



**OXFORD JOURNALS**  
OXFORD UNIVERSITY PRESS

---

Spin-offs, Divestitures, and Conglomerate Investment

Author(s): Gönül Çolak and Toni M. Whited

Source: *The Review of Financial Studies*, May, 2007, Vol. 20, No. 3 (May, 2007), pp. 557-595

Published by: Oxford University Press. Sponsor: The Society for Financial Studies.

Stable URL: <https://www.jstor.org/stable/4494781>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

and Oxford University Press are collaborating with JSTOR to digitize, preserve and extend access to *The Review of Financial Studies*

# Spin-offs, Divestitures, and Conglomerate Investment

Gönül Çolak

Wichita State University

Toni M. Whited

University of Wisconsin

We examine whether spin-offs or divestitures cause improvements in conglomerate investment efficiency. At issue are endogeneity of these restructuring decisions and correct measurement of investment efficiency. Endogeneity is a problem because the factors that induce firms to spin off or divest divisions may also improve investment efficiency; measurement error is a problem because efficiency measures employ Tobin's  $q$  as a noisy proxy for investment opportunities. We find important differences between firms that divest or spin off and a control sample. After accounting for these differences and for measurement error in  $q$ , we find no evidence of improvements in investment efficiency. (*JEL G31, G34*)

Do changes in the composition of a conglomerate affect the efficiency of the conglomerate's investment expenditures? This question is of interest not only because of its obvious implications for corporate policy but also because of its implications for the large literature on corporate diversification. Recent interest in this area began with Lang and Stulz (1994) and Berger and Ofek (1995), who documented that diversified conglomerates trade at a discount relative to matched portfolios of pure-play firms. As summarized in the survey by Stein (2003), one of the most popular explanations for this "diversification discount" has been inefficient allocation of capital expenditures across corporate divisions. The evidence on this theory has been contradictory: several articles, such as Lamont (1997) and Shin and Stulz (1998), find supportive evidence and several others, such as Whited (2001) and Chevalier (2004), find that this evidence is an artifact of either measurement error in proxies for investment opportunities or endogeneity of the decision to become a conglomerate.

In an attempt to sort out these issues, and as a departure from the literature's focus on cross-sectional variation in firm behavior, a series of

---

We thank the editor, Robert McDonald, an anonymous referee, Tor-Erik Bakke, Joyce Berg, Tim Erickson, Chris Hennessy, N. Prabhala, Leigh Riddick, Drew Roper, Min Shi, and seminar participants at the American University, the Stockholm Institute for Financial Research, the University of Iowa, the University of Toronto, the University of Wisconsin, Virginia Tech, and York University for helpful comments and suggestions. Address correspondence to Toni M. Whited, Department of Finance, University of Wisconsin, 975 University Avenue, Madison, WI 53706-1323, or e-mail: twhited@bus.wisc.edu.

recent articles has turned to an examination of *changes* in the composition of conglomerates. Ahn and Denis (2004), Gertner, Powers, and Scharfstein (2002), Dittmar and Shivdasani (2003), and Burch and Nanda (2003) argued that if paring down a conglomerate improves the efficiency of its investment, then investment inefficiency must be an important source of the diversification discount. In support of this hypothesis, all find that measures of investment efficiency increase after a conglomerate spins-off or divests a segment—actions that we hereafter refer to as refocusing or restructuring.

This article builds upon the idea of examining spin-offs and divestitures as a way to understand conglomerate investment. Its contribution is its attention to two problems that may have contaminated previous findings: endogeneity and measurement error. Endogeneity is an issue because firms choose to refocus. In other words, factors that induce firms to spin off or divest divisions may also improve investment efficiency. Measurement error is an issue because all efficiency measures require proxies for unobservable investment opportunities.

To treat endogeneity, we observe that if refocusing is endogenous, firms that refocus will differ from those that do not. We therefore ask whether observed changes in investment efficiency persist after one controls for these differences. Answering this question requires estimating an average treatment effect. Treatment-effect estimators have been used widely by labor economists to study the effects of training programs on labor market participation and wages. The analogy is apt, because workers who enter training programs choose to do so, just as firms choose to refocus. Therefore, within the context of the effect of refocusing on investment efficiency, “treatment” corresponds to a spin-off or divestiture. If assignment to the treatment is exogenous, *conditional* on a set of observable control variables, then the effect of treatment can be estimated by constructing a matched sample of control firms and then by averaging within-subpopulation differences in investment efficiency that occur between the treatment and control groups. We adopt this approach, using a recent econometric advance in the estimation of average treatment effects—specifically, the matching estimator in Abadie and Imbens (2006). Its main advantage is a correction for the asymptotic bias present in simple matching estimators. This bias can arise if the control and treated groups are insufficiently comparable; that is, if there is incomplete overlap between the distributions of control variables between the treated and the control groups.

To find a set of appropriate control variables, we run a probit regression of the spin-off and divestiture decisions on a set of observable variables. The results indicate that firms that decide to refocus indeed differ from their peers. In particular, they are larger, more diversified, and subject to more serious problems of asymmetric information.

Furthermore, they tend to be in fast-growing industries with a great deal of initial public offering (IPO) and corporate control activity. Finally, they appear to have experienced recent unanticipated shocks to profit.

With the control variables from this first step, we estimate treatment effects, finding that although spin-offs and divestitures are associated with improved investment efficiency, they do not *cause* these improvements. This result is robust to the use of two, more traditional endogeneity remedies: the Heckman (1979) correction for sample selection bias and matching based on an estimated propensity score, as in Dehejia and Wahba (1999, 2002).

Next, we turn to the measurement-error problem, which arises because measures of investment efficiency operate by gauging the reaction of investment to investment opportunities or to liquidity. These approaches, although sensible, require a proxy for unobservable investment opportunities, and the best proxy available, Tobin's  $q$ , is likely to be of low quality. For example, Erickson and Whited (2000) and Whited (2001) find that true unobservable investment opportunities account for only 20–40% of the variation in observed Tobin's  $q$ .

We treat measurement error with the measurement-error consistent estimators in Erickson and Whited (2002) (hereafter EW estimators). Shin and Stulz (1998) have argued that a good measure of conglomerate investment efficiency is the coefficient on other-segment cash flow in a regression of segment investment on industry  $q$ , own-segment cash flow, and other-segment cash flow. Dittmar and Shivdasani (2003) find that this measure rises significantly after firms divest divisions. Both before and after a divestiture, however, the EW estimators produce insignificant coefficients on both own- and other-segment cash flow.

In sum, we find that the observed causation between refocusing and efficiency is an artifact of endogenous selection and measurement error. Our findings are comforting from an economic perspective. As documented in Villalonga (2004a), in the early 1990s at least half of the economic activity occurring in publicly traded firms occurred in diversified conglomerates. If these firms behave inefficiently, it is difficult to understand why competitive forces do not cause more of them to divest or spin off segments. The number of spin-offs and divestitures occurring in any given year is a small fraction of the number of diversified firms. In addition, Maksimovic and Phillips (2000), Bernardo and Chowdry (2002), and Gomes and Livdan (2004) have demonstrated theoretically and empirically that the diversification discount and patterns of internal resource allocation perceived as inefficient can be explained by rational maximizing behavior.

The organization of the article is as follows. In Section 1, we describe the data, and in Section 2, we describe investment efficiency measures. The next two sections contain our tests of whether improvements in these

efficiency measures exist after controlling for endogeneity and measurement error. In Section 3, we treat endogeneity, and in Section 4, we treat measurement error. In Section 5, we present our conclusions. The appendix presents technical details behind several of our calculations, as well as Monte Carlo experiments to assess the finite sample performance of the estimators we use.

## 1. Data

The research design examines whether investment efficiency of the sets of firms that spin off and divest divisions significantly improves after controlling for endogeneity and measurement error. In this section, we discuss the data and then present the different measures of efficiency we examine in the next section. To examine conglomerate investment efficiency, we use three different samples of firms: one sample of firms that spins off divisions, another that divests divisions, and a third control sample that does neither. We describe the construction of each one separately.

### 1.1 Spin-off sample

We use the Security Data Corporation's (SDC) Mergers and Acquisitions Database to identify all spin-offs between 1981 and 1996, for a total of 475 spin-offs. Then, we apply the following sample selection criteria. First, COMPUSTAT firm- and segment-level data must be available for the spin-off parent and child. Second, the spin-off must not be in the financial industry (SIC between 6000 and 6999.) Third, we must be able to confirm that the transaction is a tax-free spin-off by reading the related news report from Lexis-Nexis or Factiva. Our final sample consists of 154 tax-free spin-offs from 1981 to 1996. To facilitate our analysis in the subsequent sections, we also distinguish between the spin-offs undertaken by a diversified parent firm and the ones undertaken by single-segment parent firms. We have 107 spin-offs by multi-segment firms and 47 spin-offs by single-segment firms. Because some of our data analysis requires information on analysts' earnings estimates from I/B/E/S, for this portion of the data analysis, we omit the two observations for which these data are unavailable.

### 1.2 Divestiture sample

Again, we use SDC's Mergers and Acquisitions Database to identify all divestitures from 1983 to 1994. We choose this time period to be consistent with Dittmar and Shivdasani (2003). We initially find 23,930 divestitures. However, not all result in significant structural changes in the parent. To isolate those that do, we use the COMPUSTAT Business Information File to find those that lead to a decrease in the number of

the parent's segments. We find only 1510 such divestitures in the initial sample. Then, we apply the following sample selection criteria. First, we drop divestitures in the financial industries (SIC between 6000 and 6999). Second, we drop firms incorporated outside the United States. Third, the firm must not have divested assets in multiple years over a three-year period. Fourth, we eliminate divestitures associated with a major restructuring, for example, in which the number of divested units with *different* three-digit SICs exceeds the number of segments or in which the firm changes most of the SIC codes of its new segments. Significant "reshuffling" of segment assets in these types of events renders it impossible to identify which divested units belong to which dropped segment. Fifth, because managers can yield certain discretion in classifying and reporting segments, in many cases a divestiture does not necessarily correspond to a dropped segment. To eliminate this possibility, we require that the three-digit SIC codes of dropped segment and divested unit(s) match. In ambiguous instances, we search for related news in Factiva or Lexis-Nexis to determine whether the assets sold belong to a dropped segment.

The final sample consists of 267 divestitures undertaken by 217 firms. We have 104 parent firms (128 corresponding divestitures) that operate with a single segment after the divestiture and 113 parent firms (139 corresponding divestitures) that remain diversified even after they shed assets. As before, we distinguish between these two types of parents. Indeed, to provide a cohesive categorization scheme for both spin-offs and divestitures, we define multi-segment parents as those that have multiple segments both before and after restructuring. We refer to all others as single-segment parents. Finally, as in the case of spin-offs, we omit, when appropriate, the 14 observations for which no I/B/E/S data are available.

### **1.3 Control sample**

To implement matching estimators, we need a set of control firms. We extract this set from the 2002 Standard and Poor's COMPUSTAT industrial files that are also covered by the most recent COMPUSTAT business information file and by I/B/E/S. We select the sample by first deleting any firm-year observations with missing data. Next, we delete any observations for which total assets, the gross capital stock, or sales are either zero or negative. To avoid rounding error issues, we delete firms whose total assets are less than 2 million dollars and whose gross capital stocks are less than 1 million dollars. Furthermore, we delete any observations that fail to obey standard accounting identities. Next, we also discard any potential control firms that are not in the same three-digit industry as any of our divestitures or spin-offs, and we remove any observations that lie outside the time-span of our sample of divestitures and spin-offs. Finally, we include a firm only if it has at least seven consecutive years of complete

data, if it reports more than one segment in all of these years, and if the number of segments remains constant. These selection criteria leave us with 461 firms, and 2812 firm-year observations.

## 2. Measures of Investment Efficiency

We examine multiple measures of investment efficiency to test whether refocusing decisions affect conglomerate investment efficiency. Four separate measures have been proposed in the diversification-discount literature. In this section, we describe these measures and discuss the presence of measurement error in each. Subsequent to this section, we discuss how to control for endogeneity in the types of experiments that have used these measures, as well as how to correct for measurement error. We show how these corrections affect the empirical results.

Our description of these measures is divided into three parts. The first part describes the two measures based on the coefficients from regressions of investment on  $q$  and cash flow. This part also points out measurement error-induced biases in these quantities. The second part describes the two measures based on the correlation between investment and investment opportunities across conglomerate divisions. Because the correlation-based measures are difficult to interpret, the second part also examines their economic content with simulations. The third part describes two quantities that are closely related to the four efficiency measures we consider. These additional two quantities measure the value of the conglomerate relative to the value of a portfolio of pure-play firms in the same industries as the conglomerate's segments. Although not explicit measures of investment efficiency, such "excess value" measures have been studied extensively in the context of spin-offs and divestitures and are clearly related to investment efficiency, inasmuch as capital expenditures contribute to firm value.

### 2.1 Regression-based measures of efficiency

This section describes the two regression-based measures of investment efficiency and explains how they may be contaminated by measurement error in  $q$ . The first of these two measures, introduced by Scharfstein (1998), is the sensitivity of segment investment to the median Tobin's  $q$  of the single-segment firms in that segment's industry. Also used in Gertner, Powers, and Scharfstein (2002) and Ahn and Denis (2004), this measure captures the idea that the more efficient a firm, the more its investment should respond to changes in investment opportunities. Both of these studies find increases in investment- $q$  sensitivity after spin-offs.

However, measurement error in  $q$  can cause changes in this sensitivity, if the error variance is time varying. In this case, the usual downward attenuation bias on the  $q$  coefficient can move even when the true

sensitivity of investment to  $q$  is constant. A time-varying measurement error variance is likely, because the proxy for investment opportunities is the median  $q$  of the industry to which the segment belongs and because this proxy may be better for a single-segment firm than for a segment of a conglomerate. Proxy quality may be poor for a segment, because, as pointed out in Maksimovic and Phillips (2002), if a firm is part of a conglomerate, the productivity of its assets is likely to depend on the assets of the entire conglomerate, and the industry median  $q$  will not be as good a measure of its investment opportunities as it is in the case of an otherwise identical firm that operates independently. Therefore, when a segment becomes a single-segment firm, industry median  $q$  may become a better proxy for investment opportunities, and the attenuation bias in an investment- $q$  regression may be reduced.

