

**TECHNOLOGY SHOCKS, Q, AND THE PROPENSITY TO MERGE**

by

Lihong Han and Peter L. Rousseau



**Working Paper No. 09-W14**

September 2009

DEPARTMENT OF ECONOMICS  
VANDERBILT UNIVERSITY  
NASHVILLE, TN 37235

[www.vanderbilt.edu/econ](http://www.vanderbilt.edu/econ)

# Technology Shocks, $Q$ , and the Propensity to Merge\*

Lihong Han and Peter L. Rousseau

August 2009

## Abstract

Data on U.S. mergers and acquisitions from 1987 to 2006 indicate that firms with high market-to-book values (i.e., Tobin's  $Q$ ) tend to merge with firms that have lower  $Q$ 's, but that target  $Q$ 's are on average higher than those of firms not involved in mergers at all. We capture this fact with a model in which the ratio of a bidder's  $Q$  to that of a prospective target has a non-monotone, inverted U-shaped effect on the probability of the two firms merging. Further, we find that the likelihood of a merger is positively and linearly related to the ratio of the growth potential of an acquirer and its prospective target. Using data from *Compustat*, a series of bootstrap logit regressions bear out these implications.

## 1 Introduction

According to the Securities Data Corporation (SDC), more than 3,700 U.S. firms were involved in domestic within-industry corporate mergers between 1987 and 2006, with the total value of these transactions reaching nearly 5 trillion constant 2005 dollars. Direct transactions costs (brokerage and legal) alone averaged \$6 million per merger over the same 20-year period. And while the number of mergers has fallen from a high-water mark of more than 1,200 between 1997 and 1999, they remain central to the U.S. market for corporate control. At the same time, it is also true that the vast majority of U.S. companies did not engage in mergers at all. Given the prospective gains that investors might obtain by predicting mergers, identifying factors that influence them is a topic of considerable interest among finance practitioners.

One stylized fact about merger and acquisition (M&A) activity is that acquirers on average have higher values of Tobin's  $Q$  (defined as the ratio of a firm's market value to the reproduction cost of its capital) than their targets. Andrade, Mitchell

---

\*We thank the NSF for financial support.

and Stafford (2001), for example, report that roughly two-thirds of mergers since 1973 involve acquirers with higher  $Q$ s than their targets. Further, Servaes (1991) finds that total takeover returns (i.e., the abnormal increase in the combined values of merging parties) are larger when the bidder has a higher  $Q$  than its target. Manne (1965) attributes this pattern to purchases of low value, poorly-managed firms by better-managed ones. Jovanovic and Rousseau (2008) model these facts in a framework where higher- $Q$  acquirers pass better technologies (and managements) to their lower- $Q$  targets, which allows for rapid reallocation of the aggregate capita stock in the face of technology shocks. At the same time, such reallocations come at a cost which, if high enough, can reduce the gains from merging to a point where transactions are forgone despite large differences in  $Q$ s among the counterparties.

In this paper we take a micro-based approach to the question of how productivity differences interact with Tobin's  $Q$  to determine the nature of merger matches. Figure 1 shows that the Tobin's  $Q$ 's of acquirers, though exceeding those of their targets on average, do not usually exceed them by much more than they exceed the  $Q$ 's of firms not involved in mergers at all.<sup>1</sup> In some years, especially around the 1999 peak in merger activity, target  $Q$ 's even exceed those of non-merging firms. This suggests that a high- $Q$  firm, if involved in a merger at all, will tend to purchase a firm with a lower  $Q$  but not necessarily seek out targets the lowest  $Q$ 's of all. We explain this by distinguishing the effects of technology shocks on firm-level productivity from those on Tobin's  $Q$ . In our framework, a merger is more likely when an acquirer's  $Q$  exceeds that of a prospective target by a wide margin while the productivity difference between acquirer and target is smaller. Because of the distinct role of productivity differences in determining the propensity to merge, our model delivers a 'high buys less high' pattern in terms of Tobin's  $Q$ , and a 'high buys low' pattern in terms of relative growth potentials, which we attribute to intangibles such as organization capital, management ability, and latent ideas that accompany technology shocks.

