0$. The stock price at $t = 0$ is: $E_0(X') - \pi = X_L + \vartheta\Delta_X - \vartheta(1-\vartheta)\Delta_X = X_L + \vartheta^2\Delta_X$.

Expected returns conditional on short selling. Suppose that the informed trader sells at $t = 1$, i.e. there is an amount of short selling $d_S = -d$. Then, the expected return between $t = 1$ and $t = 2$ is:

$$E(P_2 - P_1|d_S = -d) = -\frac{1}{2}\vartheta(1-\vartheta)\Delta_X = -\frac{1}{2}\pi \tag{B.4}$$

The greater the insider's profits, the more negative the stock return conditional on short selling.

Corporate hedging and short selling

Enter corporate hedging. While the literature suggests a number of alternative mechanisms through which corporate hedging can reduce uncertainty (cf. DeMarzo and Duffie, 1991, 1995), we abstract from any one particular mechanism and simply assume that hedging acts as a mean-preserving "shrinkage" of the distribution of the $t = 2$ cash flows. In other words, if the firm hedges, the terminal cash flows will be $X \in \{X_L + \varepsilon, X_H - \varepsilon\}$, $\Pr(X = X_H - \varepsilon) = \vartheta$, with $X_H - \varepsilon < \bar{X}$. Thus, it can pay no taxes, with the

[1] In cases (A) and (B), the price fully reveals the informed trader's information, and she makes no profit. In case (C), if $d_S = -d$ the informed trader's expected profit is $\frac{1}{2}(1-\vartheta) \times (X_L + \vartheta\Delta_X - X_L) = \frac{1}{2}\vartheta(1-\vartheta)\Delta_X$. The same expression obtains in case (C) if $d_S = +d$, yielding expression (B.3).

same expected before-tax terminal cash flow $(1-\vartheta)X_L+\vartheta X_H$. The expected after-tax cash flows are also $(1-\vartheta)X_L+\vartheta X_H$, while with no hedging we have $X_L+\vartheta[(1-\tau)(X_H-\bar{X})+\tau\bar{X}-X_L]$. Under these assumptions, the firm strictly prefers to hedge. This captures the incentives to hedge generated by the convex tax schedule (the larger τ, the greater the convexity of the tax schedule, the more the firm will want to hedge).[2]

What does this imply for the informed trader's profits? They will be $\hat{\pi}=\vartheta(1-\vartheta)\widehat{\Delta}_X$, with $\widehat{\Delta}_X\equiv X_H-X_L-2\varepsilon$. Recall that under no hedging, the informed trader's profits are given by $\pi=\vartheta(1-\vartheta)\Delta_X$. Now, $\widehat{\Delta}_X<\Delta_X$ iff $\varepsilon>\frac{\tau}{2}(X_H-\bar{X})$. But $\varepsilon>X_H-\bar{X}>\frac{\tau}{2}(X_H-\bar{X})$, so that this condition is always satisfied. To sum up, in the presence of corporate hedging the informed trader will trade only if $k<\hat{k}^*\equiv\hat{\pi}<\pi=k^*$, i.e. the informed trader is less likely to trade, for any given cost of becoming informed k.[3]

If the informed trader does trade, the stock price at $t=0$ will be $X_L+\varepsilon+\vartheta^2\widehat{\Delta}_X$. Moreover, the expected return between $t=1$ and $t=2$, conditional on short selling, will be:

$$-\frac{1}{2}\vartheta(1-\vartheta)\widehat{\Delta}_X \tag{B.5}$$

The above expression is greater (less negative) than the return in (B.4), because $\widehat{\Delta}_X<\Delta_X$. Thus, under corporate hedging, on average the stock price drops less subsequent to a short sale.

The "treatment effect" of corporate hedging on short selling intensity, i.e. the expected drop in short selling given corporate hedging, is given by the drop in the probability that the informed trader trades and the before-tax cash flow realization is X_L. It is thus equal to:

$$\Pr(k\le k^*\cap X=X_L)-\Pr\left(k\le\hat{k}^*\cap X=X_L\right)=(1-\vartheta)\times\frac{k^*-\hat{k}^*}{k_{\max}}>0, \tag{B.6}$$

where the inequality follows from the fact that $k^*>\hat{k}^*$.

[2] For simplicity, we assume no cost associated with corporate hedging. Alternatively, in the presence of hedging costs, the firm would trade them off with the tax benefits of hedging, leading to an optimal hedging. While this would render our argument more complicated, it would not alter our empirical predictions.

[3] An alternative interpretation could be that corporate hedging, by limiting the informed trader's expected profits, lowers the cost of information associated with the "marginal" informed trader. In other words, informed traders that face relatively higher costs of becoming informed will not want to collect private information to trade, as it is no longer profitable in the presence of corporate hedging.