The second regression-based efficiency measure is the sensitivity of segment investment to other-segment cash flow. As argued in Shin and Stulz (1998), if a conglomerate's internal capital market is efficient, the cash flow of the firm as a whole should matter more for a segment's investment than that segment's own cash flow. Therefore, in a regression of investment on  $q$ , own-segment cash flow, and other-segment cash flow, other-segment cash flow should carry a larger coefficient than own-segment cash flow. Both Shin and Stulz (1998) and Dittmar and Shivdasani (2003) confirm this hypothesis, and the latter found that the coefficient on other-segment cash flow increases significantly after a divestiture.

This measure of investment efficiency presents a problem related to the one afflicting investment- $q$  sensitivity. As discussed in Whited (2001), it is likely that cash flow sensitivities have appeared significant because typical proxies for segment  $q$  fail to capture investment opportunities. The apparent significance also occurs because investment opportunities are positively correlated with cash flow—not only a segment's own cash flow but other segments' cash flows as well. If the quality of the proxy for true segment  $q$  is time varying, any differences in the sensitivity of investment to cash flow may be because of this time variation in proxy quality.

## 2.2 Correlation-based measures of efficiency

This section describes the next two investment efficiency measures. Introduced by Rajan, Servaes, and Zingales (2000), they are based on the association between investment and investment opportunities across divisions. The stronger the association, the more efficient conglomerate investment becomes. This section describes the measures in both mathematical and intuitive terms. It then goes on to present a simulation that aids in interpreting the magnitudes of these measures.

The first measure is called the relative investment ratio ( $RINV$ ). To construct it, consider a firm with  $n$  segments, assign to each segment the median  $q$  of the single-segment firms operating in the same three-digit

SIC industry, and sort the segments on the basis of these  $q$ 's. Suppose the first  $k$  segments have industry median  $q$ 's greater than the sales-weighted average of all the segments' industry median  $q$ 's. Let  $S_j$  be the sales of segment  $j$ ,  $w_j$  be the fraction of total firm sales made by segment  $j$ ,  $I_j$  be the capital expenditures of segment  $j$ , and  $(I/S)_j^{ss}$  be the capital expenditure-to-sales ratio of the median single-segment firm operating in the same three-digit SIC industry as segment  $j$ . One version of RINV can be expressed as

$$\begin{aligned} RINV_S \equiv & \sum_{j=1}^k w_j \left\{ \frac{I_j}{S_j} - \left( \frac{I}{S} \right)_j^{ss} - \sum_{i=1}^n w_i \left[ \frac{I_i}{S_i} - \left( \frac{I}{S} \right)_i^{ss} \right] \right\} \\ & - \sum_{j=n-k+1}^n w_j \left\{ \frac{I_j}{S_j} - \left( \frac{I}{S} \right)_j^{ss} - \sum_{i=1}^n w_i \left[ \frac{I_i}{S_i} - \left( \frac{I}{S} \right)_i^{ss} \right] \right\}. \end{aligned} \quad (1)$$

To understand this expression, note that

$$\left[ \frac{I_j}{S_j} - \left( \frac{I}{S} \right)_j^{ss} \right] \quad (2)$$

is simply the industry-adjusted investment to sales ratio. Therefore, the term

$$\frac{I_j}{S_j} - \left( \frac{I}{S} \right)_j^{ss} - \sum_{i=1}^n w_i \left[ \frac{I_i}{S_i} - \left( \frac{I}{S} \right)_i^{ss} \right] \quad (3)$$

is the industry- and firm-adjusted investment to sales ratio. Equation (1) then implies that, on an industry- and firm-adjusted basis,  $RINV_S$  will be greater when firms invest more in their high- $q$  segments, that is, when they are more efficient. A second version of  $RINV$ ,  $RINV_A$  is constructed analogously, except that "assets" replaces "sales."

The second measure is called relative value added (RVA). To construct it, let  $q_j$  be the industry median  $q$  of segment  $j$ . The sales-based relative value added measure,  $RVA_S$  is

$$\sum_{j=1}^n w_j (q_j - \bar{q}) \left\{ \frac{I_j}{S_j} - \left( \frac{I}{S} \right)_j^{ss} - \sum_{i=1}^n w_i \left[ \frac{I_i}{S_i} - \left( \frac{I}{S} \right)_i^{ss} \right] \right\}, \quad (4)$$

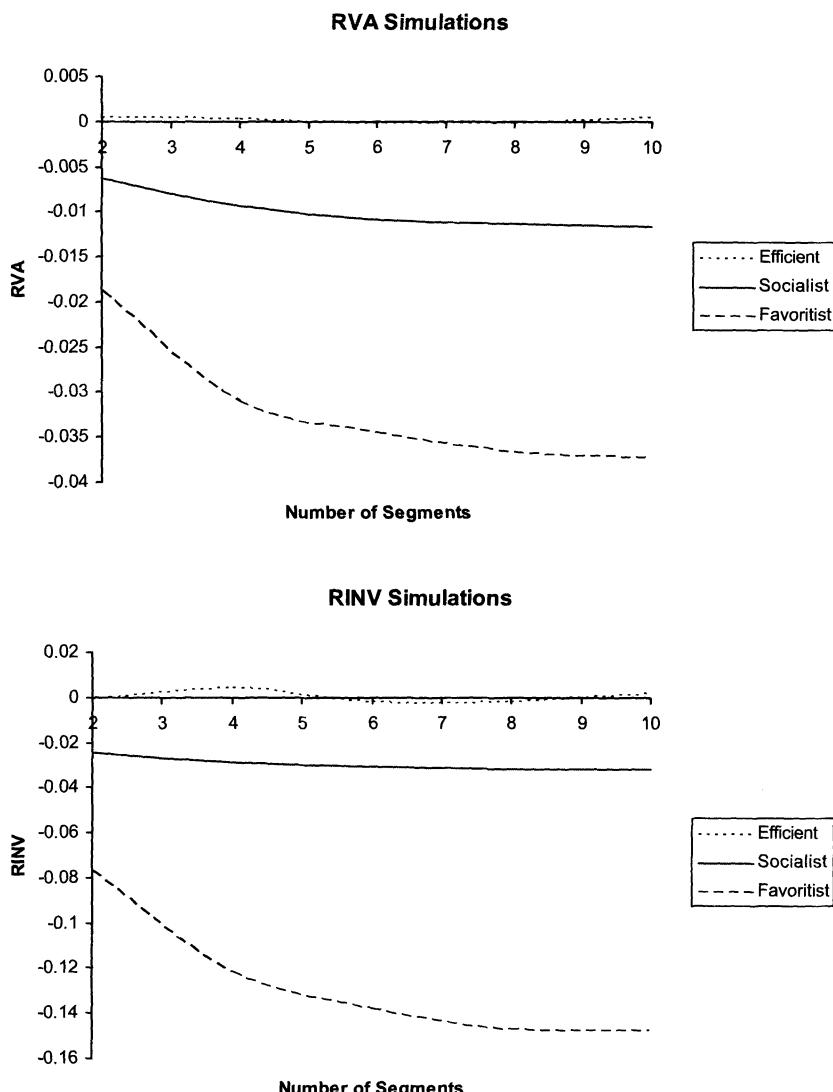
in which  $\bar{q}$  is the sales-weighted average of all of the segment industry median  $q$ 's.

To understand this expression, consider the simple case of all segments having the same sales.  $RVA_S$  is then simply the covariance between industry median  $q$  and industry-adjusted segment investment. Because all segments do not have the same sales,  $RVA_S$  is therefore the sales-weighted covariance between  $q$  and investment. As with  $RINV$ , it is higher when firms allocate more capital expenditures to their segments in high- $q$  industries. As above, an asset-based measure of  $RVA$ ,  $RVA_A$ , can be constructed by substituting assets for sales.

Having defined  $RINV$  and  $RVA$ , we now turn to the problem of interpreting observed magnitudes of these quantities. This problem is difficult, because, unlike regression coefficients, these measures cannot be interpreted in terms of elasticities. Accordingly, we turn to simulations to understand what sorts of values for  $RINV$  and  $RVA$  represent efficient or inefficient investment. The simulations compare observed values of  $RINV$  and  $RVA$  for three types of simulated firms: efficient, “favoritist,” and “socialist.” The efficient firm observes true investment opportunities ( $q$ ) for each division and allocates its capital budget accordingly. The next two types of firms represent different aspects of inefficient behavior. The favoritist firm allocates all of its capital budget to the firm with the worst investment opportunities, and the socialist firm allocates an equal capital budget to each division. It is important to note that we simulate versions of  $RINV$  and  $RVA$  that use imperfectly measured versions of  $q$ . Our intent is to approximate versions of these variables that are observable to the econometrician. We generate sets of 10,000 pseudo-conglomerate, in which each set is defined by the number of segments in each of its member conglomerates and in which we consider conglomerates with two, four, six, eight, and ten segments. The details of the construction of the simulation are in the Appendix.

Figure 1 depicts average values  $RINV$  and  $RVA$  over each set of simulated firms, as a function of the number of segments in a simulated conglomerate. First, efficient firms’ levels of  $RINV$  and  $RVA$  hover around zero, and both types of inefficiencies generate negative values for  $RINV$  and  $RVA$ . Not surprisingly, favoritist investment produces lower levels of both measures. Also, a comparison of the two figures reveals that the same inefficiency generates lower values of  $RINV$  than of  $RVA$ . Finally, because socialist and favoritist investment represent extreme inefficiencies, one would expect real-data estimates of these quantities to be smaller in absolute value than the values in Figure 1.

Figure 2 addresses a related question that helps with interpretation of the empirical results below. It illustrates whether observed values of  $RVA$  and  $RINV$  can rank efficient investment higher than either type of inefficient investment, by plotting the fraction of simulated conglomerates in which observed  $RVA$  and  $RINV$  provide the correct ranking. For both types of inefficiency,  $RVA$  and  $RINV$  only provide the correct ranking 65% of the

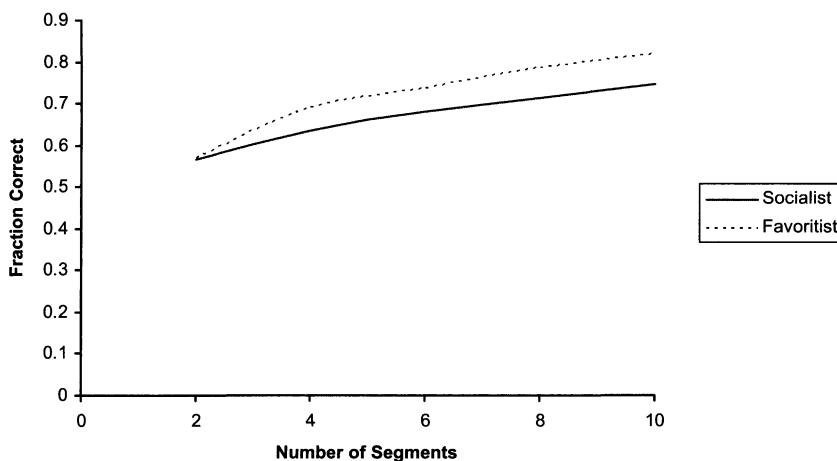


**Figure 1**

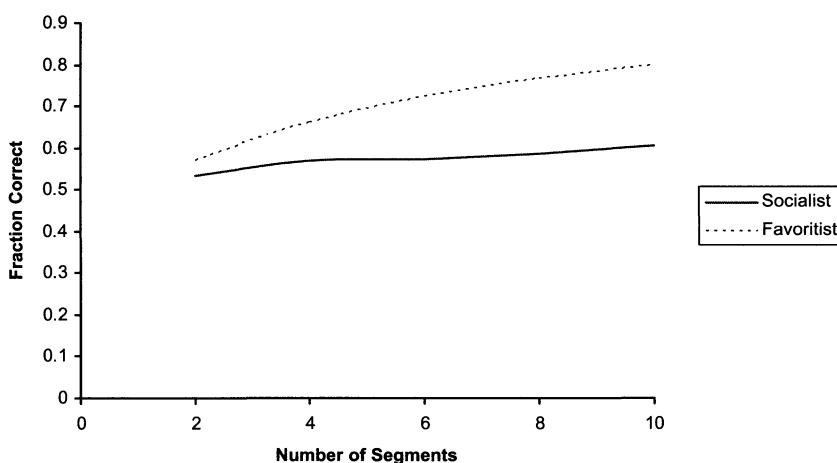
**Simulations of RVA and RINV levels**

RVA is "relative value added," and RINV is "relative investment ratio." Indicated levels are averages based on 10,000 Monte Carlo simulations. Efficient investment refers to firms that invest according to the  $q$  theory of investment. Socialist investment refers to firms that spread their capital budgets equally across segments. Favoritist investment refers to firms that give their entire capital budget to the segment with the lowest marginal  $q$ . The horizontal axis measures the number of segments in the simulated conglomerate.

### RVA Simulations



### RINV Simulations



**Figure 2**

#### Simulations of RVA and RINV orderings

RVA is “relative value added,” and RINV is “relative investment ratio.” Indicated levels are averages based on 10,000 Monte Carlo simulations. Efficient investment refers to firms that invest according to the  $q$  theory of investment. Socialist investment refers to firms that spread their capital budgets equally across segments. Favoritist investment refers to firms that give their entire capital budget to the segment with the lowest marginal  $q$ . The vertical axis reports the fraction of the simulations in which RVA is higher for efficient investment than for socialist or favoritist investment, respectively. The horizontal axis measures the number of segments in the simulated conglomerate.

time on average. This percentage increases with the number of segments—a pattern that provides useful intuition for this poor performance. Both *RVA* and *RINV* are measures of association; indeed, *RINV* is closely related to a covariance. As with any measure of association, the precision of estimation increases with sample size. Because the sample sizes here range from two to ten, it is not surprising that they perform poorly.

### 2.3 Excess value measures

To conclude this section, we describe measures of the value of a conglomerate relative to a portfolio of pure-play firms in the same industries as the conglomerate's segments. To make our study comparable to its predecessors, we use two variants of this variable. When we examine spin-offs, we follow Ahn and Denis (2004), computing a quantity usually referred to as excess value (*XVAL*). The sales-based version can be expressed as

$$\left(\frac{V}{S}\right)_i - \sum_{j=1}^n w_j \left(\frac{V}{S}\right)_j^{ss}, \quad (5)$$

in which  $w_j$  is defined as above,  $(V/S)_j^{ss}$  is the median ratio of market value to sales for the median firm in the three-digit industry to which segment  $j$  belongs, and  $(V/S)_i$  is the ratio of market value to sales for the conglomerate as a whole. When we examine divestitures, we follow Dittmar and Shivdasani (2003), computing a sales-based measure *not* of excess value but of the discount at which the conglomerate trades relative to a comparable portfolio of pure-play firms. We measure the discount as

$$-\ln \left[ \frac{\left(\frac{V}{S}\right)_i}{\sum_{j=1}^n w_j \left(\frac{V}{S}\right)_j^{ss}} \right]. \quad (6)$$

As above, asset-based measures are defined by replacing sales with assets.

## 3. The Endogeneity of Refocusing

Our first goal in re-examining divestitures and spin-offs is to understand whether the endogeneity of these actions has produced spurious causal inference about investment efficiency. As emphasized in the introduction, observed divestitures and spin-offs are not selected at random from the population of diversified firms; that is, the firms that choose these actions have characteristics that systematically differentiate them from their peers that do not refocus. Therefore, it is important to control for these characteristics when ascertaining if refocusing *causes* improvements in

investment efficiency or if refocusing is only associated with these improvements, via its correlation with firm characteristics.

A simple example helps clarify the nature of the endogeneity problem. After its spin-off from AT&T, Lucent's capital expenditures increased dramatically at the same time that the telecommunication industry was booming. One can argue that the spin-off caused this change. One can argue just as easily, however, that if Lucent had remained a segment of AT&T, because it was in a "hot" industry, its investment would still have increased relative to the other segments, thus improving AT&T's investment efficiency. In other words, it is difficult to tell whether Lucent's investment increased because of the spin-off or because of its changing investment opportunities.

We treat this problem in two ways. First, as a "sniff-test" in subsection 3.1, we examine summary statistics from our three samples. Second, and more importantly, we estimate treatment effects. This procedure requires an important preliminary step—the identification of firm and industry attributes to be used as control variables in the estimation of these treatment effects. To this end, we use probit regressions, discussing the specification in subsection 3.2 and the results in subsection 3.3. The next two subsections explain the Abadie and Imbens (2006) estimator (AI hereafter) and present the results. The next subsection presents results from two alternative treatment-effect estimators as a robustness check. The final subsection considers an important alternative interpretation of our results.

### **3.1 Summary statistics**

As a first pass at ascertaining whether endogeneity is a problem, we calculate simple summary statistics from our three samples, examining the spin-off and the divestiture samples *before* the refocusing event. The results are in Table 1, which reveals several striking patterns. First, refocusing firms have much higher values of  $q$  (better investment opportunities) than the control firms. Second, firms that divest and spin off segments are much smaller than firms in the control sample. Third, the divestiture sample, in particular, invests at a higher rate than the control sample. No appreciable differences exist in operating income or debt across the three samples. These results are consistent with the hypothesis that the refocusing firms are systematically different from the control firms and that this difference drives observed improvements in investment efficiency. This evidence, however, is merely suggestive, and we now turn to the task of uncovering more conclusive evidence.

### **3.2 Control variables**

AI emphasize the importance of matching on more than one or two variables, such as firm size and industry. As a first preliminary step in

**Table 1**  
**Summary statistics: spin-offs, divestitures, and controls**

	Assets	<i>Q</i>	Operating income	Debt	Investment	Number of segments
<b>Spin-offs</b>						
Mean	4978.503	1.886	0.140	0.222	0.069	3.051
Median	1205.679	1.410	0.137	0.221	0.061	3.000
<b>Divestitures</b>						
Mean	1840.557	1.734	0.107	0.225	0.081	3.428
Median	303.810	1.099	0.107	0.191	0.062	3.000
<b>Controls</b>						
Mean	5502.314	1.409	0.105	0.235	0.066	3.086
Median	3619.779	1.031	0.114	0.200	0.053	3.000

The sample consists of 107 spin-offs from 1983 to 1997, 139 divestitures from 1984 to 1994, and 461 control firms from 1983 to 1997. Assets are measured in millions of 1997 dollars. The numerator of *Q* contains the book value of assets less the book value of common equity plus the market value of common equity less deferred taxes. The denominator contains the book value of assets. Operating income, debt, and investment are all scaled by assets.

the implementation of the AI estimator, we therefore examine studies of motives for spin-offs and divestitures to find candidate control variables.

First, we examine work that analyzes motives for spin-offs, such as Cusatis, Miles, and Woolridge (1993), Daley, Mehrotra, and Sivakumar (1997), Krishnaswami and Subramaniam (1999), Desai and Jain (1999), and Maxwell and Rao (2003). Krishnaswami and Subramaniam (1999) argue that conglomerates operating in a diverse range of industries benefit most from the focus provided by a spin-off. They measure diversity via the entropy measure from Palepu (1985), which we also adopt. They also argue that firms spin off divisions to improve the divisions' transparency. In this case, the level of asymmetric information should be able to predict a spin-off. We adopt two of their proxies. The first is the standard deviation of analysts' earnings-per-share (EPS) forecasts obtained from I/B/E/S in the last month of the fiscal year before the spin-off. We modify this measure by using the coefficient of variation, which is a unit-free measure. The second is an EPS forecast error: the absolute value of the difference between the EPS estimate in the last month of the fiscal year before the spin-off and the actual EPS during the spin-off year. We normalize this figure by the ratio of shares outstanding to assets, to put this figure in the same units as most of the rest of our control variables. Finally, several authors, such as Parrino (1997), have noted that a possible motive for a spin-off is expropriation of wealth from bondholders to stockholders. Accordingly, we also include the ratio of debt to assets as an explanatory variable.

Next, we turn to studies of divestitures, such as Bergh (1997), Berger and Ofek (1999), Maksimovic and Phillips (2001), Schlingemann, Stulz, and Walkling (2002), and Haynes, Thompson, and Wright (2003). Variables that explain divestitures and spin-offs are similar, not surprisingly,

although some differences exist. Haynes, Thompson, and Wright (2003) argue that large firms and firms with large market shares can benefit more from the focus provided by a divestiture. We follow them in measuring size as the log of total assets and market share as the ratio of firm sales to three-digit industry sales. Next, Lang, Poulsen, and Stulz (1995) find that firms divest assets if they need cash to finance capital expenditures in their core divisions. Similarly, we include the ratio of cash flow to assets, where cash flow is defined as net income plus depreciation. We also include the ratio of the firm's financing gap to assets, where the financing gap is defined as the difference between capital expenditures and the sum of cash flow and net debt issuance. Maksimovic and Phillips (2002) show that firms sell assets after a positive demand shock in the firm's main industry. To capture this idea, we include the operating profits of the firm's largest segment divided by that segment's assets.<sup>1</sup> Finally, the model in Shleifer and Vishny (1992) predicts that asset sales are done by high-debt firms when their industries are temporarily distressed; so the debt-to-assets ratio explains divestitures as well as spin-offs, albeit for a different reason.

Next, we consider four variables designed to answer the question of why the firm decided to refocus in the particular year that it did. First, we include the forecast error from a panel autoregression of the first difference of the ratio of operating income to assets. We use the technique in Holtz-Eakin, Newey, and Rosen (1988) with twice-lagged operating income as an instrument.<sup>2</sup> The forecast error captures an unanticipated shock that may prompt a conglomerate to break apart. We also include industry sales growth in the year of the spin-off or divestiture. Because an industry-level variable is arguably exogenous to the firm, we measure it contemporaneously, thus capturing the idea of an unanticipated change in industry prospects. Next, motivated by the idea in Schlingeman, Stulz, and Walkling (2002) that the liquidity of the market for corporate assets is an important restructuring motive, we include the market value of IPO activity and the value of corporate control activity in the firm's two-digit industry. We normalize both variables by the industry market capitalization. These data are from the SDC database. As in Schlingeman, Stulz, and Walkling (2002), control activity includes all disclosed and completed leveraged buyouts, tender offers, exchange offers, minority stake purchases, acquisitions of remaining interest, privatizations, and equity

<sup>1</sup> For the most part, Maksimovic and Phillips (2001) focus on the characteristics of divested segments as opposed to the characteristics of divesting firms, which is our interest.

<sup>2</sup> Because the autoregression requires a twice-lagged instrument, we must drop three firms from the spin-off sample and five from the divestiture sample. These omissions affect the results little: if we omit the profit shock from the treatment effect estimation, the results with and without these extra observations are almost identical. We include these observations in the investment regressions, because they do not include the shock.

carve-outs. We exclude share repurchases and self-tenders. As in the case of industry sales growth, we measure these industry-level variables in the year of the spin-off or divestiture.

Finally, we add year dummies and Tobin's  $q$ , measured by the market-to-book ratio. Here,  $q$  is not a proxy for investment opportunities; rather, it is a scaled measure of value. We also add the *levels* of RINV, RVA, and excess value. If improvements in investment efficiency or value indeed motivate spin-offs and divestitures, then low levels of these variables should prompt refocusing.<sup>3</sup> Except where noted, we measure all control variables pretreatment, that is, in the year before the spin-off or divestiture.

### 3.3 Probit regressions

As a second preliminary step in choosing an appropriate vector of control variables to use in the AI estimator, we run probit regressions to predict refocusing. The intent of this exercise is to narrow down the long list of possible control variables to those that appear to have an impact on the refocusing decision. The results are in Table 2, which is divided into two sections. In the first, all "ratio" variables have assets in the denominator, and the regressions contain the asset-based measures of RINV, RVA, and XVAL. In the second, all of these variables are normalized instead by sales. Each section presents separate results on spin-offs and divestitures. Because the two sections are similar, we discuss only the first. We report the marginal effects from the probit regressions, with their standard errors in parentheses.<sup>4</sup>

Table 2 shows that several of these marginal effects are significant. Of particular interest are the asymmetric information variables: the coefficient of variation has a significant positive impact on spin-offs, and the forecast error has a significant positive impact on divestitures. Also, the measure of relatedness is important across the board; in other words, refocusing happens more often to highly diversified firms. What is surprising at first is the negative impact of leverage. Although the motive of bondholder expropriation makes refocusing more attractive for leveraged firms, it is also possible that refocusing is unattractive to a leveraged firm whose divisions have negatively correlated cash flows, because dropping a division would amplify the variance of firm cash flow and perhaps lower debt capacity. Not surprisingly, a high financing gap or low cash flow predicts a cash-producing divestiture but not a zero-revenue spin-off.

<sup>3</sup> Because one can argue that firms may refocus to fix investment inefficiency and because we wish to examine the importance of other factors besides these three variables, we also estimate this probit specification without them. The results are qualitatively unaffected. An earlier version of the article, available from the authors, contains results without these variables.

<sup>4</sup> Because we construct the forecast-error variable from the panel autoregression, following Newey (1984), we adjust the standard errors in the probit regression to account for this two-step procedure.

**Table 2**  
Probit regressions to predict spin-offs and divestitures

	Asset-based measures		Sales-based measures	
	Spin-offs	Divestitures	Spin-offs	Divestitures
Log assets (sales)	0.0145 <sup>a</sup> (0.0057)	0.0009 (0.0060)	0.0151 <sup>a</sup> (0.0088)	0.0144 (0.0149)
Market to book	0.0690 (0.0453)	0.0169 (0.0458)	0.0033 (0.0194)	0.0011 (0.0325)
Cash flow	-0.0679 (0.3409)	-0.7623 <sup>a</sup> (0.3045)	-0.1165 (0.3079)	-0.8303 <sup>b</sup> (0.4241)
Forecast error	0.1193 (0.2199)	0.8426 <sup>a</sup> (0.2967)	0.2398 (0.2143)	0.5102 <sup>b</sup> (0.2611)
Coefficient of variation	0.3184 <sup>a</sup> (0.1510)	0.0271 (0.0221)	0.0604 <sup>a</sup> (0.0127)	0.0110 (0.0151)
Debt	-0.2245 <sup>a</sup> (0.1037)	-0.7664 <sup>a</sup> (0.3026)	-0.1380 (0.0907)	-0.8631 <sup>a</sup> (0.3657)
Relative entropy	0.0167 <sup>a</sup> (0.0072)	0.1763 <sup>a</sup> (0.0639)	0.0165 <sup>b</sup> (0.0086)	0.1137 <sup>a</sup> (0.0476)
RVA	0.0384 <sup>a</sup> (0.0130)	0.0835 <sup>a</sup> (0.0402)	0.0222 <sup>a</sup> (0.0108)	0.0516 <sup>a</sup> (0.0247)
RINV	-0.0010 (0.0080)	-0.0022 (0.0107)	-0.0029 (0.0059)	-0.0020 (0.0102)
XVAL	-0.0653 <sup>b</sup> (0.0344)	-0.1965 <sup>b</sup> (0.1040)	-0.2031 <sup>a</sup> (0.0454)	-0.4608 <sup>b</sup> (0.2427)
Financing gap	-0.4056 (0.3889)	0.7568 <sup>a</sup> (0.3112)	0.0387 (0.2882)	0.1325 <sup>a</sup> (0.0666)
Profit shock	0.3571 <sup>a</sup> (0.1233)	0.3772 <sup>a</sup> (0.1185)	0.2952 <sup>a</sup> (0.1206)	0.2257 (0.1455)
IPO activity	0.2438 (0.5663)	0.5662 <sup>a</sup> (0.2079)	0.2686 (0.5963)	0.6307 <sup>a</sup> (0.3018)
Control activity	0.2321 <sup>a</sup> (0.1089)	0.0636 (0.0769)	0.3525 <sup>a</sup> (0.1566)	0.0893 (0.1528)
Industry sales growth	0.3017 <sup>b</sup> (0.1466)	-0.0613 (0.1857)	0.2968 <sup>b</sup> (0.1686)	0.0644 (0.1839)
Market share	0.3502 <sup>a</sup> (0.1299)	0.3229 <sup>a</sup> (0.1117)	0.3684 <sup>a</sup> (0.1566)	0.3680 <sup>a</sup> (0.1249)
Largest segment profit	-0.0946 (0.1923)	0.2125 <sup>a</sup> (0.0692)	-0.0928 (0.1555)	0.2205 <sup>a</sup> (0.0668)
Pseudo- <i>R</i> <sup>2</sup>	0.2724	0.4816	0.2207	0.3532

Mean marginal effects are reported, and standard errors are in parentheses under the parameter estimates. The sample consists of 102 spin-offs from 1983 to 1997, 120 divestitures from 1984 to 1994, and 461 control firms from 1983 to 1997. XVAL, RINV and RVA are calculated according to Equations (5), (4), and (1). The entropy measure of relatedness of segments is from Palepu (1985). The forecast error is the absolute value of the difference between the actual value and the mean analyst estimate of earnings per share in the fiscal year preceding the spin-off or divestiture. This variable is normalized by the ratio of shares to either assets or sales. The coefficient of variation is the coefficient of variation of analysts' estimates of earnings per share in the fiscal year preceding the spin-off or divestiture. Market share is firm sales divided by three-digit industry sales. Assets is the natural logarithm of the book value of assets. Sales is the natural logarithm of the book value of sales. Assets are used in the first two regressions, and sales are used in the second two. Cash flow is income plus depreciation, and the financing gap is cash flow plus net debt issuance minus net capital expenditures. Profit shock is the forecast error to operating income calculated from a panel autoregression. IPO activity is the value of all IPOs in a two-digit industry, and control activity is the value of corporate control transactions in two-digit industry. Both are divided by the sum of the market capitalizations of the firms in that industry. Industry sales growth is measured at the two-digit level. Largest segment profit is the operating income of the largest segment of the conglomerate. Cash flow, debt, largest segment profit, and the financing gap are all normalized by either the book value of total assets or the book value of total sales. The letter "a" indicates significance at the 5% level, and the letter "b" at the 10% level.

Several measures of profitability and value are significant, namely, the shock to profits, the profitability of the largest segment, RVA and excess value. The negative impact of excess value is consistent with the notion that firms refocus to raise value. The positive impact of the profit shock suggests that if a firm's marginal decision to refocus depends on the realization of this surprise shock, it will not experience the same improvement in investment efficiency as an otherwise identical firm that would have refocused in the absence of any such shocks. In contrast, the market-to-book ratio and RINV are not significant. Finally, three industry variables are significant: higher industry sales growth and more corporate control transactions matter for spin-offs and more IPO activity matters for divestitures.

In the treatment-effect estimation that follows, the control variables are those variables that are significant at the 10% level in the relevant probit regression. For example, for the asset-based measures of efficiency in the divestiture sample, we use all of the control variables except the log of assets, the market-to-book ratio, the coefficient of variation, RINV, industry sales growth, and corporate control activity.

### 3.4 Treatment effect estimation methodology

With an appropriate vector of control variables in hand, we now use the AI technique to analyze the “treatment” of a conglomerate with a spin-off or divestiture. We prefer this technique over previous approaches because it puts no parametric assumptions on the distributions of the variables. This flexibility is especially important for COMPUSTAT data, because incorrect distributional assumptions are well known to bias standard errors, and because the distributions of many balance-sheet and income-statement items cannot be approximated by the normal or logistic distributions—the two distributions most often used in this area.

Before describing the AI method, we present the generic problem of obtaining consistent estimates of treatment effects. We start with notation. Let  $R$  be a binary variable that is 1 if the firm refocuses and 0 otherwise. Let  $F_i(R)$  denote a measure of the *level* of investment efficiency as a function of  $R$  for observation  $i$ . Given this notation,  $E(F_i(1)|R = 1)$  indicates the expected value of the treatment on the treated group, given that treatment happens. Similarly,  $E(F_i(0)|R = 1)$  indicates the (hypothetical and unobservable) expected value of no treatment, given that treatment takes place.

In most of our applications, we look at the change in  $F_i(R)$  relative to its value before treatment, which we denote as  $\Delta F_i(R)$ . This “difference-in-difference” method accounts for time-invariant unobservable differences between the treatment and the control groups. The method is analogous to differencing to eliminate-fixed effects in panel data.

We estimate the average effect of a spin-off or divestiture on investment efficiency for the sample of firms that refocus, that is, the average treatment effect on the treated:

$$\tau|_{R=1} \equiv E(\Delta F_i(1) - \Delta F_i(0) | R = 1). \quad (7)$$

Because we cannot observe the effect of treatment on a firm and then roll back time to observe the effect of no treatment on that same firm,  $E(\Delta F_i(0)|R = 1)$  is unobservable.

In contrast, the recent work on spin-offs and divestitures has calculated

$$E(\Delta F_i(1) | R = 1), \quad (8)$$

by averaging the difference in efficiency of treated firms before and after treatment. However, Equation (8) is a biased estimator of Equation (7), unless  $E(\Delta F_i(0)|R = 1) = 0$ , which occurs only if the firms that chose to refocus would not have experienced any change in investment efficiency in the absence of refocusing. This requirement only holds if refocusing is the sole way to improve investment efficiency or if the refocusing firms have no other attributes that affect investment efficiency. The first condition is clearly false; and the second one, although unlikely, is a matter for empirical investigation—a matter that requires estimating  $\tau|_{R=1}$ .

Because  $\tau|_{R=1}$  is inherently unobservable, special assumptions are necessary to estimate its unobservable component:  $E(\Delta F_i(0)|R = 1)$ . The most common assumption of this type in the literature is that assignment to treatment is unconfounded, which means that assignment to treatment is random, conditional on a set of observable pretreatment variables,  $Z$ . Simple matching estimators exploit this assumption by matching each treated unit to one or more untreated units with similar values for the pretreatment variables,  $Z$ . Then,  $E(\Delta F_i(0)|R = 1)$  is estimated by averaging  $\Delta F_i(0)$  over the matches (the control observations). An estimate of  $\tau|_{R=1}$  is obtained by subtracting this estimate of  $E(\Delta F_i(0)|R = 1)$  from the average of  $\Delta F_i(1)$  over the treated observations. The matching is typically done without replacement. [See, for example, Rosenbaum (1989, 1995), Rubin (1973a,b), Dehejia and Wahba (1999)].

Having presented the general issues involved in estimating treatment effects, we now turn to the AI technique, providing a heuristic description.<sup>5</sup> The AI estimator is identical to the simple matching estimator described above, with two important extensions, which are motivated by the

---

<sup>5</sup> A more precise and technical description is in the appendix, along with Monte Carlo experiments to evaluate the finite sample properties of this estimator.

demonstration in AI that simple matching estimators are asymptotically biased if the vector  $Z$  contains more than one element. Asymptotic bias arises in matching estimators because the matches are not exact. The first extension is matching with replacement to reduce asymptotic bias. The second is the construction of an additional bias correction term.

The bias adjustment need only be made for the estimate of term  $E(\Delta F_i(0)|R = 1)$  in Equation (7), because the term  $E(\Delta F_i(1)|R = 1)$  is observable. The adjustment is an estimate of the difference between two terms. The first is what the effect of treatment on the control group would be in the case of perfect matching, and the second is what the effect actually is. The estimates of these two components are based on the conditional expectation of  $\Delta F_i(0)$  given  $Z_i$ , which we estimate by regressing  $\Delta F_i(0)$  on  $Z_i$  in the control sample. The estimate of the conditional expectation is then constructed as  $\hat{m}_0(Z_i) \equiv \hat{b}_0 + \hat{b}_1 Z_i$ , where  $\hat{b}_0$  and  $\hat{b}_1$  are the estimated regression coefficients, with  $\hat{b}_0$  a scalar and  $\hat{b}_1$  a vector with the same dimension as  $Z_i$ . The bias-adjusted estimate of  $E(\Delta F_i(0)|R = 1)$  is then the simple estimate described above plus a term we denote as  $\hat{m}_0(Z_i) - \hat{m}_0(Z_j)$ . This term is the difference in the predicted values of  $\Delta F_i(0)$  using the vector of controls for the  $i^{\text{th}}$  treated unit and the vector of controls for its corresponding match, indexed by  $j$ .

### 3.5 Treatment effect results

With this estimator, we now turn to the central question of whether refocusing causes improvements in investment efficiency. To understand the importance of the treatment-effect estimates that follow, it is useful to review previous results in the literature. Burch and Nanda (2003) find that an asset-based measure of RVA increases significantly after a spin-off. Ahn and Denis (2004) confirm this result using a sales-based measure of RVA, a sales-based measure of RINV, and a measure of excess value. They also confirm the result in Gertner, Powers, and Scharfstein (2002) that the sensitivity of investment to industry median  $q$  rises after a spin-off. To be comparable to these studies, for our post-spin-off calculations of excess value, RINV, and RVA, we treat the spin-off unit as if it were still a part of the firm. Similarly, we confine the analysis to multi-segment parent firms, because measures of excess value and investment efficiency are meaningless for single-segment parents.

Table 3 summarizes our replication of these results for spin-offs. It also compares this replication to the treatment effects estimated with the AI technique. The table is organized as follows. The first section presents average levels of RVA, RINV, and excess value before and after the spin-off, as well as the average changes in these three variables. The average change is an estimate of the quantity  $E(\Delta F_i(1)|R = 1)$ . The next section contains the level treatment effects, which are estimates of

**Table 3**  
**Measures of firm value and investment efficiency before and after a spin-off**

	Asset-based measures			Sales-based measures			Investment- <i>q</i> sensitivity
	Excess value	RINV	RVA	Excess value	RINV	RVA	
<b>Means</b>							
Before	-0.1088 <sup>a</sup> (0.0375)	-0.0099 <sup>a</sup> (0.0030)	-0.0181 <sup>a</sup> (0.0047)	-0.1737 <sup>a</sup> (0.0975)	-0.0073 (0.0047)	-0.0016 <sup>a</sup> (0.0007)	0.0059 <sup>a</sup> (0.0010)
After	0.1231 <sup>a</sup> (0.0174)	-0.0032 (0.0023)	0.0022 <sup>a</sup> (0.0011)	0.0330 (0.0542)	0.0011 (0.0012)	0.0009 <sup>a</sup> (0.0002)	0.0170 <sup>a</sup> (0.0056)
Change	0.2319 <sup>a</sup> (0.0413)	0.0067 <sup>b</sup> (0.0038)	0.0203 <sup>a</sup> (0.0048)	0.2067 <sup>b</sup> (0.1116)	0.0084 <sup>b</sup> (0.0049)	0.0025 <sup>a</sup> (0.0007)	0.0111 <sup>b</sup> (0.0057)
<b>Level Treatment Effects</b>							
Before	-0.1088 <sup>a</sup> (0.0375)	-0.0099 <sup>a</sup> (0.0030)	-0.0181 <sup>a</sup> (0.0047)	-0.1737 <sup>a</sup> (0.0975)	-0.0073 (0.0047)	-0.0016 <sup>a</sup> (0.0007)	0.0059 <sup>a</sup> (0.0010)
After	-0.0364 (0.0068)	-0.0068 (0.0115)	-0.0055 (0.0044)	-0.0549 (0.0485)	-0.0021 (0.0046)	-0.0005 (0.0006)	0.0075 (0.0026)
Bias adjustment	0.0295	-0.0096	-0.0004	-0.1488	-0.0019	-0.0003	-0.0020
Change	0.0624 (0.0381)	0.0031 (0.0119)	0.0126 <sup>b</sup> (0.0064)	0.1188 (0.1089)	0.0052 (0.0066)	0.0011 (0.0009)	0.0016 (0.0028)
Difference in difference treatment effects	0.0203 (0.0577)	0.0024 (0.0015)	0.0041 (0.0025)	0.0848 (0.0568)	0.0046 (0.0031)	0.0016 (0.0011)	0.0021 (0.0050)
	0.0058	-0.0003	-0.0007	-0.0231	-0.0005	-0.0003	-0.0015

The sample consists of 102 spin-offs from 1983 to 1997. Excess Value, RINV, and RVA are calculated according to Equations (5), (1), and (4). The figures reported in the first section are sample means. Standard errors are in parentheses under the parameter estimates. The figures labeled as "Before" in the second section are identical to their counterparts in the first section. The remaining figures in the second and third sections are treatment effects, calculated with a matched sample of diversified firms using the procedure in Abadie and Imbens (2006). Standard errors are in parentheses under the parameter estimates. Bias adjustment terms are in parentheses under the standard errors. The letter "a" indicates significance at the 5% level, and the letter "b" at the 10% level.

$E(F_i(1)|R = 1)$  or, equivalently, the average post spin-off level of these three variables relative to the level in the control sample.<sup>6</sup>

In the final section we present the difference-in-difference treatment effects, which are estimates of  $E(\Delta F_i(1) - \Delta F_i(0)|R = 1)$  or, equivalently, the average changes in these three variables relative to the changes in the control sample. As noted above, the difference-in-difference estimator accounts for time-invariant but unobservable control variables, whereas the level treatment-effect estimator does not.

The first section of Table 3 replicates the qualitative conclusions in Ahn and Denis (2004), Burch and Nanda (2003), and Gertner, Powers, and Scharfstein (2002). The replication is not exact because our sample differs slightly. The first two rows present the averages of excess value, RINV, RVA, and investment-*q* sensitivity over the three years preceding the

<sup>6</sup> Using the levels of RINV, RVA, and excess value as controls is not valid in the *level* treatment-effect estimations, because these variables would be used to explain themselves. However, their use is valid in the *difference-in-difference* treatment-effect estimations. As in a time-series error-correction model, we use the lagged levels of these variables to explain their current changes.

spin-off and over the three years after the spin-off.<sup>7</sup> Before the spin-off, both the asset-based and sales-based measures of excess value, RINV, and RVA are negative and significantly different from zero. The simulations in Section 1 assist in interpreting the economic significance of RINV and RVA. The estimates of RINV are slightly less than half the average figure for the simulated socialist firm. Although the sales-based estimate of RVA is close to the average figure for the simulated efficient firm, the asset-based estimate is at a level consistent with the simulated socialist firm. In sum, these figures reflect a modest amount of inefficiency before the spin-off. In contrast, after the spin-off, the asset-based measure of excess value becomes significantly positive; and the sales-based measure of excess value becomes insignificantly different from zero, as do both the asset- and sales-based measures of RVA and RINV. Because the simulations from Section 1 indicate that observed values of 0 for RINV and RVA constitute efficient investment, most of this evidence represents an economically meaningful improvement in efficiency. Finally, the estimated sensitivity of investment to industry median  $q$  almost triples after the spin-off. The changes in all of these variables, reported in the third row, are significantly different from zero at either the 5% or the 10% level.

Although this evidence shows that improvements in efficiency accompany spin-offs, inferring causation from spin-offs to efficiency improvements requires estimation of a treatment effect.<sup>8</sup> The next section of Table 3 repeats the first row of the first section but replaces the second row with the level treatment-effect estimates. To interpret these estimates, note that the reference point is not 0 but is at the pretreatment level of the variable in question. All of the estimated level treatment effects are closer to the pretreatment levels. In addition, all have large standard errors; only one of the differences between the “before” and the “after” rows is significantly different from zero and only at the 10% level. The next line of the table presents bias corrections from the AI estimators. These figures are estimates of the term  $\hat{m}_r(Z_i) - \hat{m}_r(Z_j)$ , which, as explained above, is an adjustment for the differences between the treated and the control firms. A positive value for the term indicates that the unadjusted treatment effect is lower than the adjusted and vice versa. Because these bias adjustments are often substantial relative to the magnitude of the estimated effect, it is clear that making the adjustments is important for inference.

<sup>7</sup> Although, in principle, it is possible to address the measurement error problem in the regressions of investment on  $q$ , our spin-off sample is too small to use the Erickson and Whited (2002) estimators. The tests for model identification never produce rejections of the null hypothesis that the model is unidentified.

<sup>8</sup> Calculating treatment effects for excess value, RINV, and RVA requires a straightforward application of the AI technique, because it applies to scalar variables. However, examining investment- $q$  sensitivity requires creation of a new scalar variable whose mean is the slope-coefficient from a regression of investment on  $q$ :  $Ny_i x_i / \sum_i x_i^2$ , where  $y_i$  and  $x_i$  are the de-meaned rate of investment and industry median  $q$ .

The last row of Table 3 contains the difference-in-difference treatment effects, reporting the estimated changes in the efficiency measures relative to the changes in these variables in a control group. None of these estimates is significantly greater than 0. Furthermore, none is larger than the changes reported in the first or second sections. These results indicate the importance of controlling for time-invariant, unobserved control variables. More importantly, these results are precisely what one would expect if  $E(\Delta F_i(0)|R = 1) \neq 0$ . In other words, the results confirm that firms that choose to spin off are indeed different from their counterparts that do not, and this difference drives the observed improvements in investment efficiency rather than the spin-off per se.

Next, we conduct a similar analysis for divestitures, once again restricting ourselves to multi-segment parents. Table 4 is analogous to Table 3, with two differences. As explained in Section 2, for consistency with the results in Dittmar and Shivdasani (2003), we use a slightly different measure of the diversification discount: Equation (6) instead of Equation (5). For consistency, we also do not examine investment- $q$  sensitivity, which is not part of the Dittmar and Shivdasani analysis. The first section shows

**Table 4**  
Measures of firm value and investment efficiency before and after a divestiture

Means	Asset-based measures			Sales-based measures		
	Discount	RINV	RVA	Discount	RINV	RVA
Before	0.1863 <sup>b</sup> (0.1107)	-0.0082 <sup>a</sup> (0.0024)	-0.0011 (0.0007)	0.1284 <sup>a</sup> (0.0408)	-0.0032 (0.0031)	-0.0026 <sup>a</sup> (0.0009)
After	-0.0248 <sup>b</sup> (0.0130)	-0.0032 <sup>a</sup> (0.0014)	0.0012 <sup>a</sup> (0.0006)	-0.1820 <sup>a</sup> (0.0321)	-0.0018 <sup>a</sup> (0.0004)	0.0008 <sup>a</sup> (0.0002)
Change	-0.2111 <sup>b</sup> (0.1115)	0.0050 <sup>b</sup> (0.0028)	0.0023 <sup>a</sup> (0.0009)	-0.3104 <sup>a</sup> (0.0519)	0.0014 (0.0031)	0.0034 <sup>b</sup> (0.0009)
Level treatment effects						
Before	0.1863 <sup>b</sup> (0.1107)	-0.0082 <sup>a</sup> (0.0024)	-0.0011 (0.0007)	0.1284 <sup>a</sup> (0.0408)	-0.0032 (0.0031)	-0.0026 <sup>a</sup> (0.0009)
After	0.0093 (0.0060)	-0.0067 (0.0204)	-0.0007 (0.0005)	0.0630 <sup>a</sup> (0.0279)	-0.0038 (0.0038)	-0.0023 (0.0021)
Bias adjustment	0.0085	-0.0023	-0.0001	-0.0142	0.0023	-0.0008
Change	-0.1770 (0.1109)	0.0015 (0.0205)	0.0004 (0.0009)	-0.0654 (0.0494)	-0.0006 (0.0049)	0.0003 (0.0023)
Difference in difference treatment effects	-0.1209 (0.0851)	0.0010 (0.0023)	-0.0008 (0.0007)	-0.0592 (0.0519)	-0.0005 (0.0009)	-0.0008 (0.0034)
	0.0162	-0.0013	-0.0009	0.0131	-0.0003	-0.0007

The sample consists of 120 divestitures from 1984 to 1994. The discount, RINV, and RVA are calculated according to Equations (6), (1) and (4). The figures reported in the first section are sample means. Standard errors are in parentheses under the parameter estimates. The figures labeled as "Before" in the second section are identical to their counterparts in the first section. The remaining figures in the second and third sections are treatment effects, calculated with a matched sample of diversified firms using the procedure in Abadie and Imbens (2006). Standard errors are in parentheses under the parameter estimates. Bias adjustment terms are in parentheses under the standard errors. The letter "a" indicates significance at the 5% level, and the letter "b" at the 10% level.

the same pattern as in the case of spin-offs: the diversification discount falls after a divestiture, and measures of investment efficiency rise. However, the economic significance of these changes is smaller than in the case of spin-offs; indeed, the pretreatment levels of RVA and RINV are close to levels associated with efficient investment in the Section 1 simulations. Nonetheless, as before, the second two sections show that this evidence is likely an artifact of endogeneity. We find no statistically significant differences between the control group and the divesting group.

It is possible that our findings of insignificant treatment-effect estimates are an artifact of poor power of this estimator to detect actual non-zero treatment effects. The Monte Carlo simulations in the Appendix address this question. The results show that for a true treatment equal to 0.02, the AI tests have excellent power to reject the null of no treatment. To interpret this result, consider the extreme scenario in which the simple efficiency changes in the first sections of Tables 3 and 4 represent the "truth." All of the changes in excess value and the discount, as well as one of the changes in RVA reported here are greater than 0.02; so we can be confident that the zero treatment-effect estimates corresponding to these particular figures are not spurious. The remaining eight changes reported in the first sections of Tables 3 and 4 are smaller, however, ranging from 0.0014 to 0.0111. The Monte Carlo simulations show that for true treatments ranging from 0.001 to 0.01, power ranges between 0.187 to 0.470. Although these power figures appear small, a simple calculation shows that they are useful for interpreting the results. The average of the smallest eight changes in the top sections of Tables 3 and 4 is 0.051, and the power to detect a treatment effect of 0.05 is 0.375. If these reported changes are the "truth," then the expected number of significant treatment-effect estimates is three out of eight. In contrast, in Tables 3 and 4, none of the difference-in-difference estimates are significant.

Two main points can be taken from this exercise. First, some of our results are clearly not the result of poor power. Second, although some of the rest of our results may be because of poor power, it is unlikely that all are. Two further points are relevant. First, the results that may be because of poor power represent cases in which the original observed inefficiencies are quite small and, according to the simulations in Section 1, of little economic significance. Second, these small estimated inefficiencies may actually reflect efficient investment, because the Section 1 simulations indicate that 20–30% of the time estimates of RINV and RVA create false rankings of efficient and inefficient firms.

### 3.6 Alternative estimators

We now use two more familiar but also more parametric estimators to examine the robustness of our qualitative results. The first alternative is based on an analogy to the Heckman (1979) procedure for correcting for

endogeneity. Because estimating a mean is the same as running a regression on a constant term, one can characterize the comparison of average investment efficiency before and after refocusing as the following regression

$$\Delta F_i(R_i) = \alpha + \delta R_i + e_i, \quad (9)$$

in which  $R_i = 1$  if the firm is in the treatment sample and  $R_i = 0$  if the firm is in the control sample. The coefficient on the constant,  $\alpha$ , is the average change in investment efficiency in the control sample, and the sum of the coefficients ( $\alpha + \delta$ ) is the average change in efficiency in the treated sample. If refocusing is endogenous,  $e_i$  is correlated with  $R_i$ , and the OLS estimate of  $\delta$  is therefore biased. As pointed out by Heckman (1979), this problem is equivalent to an omitted variable problem, in which the omitted variable is the inverse Mills ratio associated with refocusing. A consistent estimate of  $\delta$  is obtained by estimating the inverse Mills ratio with a probit regression, as in Section 3.3, and then including this estimate in Equation (9). Using this method for investigating investment- $q$  sensitivity is analogous, except that the coefficient of interest is the coefficient on the interaction of the refocusing dummy and  $q$  in a regression of investment on  $q$  and this interaction term.

The second alternative is the propensity score matching method in Dehejia and Wahba (1999, 2002). Because this method is spelled out in detail in Villalonga (2004b), we present only a brief outline. Essentially, the technique is a matching estimator in which the only observable matching variable is the propensity score to refocus. Rosenbaum and Rubin (1983) showed that if treatment assignment is unconfounded conditional on  $Z_i$ , then it is also unconfounded conditional on the propensity score associated with  $Z_i$ . Dehejia and Wahba (1999, 2002) proposed the following algorithm for implementing this simple idea. First, separate observations into treated and control groups. Second, estimate the propensity scores for both the treated and the control samples. In this step, we use the fitted value from a simple probit model that contains the same control variables used for the AI estimator. Third, discard all control observations whose propensity scores fall outside the interval defined by the minimum and maximum propensity scores for the treated group. Fourth, sort firms into groups based on the propensity score for the refocusing firms. Fifth, for each control variable, do a simple  $t$ -test on the difference in means between the treated and control observations within each group. If the null of the same mean is rejected, refine the groups. (Here, we find that fourths are a sufficiently fine classification scheme.) Finally, take the weighted average of the within-group means of  $\Delta F(\cdot)$  between the treated and the untreated units.

**Table 5**  
Alternative estimates of treatment effects

	Asset-based measures			Sales-based measures			Investment- <i>q</i> sensitivity
	Excess value	RINV	RVA	Excess value	RINV	RVA	
<b>Spin-offs</b>							
Dehejia and Wahba	0.0065 (0.0088)	-0.0024 (0.0255)	-0.0009 (0.0071)	-0.0021 (0.0015)	0.0014 (0.0321)	0.0026 (0.0112)	-0.0007 (0.0010)
Heckman	0.0060 (0.0203)	-0.0020 (0.0185)	0.0069 <sup>b</sup> (0.0040)	-0.0026 (0.0210)	0.0017 (0.0179)	0.0014 (0.0066)	0.0019 (0.0024)
Inverse Mills ratio	0.0366 <sup>b</sup> (0.0215)	0.0621 <sup>a</sup> (0.0203)	0.0454 <sup>a</sup> (0.0052)	0.0506 <sup>a</sup> (0.0196)	0.0470 <sup>a</sup> (0.0210)	0.0537 <sup>a</sup> (0.0055)	0.0409 <sup>a</sup> (0.0187)
<b>Divestitures</b>							
Dehejia and Wahba	-0.0773 (0.0420)	0.0021 (0.0510)	0.0021 (0.0123)	-0.0353 (0.0490)	-0.0046 (0.0023)	-0.0006 (0.0101)	
Heckman	0.0027 (0.0191)	-0.0083 (0.0142)	-0.0058 (0.0055)	-0.0075 (0.0192)	-0.0133 (0.0148)	0.0007 (0.0005)	
Inverse Mills ratio	0.0672 <sup>a</sup> (0.0176)	0.0604 <sup>a</sup> (0.0208)	0.0211 <sup>a</sup> (0.0050)	0.0491 <sup>a</sup> (0.0187)	0.0235 (0.0192)	0.0220 <sup>a</sup> (0.0051)	

The sample consists of 102 spin-offs from 1983 to 1997 and 120 divestitures from 1984 to 1994. Excess value, RINV, and RVA are calculated according to Equations (5), (1), and (4). The figures reported are treatment effects. Standard errors are in parentheses under the parameter estimates. Treatment effects are calculated with a matched sample of diversified firms using the procedures in Heckman (1979) and Dehejia and Wahba (1999, 2002). The rows labeled "Inverse Mills Ratio" present the coefficients on this variable in the Heckman regressions. The letter "a" indicates significance at 5% level, and the letter "b" at the 10% level.

Our results from these two procedures are in Table 5. For brevity, we report only the difference-in-difference estimates. The results here confirm our earlier findings. The Dehejia–Wahba treatment-effect estimates are all insignificantly different from zero, and only one of the Heckman-style estimates is significantly negative. This table also summarizes the coefficients on the inverse Mills ratio. All of the estimated coefficients on the inverse Mills ratio are in the 0.04–0.06 range, and all but one are significant at the 5% level. This result indicates the importance of self-selection; that is, characteristics that make firms choose to refocus are, on average, positively correlated with investment efficiency. Monte Carlo results in Table A1 in the Appendix indicate that these estimators have similar power to the AI estimator to detect a non-zero treatment effect.

### 3.7 Robustness of results to alternative interpretations

One final issue needs to be addressed in examining endogeneity. We have thus far concluded that refocusing appears to affect investment efficiency only because of its correlation with other variables that affect investment efficiency. It is also possible, however, that refocusing has no incremental effect on investment efficiency beyond other unobserved "treatments" used in the group of control firms. The possibility of no incremental effect of treatment arises because of the negative impact of XVAL on the

decision to refocus reported in Table 2. In other words, the negative impact of XVAL introduces the possibility that conglomerate investment is ex ante inefficient in both the treated and the control groups and that the only difference between the two groups is their choice of treatment. The treated group chooses refocusing and the control group chooses other, unobserved treatments.

We now try to differentiate between this possibility of the equivalence of alternative treatments and our original conclusion that the observed causation between refocusing and efficiency is an artifact of endogenous selection. To do so, we omit XVAL, RINV, and RVA from the vector of control variables. If it is industry shocks and other variables that are really driving the results, and not ex ante conglomerate inefficiency, then we should see no difference in the matching results. Table 6 summarizes the results from implementing the AI, the Dehejia–Wahba, and the Heckman estimators on the spin-off and divestiture samples without these three variables in the vector of controls. For brevity, we report only the difference-in-difference results. We find almost no difference between these results and the corresponding results in Tables 3, 4, and 5. This result is not surprising. In unreported probit regressions of the

**Table 6**  
Treatment effect estimates with alternative controls

	Asset-based measures			Sales-based measures			Investment- <i>q</i> Sensitivity
	Excess value	RINV	RVA	Excess value	RINV	RVA	
<b>Spin-offs</b>							
Abadie and Imbens	0.0371 (0.0660)	0.0038 (0.0026)	0.0052 (0.0038)	0.0945 (0.0765)	0.0053 (0.0040)	0.0017 (0.0016)	0.0023 (0.0034)
Dehejia and Wahba	0.0239 (0.0157)	0.0008 (0.0036)	0.0019 (0.0069)	0.0016 (0.0014)	0.0037 (0.0384)	0.0037 (0.0214)	0.0011 (0.0009)
Heckman	0.0085 (0.0213)	-0.0018 (0.0179)	0.0149 <sup>b</sup> (0.0085)	0.0054 (0.0211)	0.0047 (0.0164)	0.0016 (0.0056)	0.0021 (0.0023)
Inverse Mills ratio	0.0867 <sup>b</sup> (0.0223)	0.0308 <sup>b</sup> (0.0161)	0.0562 <sup>a</sup> (0.0055)	0.0694 <sup>a</sup> (0.0247)	0.0295 <sup>a</sup> (0.0104)	0.0533 <sup>a</sup> (0.0054)	0.0362 <sup>a</sup> (0.0157)
<b>Divestitures</b>							
Abadie and Imbens	-0.1333 (0.0907)	0.0013 (0.0024)	-0.0002 (0.0008)	-0.0612 (0.0654)	0.0003 (0.0008)	-0.0005 (0.0006)	
Dehejia and Wahba	-0.0818 (0.0385)	0.0035 (0.0068)	0.0021 (0.0098)	-0.0620 (0.0623)	-0.0010 (0.0026)	0.0020 (0.0114)	
Heckman	0.0023 (0.0186)	-0.0065 (0.0136)	0.0004 (0.0005)	-0.0085 (0.0188)	-0.0095 (0.0139)	0.0016 (0.0012)	
Inverse Mills Ratio	0.0521 <sup>a</sup> (0.0178)	0.0495 <sup>a</sup> (0.0180)	0.0246 <sup>a</sup> (0.0048)	0.0673 <sup>a</sup> (0.0172)	0.0273 (0.0188)	0.0228 <sup>a</sup> (0.0049)	

The sample consists of 102 spin-offs from 1983 to 1997 and 120 divestitures from 1984 to 1994. Excess value, RINV, and RVA are calculated according to Equations (5), (1), and (4). The figures reported are treatment effects. Standard errors are in parentheses under the parameter estimates. Treatment effects are calculated with a matched sample of diversified firms using the procedures in Heckman (1979) and Dehejia and Wahba (1999, 2002). The rows labeled “Inverse Mills Ratio” present the coefficients on this variable in the Heckman regressions. The letter “a” indicates significance at the 5% level, and the letter “b” at the 10% level.

decision to spin off or divest on a vector of controls that excludes these three variables, we find that their exclusion lowers the pseudo- $R^2$  by very little. We therefore conclude that exogenous shocks are driving our results rather than many poorly performing conglomerates seeking different sorts of treatments.

#### 4. Measurement Error

Having demonstrated that refocusing, per se, does not *cause* increases in RINV, RVA, and investment- $q$  sensitivities, we turn to the second important problem that may have contaminated previous inference about conglomerate investment. In particular, we analyze the following regression used by Dittmar and Shivdasani (2003):

$$y_i = \hat{\alpha}_0 + \hat{\beta}x_i + \hat{\alpha}_1z_i + u_i^*. \quad (10)$$

Here,  $y_i$  is the rate of investment of segment  $i$ ;  $x_i$  is its observed  $q$ ;  $\beta$  is the sensitivity of investment to observed  $q$ ;  $u_i^*$  is the OLS regression disturbance;  $z_i$  is a vector containing a segment's own cash flow and other-segment cash flow;  $\alpha_1$  is its corresponding coefficient vector; and  $\alpha_0$  is the regression intercept. A caret indicates that the coefficients are not necessarily the true coefficients that would be obtained in the absence of measurement error in  $q$ , but the possibly inconsistent OLS estimates. As explained in Section 2, several authors have argued that a high coefficient on other-segment cash flow indicates efficient investment with a smoothly functioning internal capital market.

As described in detail in Whited (2001), significant measurement error is likely to exist in the usual measure of  $x_i$ : the median  $q$  in the industry to which a segment belongs. Measurement error in  $q$  can bias the coefficient on other-segment cash flow, if cash flow in one division is correlated with the  $q$  of another. This correlation can occur even in diversified firms, because, as pointed out in Chevalier (2004) and Whited (2001), the divisions of these firms can be related.

We examine the measurement error problem with the third and higher order moment estimators in Erickson and Whited (2002). We opt for this technique because, as argued in Whited (2001) and Erickson and Whited (2000), more conventional remedies such as instrumental variables require assumptions that are both implausible and untestable.<sup>9</sup> The Erickson and Whited (2002) estimators used the following statistical model. Let  $x_i$  denote true unobserved  $q$ ; then the relationship between segment

<sup>9</sup> See Hennessy (2004) for another application of this technique.

investment, true unobserved  $q$ , its proxy, segment cash flow, and other-segment cash flow is given by

$$y_i = \alpha_0 + \beta\chi_i + \alpha_1z_i + u_i \quad (11)$$

$$x_i = \gamma_0 + \chi_i + \varepsilon_i. \quad (12)$$

Here,  $\varepsilon_i$  is a mean-zero error independent of  $(u_i, z_i, \chi_i)$ , and  $u_i$  is independent of  $(\chi_i, z_i)$ . In this model, the EW estimators provide consistent estimates of  $\beta$  and  $\alpha_1$ , even in the presence of the measurement error,  $\varepsilon_i$ .

The Appendix contains a brief description of the EW estimators, as well as the results of Monte Carlo simulations based on our data set. These simulations indicate that these estimators have fair power to detect non-zero cash-flow coefficients of the magnitude we find in the OLS estimates presented below. They have excellent power, however, to detect non-zero cash-flow coefficients of the magnitude found in Dittmar and Shivdasani (2003).

Table 7 summarizes the results from estimating Equation (10) with ordinary least squares, as in Dittmar and Shivdasani (2003). As in their article, we use only the sample of firms that remain diversified after the divestiture, because “other-segment cash flow” is meaningless for a single-segment firm. As a departure from their article, we do not include firm-fixed effects, because the EW estimators are not identified when we do so.<sup>10</sup> Another departure from their article is our use of assets instead of sales to deflate investment and cash flow. We choose assets because when we deflate by sales, the overidentifying restrictions implied by the EW estimators are strongly rejected by the data. The reason is intuitive: deflating  $q$  by assets and the other regression variables by sales, as in Dittmar and Shivdasani (2003), results in a heteroskedastic regression disturbance. The EW estimators require a homoskedastic regression error; and as shown in Erickson and Whited (2000), the generalized method of moments (GMM)  $J$ -test has good power in picking up this source of misspecification. We report the regression results for three years before the divestiture and for three years after the divestiture.

Table 7 replicates the general flavor of results in Dittmar and Shivdasani (2003): the sensitivity of investment to other-segment cash flow is significant both before and after a divestiture, and the magnitudes of the coefficients on other-segment cash flow increase significantly in the years after the divestiture. Note, however, that the  $R^2$ s of these regressions are quite small—a result to which we shall return below. One minor deviation of our results from those in Dittmar and Shivdasani (2003) is the slightly smaller magnitudes of our cash flow coefficients. Nonetheless, we are satisfied with our replication of their results, because we are

---

<sup>10</sup> The omission or inclusion of firm-fixed effects makes little difference for our OLS results.

**Table 7**  
OLS segment investment regressions

	Industry $q$	Own-segment cash flow	Other-segment cash flow	$R^2$
$T - 3$	0.006 <sup>a</sup> (0.001)	0.013 <sup>a</sup> (0.003)	0.008 <sup>a</sup> (0.002)	0.137 (0.114)
$T - 2$	0.008 <sup>a</sup> (0.003)	0.014 <sup>a</sup> (0.006)	0.010 <sup>a</sup> (0.003)	0.175 <sup>a</sup> (0.055)
$T - 1$	0.000 (0.009)	0.013 (0.016)	0.012 <sup>a</sup> (0.002)	0.041 (0.034)
$T + 1$	0.012 <sup>a</sup> (0.004)	0.033 <sup>a</sup> (0.015)	0.017 <sup>a</sup> (0.004)	0.155 (0.137)
$T + 2$	0.015 <sup>b</sup> (0.009)	0.034 <sup>b</sup> (0.019)	0.020 <sup>a</sup> (0.008)	0.041 <sup>a</sup> (0.013)
$T + 3$	0.008 <sup>a</sup> (0.002)	0.017 <sup>a</sup> (0.005)	0.021 <sup>a</sup> (0.009)	0.017 (0.012)

The sample consists of 139 divestitures from 1984 to 1994. We estimate a regression of segment investment on industry  $q$ , own-segment cash flow, and other-segment cash flow for three years preceding and three years following a divestiture. The regressions include calendar year effects and the standard errors are adjusted for heteroskedasticity as in White (1980). The letter "a" indicates significance at the 5% level, and the letter "b" at the 10% level.

**Table 8**  
GMM segment investment regressions

	Industry $q$	Own-segment cash flow	Other-segment cash flow	$\rho^2$	$\tau^2$	J-statistic	Identification test
$T - 3$	0.092 (0.109)	-0.013 (0.013)	0.002 (0.002)	0.554 (0.425)	0.167 (0.167)	0.047 (0.977)	0.611 (0.737)
$T - 2$	0.050 (0.038)	-0.004 (0.016)	0.006 <sup>a</sup> (0.003)	0.480 (0.520)	0.204 (0.865)	2.895 (0.235)	2.510 (0.285)
$T - 1$	0.077 (0.065)	0.006 (0.014)	0.002 (0.002)	0.444 <sup>a</sup> (0.137)	0.151 <sup>a</sup> (0.048)	1.124 (0.570)	5.077 (0.079)
$T + 1$	0.057 <sup>b</sup> (0.031)	0.029 (0.049)	0.009 (0.009)	0.628 <sup>b</sup> (0.375)	0.112 <sup>a</sup> (0.050)	0.993 (0.609)	5.193 (0.075)
$T + 2$	0.048 <sup>b</sup> (0.029)	0.021 (0.038)	-0.001 (0.001)	0.527 <sup>a</sup> (0.257)	0.152 <sup>a</sup> (0.056)	0.903 (0.637)	10.575 (0.005)
$T + 3$	0.048 <sup>a</sup> (0.023)	-0.007 (0.024)	0.001 (0.001)	0.444 <sup>b</sup> (0.235)	0.216 <sup>b</sup> (0.133)	0.578 (0.749)	25.546 (0.000)

The sample consists of 139 divestitures from 1984 to 1994. We estimate a regression of segment investment on industry  $q$ , own-segment cash flow, and other-segment cash flow for three years preceding and three years following a divestiture, using the fourth order GMM estimator in Erickson and Whited (2002).  $\rho^2$  is a consistent estimate of the  $R^2$  of the regression, and  $\tau^2$  is an index of measurement quality for industry  $q$  that ranges between 0 and 1. The regressions include calendar year effects. Standard errors are in parentheses under the parameter estimates, and p-values are in parentheses under the chi-squared tests. The letter "a" indicates significance at the 5% level, and the letter "b" at the 10% level.

unaware of any theoretical model that gives guidance in interpreting cash-flow coefficients in terms of frictions in internal capital markets.

Table 8 summarizes the results from estimating Equations (11) and (12) with the fourth-order moment estimator in Erickson and Whited (2002). We estimate five parameters: the coefficient on  $q$ , the regression  $R^2$ , the

$R^2$  of Equation (12), the coefficient on own-segment cash flow, and the coefficient on other-segment cash flow. As explained in Erickson and Whited (2000), the GMM estimate of the  $R^2$  of Equation (11), labeled  $\rho^2$ , is consistent and the OLS  $R^2$  is biased downward. The  $R^2$  of (12), labeled  $\tau^2$ , is a useful index of measurement quality for industry median  $q$ . A value of one indicates a perfect proxy and a value of 0 indicates a worthless one.

In short, this table confirms the evidence in Whited (2001) that OLS segment-level regressions of investment on industry  $q$  are contaminated by a substantial amount of measurement error. For example, the highest estimate of  $\tau^2$  is 0.216. Remedyng this poor measurement quality manifests itself in  $R^2$ s and coefficients on  $q$  that are several times higher than their OLS counterparts. Furthermore, all but one of the coefficients on own-segment cash flow and other-segment cash flow are insignificantly different from zero. These coefficients are all smaller in absolute value than their OLS counterparts. Finally, the GMM  $J$ -test never produces a rejection of the overidentifying restrictions. The model identification test described in Erickson and Whited (2002) rejects the null of no identification in all years after the divestiture and in one year before the divestiture. Nonetheless, the results in the two years for which the model is unidentified support the conclusion of no cash-flow effects. In sum, changes in cash-flow sensitivities estimated by OLS can be attributed to measurement error, as opposed to improvements in investment efficiency.

## 5. Conclusion

In writing this article, we have sought to understand the conflicting evidence on the causes of the diversification discount, in particular, inefficient investment. We challenge the findings in a series of recent articles that examine spin-offs and divestitures. Gertner, Powers, and Scharfstein (2002), Ahn and Denis (2004), Dittmar and Shivdasani (2003), and Burch and Nanda (2003) find that refocusing increases conglomerate value and investment efficiency. The logical implication of this finding is that conglomerates, on average, invest inefficiently. Spin-offs and divestitures are endogenous decisions, however. We treat endogeneity by measuring improvements in efficiency relative to the improvements observed in a group of control firms with similar characteristics. We choose these characteristics via their ability to predict divestitures and spin-offs. The characteristics include size, diversification, industry growth, industry IPO and control activity, and recent unanticipated shocks to profit. Although we confirm that improvements in efficiency follow refocusing, we demonstrate that it is unlikely that refocusing causes these improvements. Instead, the characteristics associated with the decision to refocus are also associated with improvements in investment efficiency. Finally, we show that some of the observed

improvements in investment efficiency are due to measurement error in  $q$  rather than to any economically meaningful forces.

Because it is unlikely that misallocation of capital expenditures is sufficiently important to influence conglomerate value, it may be fruitful to examine the allocation within conglomerates of other types of resources, such as managerial expertise. Data availability is a challenge for any efforts in this direction. Data on capital expenditures of conglomerate divisions are easily available. In contrast, other potentially useful information, such as the compensation of divisional managers, is much more difficult to collect. Future research could be directed toward such efforts.

## Appendix

This Appendix is divided into four parts. The first describes the simulations used in Section 1. The second presents a simple matching estimator. The third explains some of the technical details of the AI estimator, as well as relevant Monte Carlo simulations. The fourth exposit the EW estimators, once again with relevant Monte Carlo experiments.

### Efficiency simulations

We start by describing the simulations in Section 1 in general terms for the entire set of simulated conglomerates. We generate sets of 10,000 pseudo-conglomerates, in which each set is defined by the number of segments in each of its member conglomerates and in which we consider conglomerates with two, four, six, eight, and ten segments. Given this number, we simulate a conglomerate by assigning a simulated industry at random to each segment and then by picking a random firm out of the assigned industry to become a segment.

Next, we describe the “building blocks” of each conglomerate; that is, the simulated industries and the individual production units within each industry. In each simulated economy, we construct 100 “industries,” in which each industry is defined by its mean true  $q$ ,<sup>11</sup> and in which these means in turn have a lognormal distribution, with a mean equal to the sample average observed  $q$  in our data set and with a standard deviation described below.

Within each simulated industry, we generate 100 single-segment “firms,” each of which may be assigned at random as a segment of a conglomerate. Each firm is of the form  $(y_i, \chi_i, x_i, \bar{y}_i, \bar{x}_i)$ , in which  $i$  indexes a firm in an industry,  $y_i$  is the rate of investment of firm  $i$ ,  $\chi_i$  is its true unobserved  $q$ , and  $x_i$  is its observed  $q$ , with

$$y_i = \alpha_0 + \beta\chi_i + u_i \quad (13)$$

$$x_i = \chi_i + \varepsilon_i. \quad (14)$$

$u_i$  is a zero-mean regression disturbance, and  $\varepsilon_i$  is a zero-mean measurement error, with  $(u_i, \varepsilon_i)$  independent of one another and also of  $\chi_i$  and with the population  $R^2$  of Equation (14) equal to 0.2, which is consistent with the evidence on conglomerates in Whited (2001). We set the mean of  $y_i$  and covariance matrix of  $(x_i, y_i)$  equal to their sample counterparts in our data set. Note that because the mean of  $\varepsilon_i$  is 0, the mean of  $x_i$  is simply the mean of  $\chi_i$ . We let both  $u_i$  and  $\varepsilon_i$  have normal distributions. The distribution for  $\chi_i$  within an industry is the same as the distribution of the mean of  $\chi_i$  across industries.

<sup>11</sup> One can argue that our simulations are not relevant to an empirical literature that defines an industry by its median (not mean)  $q$ . However, because there exists a one-to-one mapping between the mean and the median of a lognormal distribution, our definition of an industry on the basis of its mean  $q$  is harmless.

These settings and our independence assumptions allow us to solve for  $\beta = 0.05$  and the population  $R^2$  of Equation (13) equal to 0.15. The settings also pin down the variances of  $u_i$ ,  $\varepsilon_i$ , and  $\chi_i$ , the latter of which is the variance of  $x_i$  times the population  $R^2$  of Equation (14). The value of  $\alpha_0$  depends on the mean of  $x_i$  and therefore varies from industry to industry. See Erickson and Whited (2000), for example, for the formulas necessary to do these calculations. Finally,  $\bar{y}_i$  is the industry median rate of investment and  $\bar{x}_i$  is the industry median observed  $q_i$ .

With these ingredients in hand, we can calculate for each set of conglomerates “observed” RINV and RVA. The observed values of RINV and RVA are calculated according to Equation (1) and (4), where  $\bar{y}_i$  corresponds to  $(I/S)_j^{ss}$  and  $\bar{x}_i$  corresponds to  $q_j$ .

### Simple matching estimator

As background for our presentation of the AI estimator, we start with a simple matching estimator. Let  $N$  be the number of observations in the sample: both the firms that have refocused and all of the potential control firms. Now consider the matching. Let  $\|z\|_V \equiv (\vec{z}' V \vec{z})^{1/2}$  be the vector norm with positive definite weight matrix  $V$ . In our application,  $V$  has the inverse of the variances of  $Z$  along the diagonal. Let  $d_M(i)$  be the distance from the covariate (control-variable) value for unit  $i$ ,  $Z_i$ , to the  $M^{\text{th}}$  nearest match. Allowing for the possibility of ties,  $d_M(i)$  is the distance such that fewer than  $M$  units are closer to unit  $i$  than  $d_M(i)$  and at least  $M$  are as close as  $d_M(i)$ . With no ties, there will be exactly  $M$  matches within the distance  $d_M(i)$  of  $Z_i$ . In our application,  $M = 1$ .

Let  $\mathcal{J}_M(i)$  denote the set of indices for the matches for unit  $i$  that are at least as close as the  $M^{\text{th}}$  match. If there are no ties, the number of elements in  $\mathcal{J}_M(i)$  is  $M$ , although in general it may be larger. Let the number of elements be denoted by  $\#\mathcal{J}_M(i)$ . Because the AI estimator allows a unit to be used more than once as a match, it is useful to define a quantity  $\mathcal{K}_M(j)$  as the number of times unit  $j$  is used as a match, given that  $M$  matches per unit are used, divided by the total number of matches. Note that  $\sum_j \mathcal{K}_M(j)$  is equal to the number of treated units.

The simple matching estimator uses the following approach to estimate the effects of the pair of potential outcomes—treated and not treated. For each unit  $i$ , the observed outcome gives us one of the two potential outcomes. For example, if the unit is treated, then we can directly observe  $\Delta F_i(1)$ , but not  $\Delta F_i(0)$ . The other, unobserved, potential outcome is estimated by averaging over its matches. For example, if the unit is treated, to estimate  $\Delta \hat{F}_i(0)$ , we need to average  $\Delta F_j(0)$  over the set of matches. More formally,

$$\Delta \hat{F}_i(0) = \begin{cases} \Delta F_i & \text{if } R_i = 0 \\ \frac{1}{\#\mathcal{J}_M(i)} \sum_{j \in \mathcal{J}_M(i)} \Delta F_j & \text{if } R_i = 1, \end{cases} \quad (15)$$

and

$$\Delta \hat{F}_i(1) = \begin{cases} \frac{1}{\#\mathcal{J}_M(i)} \sum_{j \in \mathcal{J}_M(i)} \Delta F_j & \text{if } R_i = 0 \\ \Delta F_i & \text{if } R_i = 1. \end{cases} \quad (16)$$

For every unit  $i$  in the treated sample, the expression in the second line of (15) gives the average value of  $\Delta F_j$  for the matches to unit  $i$  in the control sample. The simple matching estimator for the sample average treatment effect for the treated unit is

$$\hat{\tau} |_{R=1} = \frac{1}{N_1} \sum_{i=1}^{N_1} [\Delta F_i(1) - \Delta \hat{F}_i(0)], \quad (17)$$

in which  $N_1$  is the number of treated units. Note that averaging only over the treated units gives the sample equivalent of the conditional expectation in Equation (7).

### Abadie and Imbens estimator

The simple matching estimator is biased in finite samples when the matching is not exact. AI shows that with  $k$  continuous control variables, the estimator contains a term corresponding to the matching discrepancies (the difference in covariates between matched units and their matches) that is of the order  $O(N^{-1/k})$ . In other words, this term vanishes at a rate lower than the increase in sample size.

The bias-corrected matching estimator adjusts the difference within the matches for the differences in their covariate values. The adjustment is based on a consistent estimate of the conditional mean of  $\Delta F(R)$  given  $Z = z$ , defined precisely as  $m_R(z) = E[\Delta F(R)|Z = z]$ . Following the empirical application in AI, we approximate these conditional expectations by linear functions and estimate them using least squares. Because we are interested in estimating the sample average treated effect for the treated, we only need to estimate the regression function for the control observations:  $m_0(z)$ . This estimation is done using only the control observations that are used as matches, because the non-matched control observations do not appear in the matched sample. The population regression is

$$m_R(z) = b_0 + b_1 z, \quad (18)$$

and the coefficient estimates are given by:

$$(\hat{b}_0, \hat{b}_1) = \operatorname{argmin}_{\hat{b}: R_i=0} \sum_{i: R_i=0} \mathcal{K}_M(i) (\Delta F_i - b_0 - b_1 Z_i)^2. \quad (19)$$

As in AI, we weight the observations in the regression by  $\mathcal{K}_M(i)$ , the number of times the unit is used as a match. Let  $\hat{m}_0(z)$  be the fitted value of  $m_0(z)$  using  $(\hat{b}_0, \hat{b}_1)$  given by Equation (19). Then bias adjusted estimates of  $F_i(0)$  are given by:

$$\tilde{F}_i(0) = \begin{cases} \Delta F_i & \text{if } R_i = 0 \\ \frac{1}{\#\mathcal{J}_M(i)} \sum_{j \in \mathcal{J}_M(i)} [\Delta F_j + \hat{m}_0(Z_i) - \hat{m}_0(Z_j)] & \text{if } R_i = 1 \end{cases} \quad (20)$$

The term  $\hat{m}_r(Z_i) - \hat{m}_r(Z_j)$  in the second line of Equation (20) is the difference in the predicted values of  $\Delta F_i$ , given the vector of control variables for the  $i^{\text{th}}$  treated unit and the vector of control variables for the  $j^{\text{th}}$  match. An estimate of the average treatment effect for the treated is then obtained by using  $\tilde{F}_i(0)$  in Equation (17) instead of  $\hat{F}_i(0)$ .

We perform several Monte Carlo experiments to evaluate the finite sample performance of this estimator. We pattern our experiments after the data that are deflated by total assets. For all of our experiments, we generate covariates with sample sizes, means, and a covariance matrix equal to those of our actual data on control firms and spin-offs. We simulate only the covariates that appear significant in our probit regression reported in Table 3. We choose the distributions for the covariates, so that they have third and fourth moments approximately equal to the corresponding actual data moments.

**Table A1**  
**Monte Carlo simulation of treatment effect estimators**

	True treatment effect	0	0.001	0.002	0.005	0.01	0.02
Abadie and Imbens	Estimated treatment effect	0.000	0.001	0.002	0.005	0.010	0.020
	Probability of rejection	0.057	0.187	0.223	0.319	0.470	0.897
Dehejia and Wahba	Estimated treatment effect	0.000	0.001	0.002	0.005	0.010	0.020
	Probability of rejection	0.049	0.201	0.245	0.381	0.498	0.951
Heckman	Estimated treatment effect	0.000	0.001	0.002	0.005	0.010	0.020
	Probability of rejection	0.051	0.203	0.241	0.375	0.474	0.902

Indicated averages and probabilities are based on 10,000 Monte Carlo trials. "Probability of Rejection" indicates the rate of rejection of the null hypothesis that the treatment effect is 0. All rejection rates are calculated using an asymptotic .05 significance level critical value.

Specifically, the variable representing the log of assets has a normal distribution; the variable representing the debt-to-assets ratio has a chi-squared distribution with eight degrees of freedom; the variables representing cash flow, the financing gap, the profit shock, largest-segment profit, and industry sales growth all have normal distributions; the variables representing excess value, market share, and control activity have chi-squared distributions with two degrees of freedom; the variable representing RVA is the sum of two lognormals; the variable representing the forecast error has a chi-squared distribution with two degrees of freedom; and the variable representing entropy has a chi-squared distribution with two degrees of freedom. We omit simulated regressors representing the coefficient of variation and IPO activity, because these variables have almost identical distributions to those for the forecast error and control activity.

We use these covariates to generate simulated values for the change in RVA, using regression coefficients and the  $R^2$  from a regression of RVA on these covariates in our actual data. We also use these covariates to generate a treatment dummy, where the treated group contains the observations with the  $n$  highest propensity scores, which are estimated with probit regressions. We set  $n = 102$ , which is the actual number of spin-offs.

We fix the true treatment effect at several levels: 0, 0.001, 0.002, 0.005, 0.010, and 0.020. The first level corresponds to the null hypothesis of no treatment effect. The rest of the levels represent the range of before-and-after differences reported in Tables 3 and 4. These figures are used to determine the power of these estimators to detect a non-zero treatment effect.

The results are in Table A1, which reports the mean treatment effect estimate over 10,000 Monte Carlo trials, as well as the actual size of a one-sided nominal 5%  $t$ -test of the null hypothesis that the treatment effect is 0. The results indicate that the estimator is unbiased. Furthermore, the one-sided nominal 5%  $t$ -test for the null hypothesis of a zero treatment effect provides rejections 5.7% of the time. This figure is slightly higher than the nominal size. When the actual treatment effect is non-zero, this  $t$ -test provides an increasing percentage of rejections as true treatment effect moves away from zero. Table A1 also shows that the results from evaluating the Heckman and Dehejia and Wahba techniques are similar.

### Erickson–Whited estimators

For reference, we reproduce Equations (11) and (12) from the text

$$y_i = \alpha_0 + \beta\chi_i + \alpha_1 z_i + u_i \quad (21)$$

$$x_i = \gamma_0 + \chi_i + \varepsilon_i. \quad (22)$$

$\varepsilon_i$  is a mean-zero error independent of  $(u_i, z_i, \chi_i)$ , and  $u_i$  is independent of  $(\chi_i, z_i)$ . The intercept  $\gamma_0$  allows for the non-zero means of some sources of measurement error. The EW estimators also require the assumption that  $(\varepsilon_i, u_i, z_i, \chi_i)$ ,  $i = 1, \dots, n$ , are independently and identically distributed, that the residual from the projection of  $\chi_i$  on  $z_i$  has a skewed distribution, and that  $\beta \neq 0$ . The last two assumptions are required for estimator identification and are testable. The one questionable assumption here is the independence of  $u_i$  and  $(\chi_i, z_i)$ . Clearly, investment,  $q$ , and cash flow are determined simultaneously. However, as delineated in Erickson and Whited (2000), plausible economic assumptions do exist under which the independence assumption holds. Furthermore, because the EW estimators are based on GMM, the  $J$ -test can be used to detect assumption violations.

Let  $(\dot{y}_i, \dot{x}_i, \dot{\chi}_i)$  be the residuals from the linear projection of  $(y_i, x_i, \chi_i)$  on  $z_i$ . Then Equation (21) and (22) can be written as

$$\dot{y}_i = \beta \dot{\chi}_i + u_i \quad (23)$$

$$\dot{x}_i = \dot{\chi}_i + \varepsilon_i. \quad (24)$$

If we square Equation (23), multiply the result by Equation (24), and take unconditional expectations of both sides, we obtain

$$E(\dot{y}_i^2 \dot{x}_i) = \beta^2 E(\dot{\chi}_i^3). \quad (25)$$

Analogously, if we square Equation (24), multiply the result by Equation (23), and take unconditional expectations of both sides, we obtain

$$E(\dot{y}_i \dot{x}_i^2) = \beta E(\dot{\chi}_i^3). \quad (26)$$

As shown in Geary (1942), if  $\beta \neq 0$  and  $E(\dot{\chi}_i^3) \neq 0$ , dividing Equation (25) by Equation (26) produces a consistent estimator for  $\beta$ . Equations (25) and (26) are third-order moment equations. The innovation in Erickson and Whited (2002) consists of combining the information in moment equations of order 2 up through 7 via GMM to obtain a more efficient estimator for  $\beta$ . Note that  $\alpha_1$  can be recovered by the identity

$$\alpha_1 = \mu_y - \beta \mu_x,$$

in which  $(\mu_y, \mu_x)$  are the slope coefficients in the projection of  $(y_i, x_i)$  on  $z_i$ .

The coefficients of determination ( $R^2$ 's) for Equations (21) and (22) are calculated as

$$\rho^2 = \frac{\mu_y \text{var}(z_i) \mu_y + E(\dot{\chi}_i^2) \beta^2}{\mu_y \text{var}(z_i) \mu_y + E(\dot{\chi}_i^2) \beta^2 + E(u_i^2)} \quad (27)$$

$$\tau^2 = \frac{\mu_x \text{var}(z_i) \mu_x + E(\dot{\chi}_i^2)}{\mu_x \text{var}(z_i) \mu_x + E(\dot{\chi}_i^2) + E(\varepsilon_i^2)} \quad (28)$$

We perform Monte Carlo simulations based on a data-generating process identical to the one used in the simulations in Section 1, with the following addition. We include two perfectly measured variables that have means and a covariance matrix equal to “own-segment cash flow” and “other-segment cash flow” from our divestiture sample.

In our first simulation, we set the coefficients on these two variables equal to 0. Here, we obtain optimum finite sample performance using moments of up to order 4, where performance is measured as the mean absolute deviation over 10,000 Monte Carlo samples of the estimate of either element of  $\alpha_1$ . Also, a one-sided nominal 5% *t*-test for the null hypothesis that these two coefficients are 0 provides rejections 6.5% and 5.9% of the time, respectively. In other words, even when the true coefficient on cash flow or other segment cash flow is 0, these estimators will tend to err slightly in the direction of finding a significant positive coefficient. To examine the power of this *t*-test, we run two further simulations. In the first, we set the true coefficients on own-segment cash flow and other-segment cash flow equal to their sample averages from Table 7. In the second, we set these coefficients equal to the average estimates found in Dittmar and Shivdasani (2003). Their estimate of the coefficient on own-segment cash flow is approximately six times as large as our own, and their estimate of the coefficient on other-segment cash flow is approximately 11 times our own. In the first simulation, we find that a one-sided *t*-test provides rejections of the null hypothesis that these coefficients are zero 39.1% and 37.8% of the time, respectively. These rejection rates are comparable to those in Erickson and Whited (2000). In the second, simulation, we find that these rejection rates rise to 79.2% and 85.8%.

### References

- Abadie, A., and G. W. Imbens, 2006, "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, 75, 235–267.
- Ahn, S., and D. J. Denis, 2004, "Internal Capital Markets and Investment Policy: Evidence from Corporate Spin-offs," *Journal of Financial Economics*, 71, 489–516.
- Berger, P. G., and E. Ofek, 1995, "Diversification's Effect on Firm Value," *Journal of Financial Economics*, 37, 39–65.
- Berger, P. G., and E. Ofek, 1999, "Causes and Effects of Corporate Refocusing Programs," *Review of Financial Studies*, 12, 311–345.
- Bergh, D. D., 1997, "Predicting Divestiture of Unrelated Acquisitions: An Integrative Model of Ex Ante Conditions," *Strategic Management Journal*, 18, 715–731.
- Bernardo, A. E., and B. Chowdry, 2002, "Resources, Real Options, and Corporate Strategy," *Journal of Financial Economics*, 63, 211–234.
- Burch, T. R., and V. Nanda, 2003, "Divisional Diversity and the Conglomerate Discount: Evidence from Spin-offs," *Journal of Financial Economics*, 70, 69–98.
- Chevalier, J., 2004, "What Do We Know About Cross-Subsidization? Evidence from Merging Firms," *Advances in Economic Analysis & Policy*, 4(1), Article 3.
- Cusatis, P. J., J. A. Miles, and J. R. Woolridge, 1993, "Restructuring Through Spin-Offs: The Stock-Market Evidence," *Journal of Financial Economics*, 33, 293–311.
- Daley, L., V. Mehrotra, and R. Sivakumar, 1997, "Corporate Focus and Value Creation: Evidence from Spin-offs," *Journal of Financial Economics*, 45, 257–281.
- Dehejia, R. H., and S. Wahba, 1999, "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, 94, 1053–1062.
- Dehejia, R. H., and S. Wahba, 2002, "Propensity Score-Matching Methods for Nonexperimental Causal Studies," *Review of Economics and Statistics*, 84, 151–161.
- Desai, H., and P. C. Jain, 1999, "Firm Performance and Focus: Long-Run Stock Market Performance Following Spin-offs," *Journal of Financial Economics*, 54, 75–101.
- Dittmar, A. K., and A. Shivdasani, 2003, "Divestitures and Divisional Investment Policies," *Journal of Finance*, 58, 2711–2744.

- Erickson, T., and T. M. Whited, 2000, "Measurement Error and the Relationship Between Investment and  $q$ ," *Journal of Political Economy*, 108, 1027–1057.
- Erickson, T., and T. M. Whited, 2002, "Two-Step GMM Estimation of the Errors-in-Variables Model Using High-Order Moments," *Econometric Theory*, 18, 776–799.
- Geary, R. C., 1942, "Inherent Relations Between Random Variables," *Proceedings of the Royal Irish Academy*, 47, 63–76.
- Gertner, R., E. A. Powers, and D. S. Scharfstein, 2002, "Learning About Internal Capital Markets From Corporate Spin-offs," *Journal of Finance*, 57, 2479–2506.
- Gomes, J. F., and D. Livdan, 2004, "Optimal Diversification: Reconciling Theory and Evidence," *Journal of Finance*, 59, 507–535.
- Haynes M., S. Thompson, and M. Wright, 2003, "The Determinants of Corporate Divestment: Evidence from a Panel of UK Firms," *Journal of Economic Behavior and Organization*, 52, 147–166.
- Heckman, J., 1979, "Sample Selection Bias As Specification Error," *Econometrica*, 47, 153–161.
- Hennessy, C. A., 2004, "Tobin's Q, Debt Overhang, and Investment," *Journal of Finance*, 59, 1717–1742.
- Holtz-Eakin, D., W. Newey, and H. S. Rosen, 1988, "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 56, 1371–1395.
- Krishnaswami, S., and V. Subramaniam, 1999, "Information Asymmetry, Valuation, and the Corporate Spin-Off Decision," *Journal of Financial Economics*, 53, 73–112.
- Lamont, O. A., 1997, "Cash Flow and Investment: Evidence from Internal Capital Markets," *Journal of Finance*, 52, 83–109.
- Lang, L., A. Poulsen, and R. M. Stulz, 1995, "Asset Sales, Firm Performance, and the Agency Costs of Managerial Discretion," *Journal of Financial Economics*, 37, 3–37.
- Lang, L., and R. M. Stulz, 1994, "Tobin's  $q$ , Corporate Diversification, and Firm Performance," *Journal of Political Economy*, 102, 1248–1280.
- Maksimovic, V., and G. M. Phillips, 2000, "Optimal Firm Size and the Growth of Conglomerate and Single-Industry Firms," *Journal of Finance*, 55, 721–767.
- Maksimovic, V., and G. M. Phillips, 2001, "The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains?" *Journal of Finance*, 56, 2019–2065.
- Maksimovic, V., and G. M. Phillips, 2002, "Do Conglomerate Firms Allocate Resources Inefficiently Across Industries? Theory and Evidence," *Journal of Finance*, 57, 721–767.
- Maxwell, W. F., and R. P. Rao, 2003, "Do Spin-offs Expropriate Wealth From Bondholders?," *Journal of Finance*, 58, 2087–2108.
- Newey, W., 1984, "A Methods of Moments Interpretation of Sequential Estimators," *Economics Letters*, 14, 201–206.
- Palepu, K., 1985, "Diversification Strategy, Profit Performance and the Entropy Measure," *Strategic Management Journal*, 6, 239–255.
- Parrino, R., 1997, "Spin-offs and Wealth Transfers: The Marriott Case," *Journal of Financial Economics*, 43, 241–274.
- Rajan, R. G., H. Servaes, and L. Zingales, 2000, "The Diversification Discount and Inefficient Investment," *Journal of Finance*, 55, 35–80.
- Rosenbaum, P., 1989, "Optimal Matching in Observational Studies," *Journal of the American Statistical Association*, 84, 1024–1032.

- Rosenbaum, P., 1995, *Observational Studies*, Springer Verlag, New York.
- Rosenbaum, P. R., and D. B. Rubin, 1983, "The Central Role of Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41–55.
- Rubin, D. B., 1973a, "Matching to Remove Bias in Observational Studies," *Biometrics*, 29, 159–183.
- Rubin, D. B., 1973b, "The Use of Matched Sampling and Regression Adjustments to Remove Bias in Observational Studies," *Biometrics*, 29, 185–203.
- Scharfstein, D. S., 1998, "The Dark Side of Internal Capital Markets II: Evidence from Diversified Conglomerates," working paper, M.I.T.
- Schlingemann, F. P., R. M. Stulz, and R. A. Walkling, 2002, "Divestitures and the Liquidity of the Market for Corporate Assets," *Journal of Financial Economics*, 64, 117–144.
- Shin, H. H., and R. M. Stulz, 1998, "Are Internal Capital Markets Efficient?" *Quarterly Journal of Economics*, 113, 531–552.
- Shleifer, A., and R. W. Vishny, 1992, "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," *Journal of Finance*, 47, 1343–1365.
- Stein, J. C., 2003, "Agency, Information and Corporate Investment," in G. Constantinides, M. Harris, and R. Stulz (eds), *Handbook of the Economics of Finance*, Elsevier, Amsterdam.
- Villalonga, B., 2004a, "Diversification Discount or Premium? New Evidence from the Business Information Tracking Series," *Journal of Finance*, 59, 479–506.
- Villalonga, B., 2004b, "Does Diversification Cause the 'Diversification Discount'?" *Financial Management*, 33, 5–27.
- White, H., 1980, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817–838.
- Whited, T. M., 2001, "Is it Inefficient Investment that Causes the Diversification Discount?" *Journal of Finance*, 56, 1667–1692.