This 'high buys less high' pattern in terms of  $Q$  has been noted by Rhodes-Kropf, Robinson and Viswanathan (2005) and Rhodes-Kropf and Viswanathan (2007) where emphasis is on similarities among merging firms (in their words a 'like buys like' pattern in terms of firm characteristics) and a desire of managers to place complementary assets under common control. We on the other hand suggest that the pattern emerges from the firm specificity of capital and the nature of costs associated with converting a target's capital into a form usable by the acquirer.

Our model makes three key assumptions. First, there are positive technology shocks that affect one group of firms more than the others, and that two kinds of firms will coexist after a shock. Firms that receive the shock will see an increase in both total factor productivity (TFP) and Tobin's  $Q$ , though the extent to which  $Q$  will rise differs across firms. Second, we assume that mergers are a channel through which

---

<sup>1</sup>The annual averages presented in Figure 1 include all U.S. firms listed on Standard & Poor's *Compustat* database. We identify mergers among these firms using lists of mergers from the SDC.

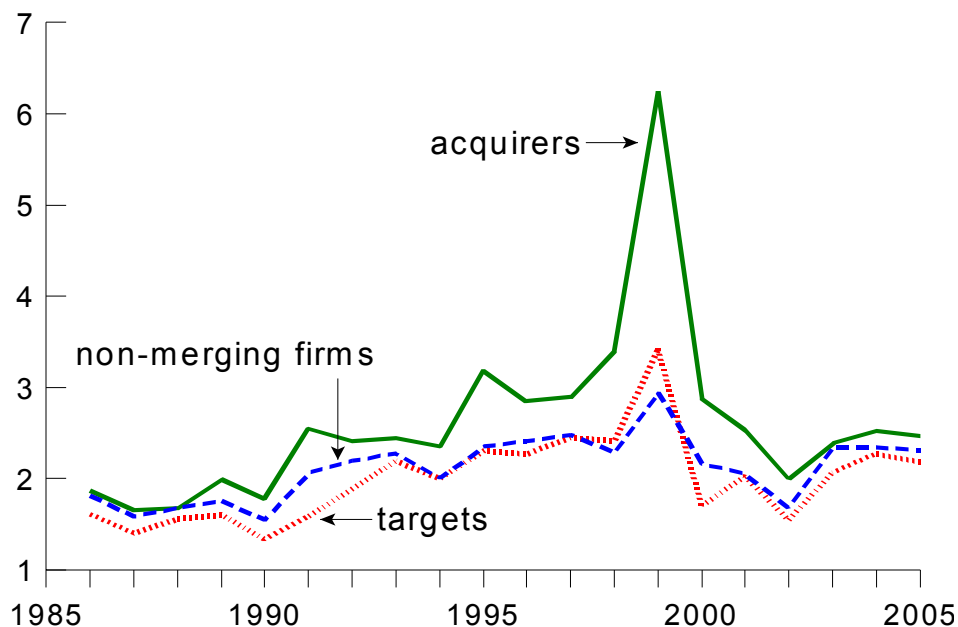


Figure 1: Tobin's  $Q$  for subgroups of Compustat firms, 1986-2005

capital flows from firms with lower growth potential (defined as the ratio of Tobin's  $Q$  to TFP) to those with higher growth potential. Finally, we assume that conversion costs are a convex function of the productivity difference between an acquirer and its potential target. Costs are highest when the acquirer is much more productive than the target because the transformation of the target's low productivity assets for use by the acquirer suffers from a compatibility problem. These costs could take the forms of employee retraining and the refitting of plant and equipment. Costs decline as the productivity difference decreases because the target's higher productivity assets can be more easily integrated as a bundle into the acquiring firm. The counterparties negotiate over how to divide the surplus generated by the merger and if both parties can make greater profits under common control than they can separately, they will merge. Specifically, our model illustrates that:

(1) The ratio of a bidder's  $Q$  to that of a potential target has an inverted U-shaped effect on the probability of the two firms merging.

(2) The likelihood of a merger is positively and linearly related to the difference in the relative growth potentials of an acquirer and its prospective target. The measure of growth potential is the ratio of a firm's Tobin's  $Q$  to its total factor productivity. Our finding that deflation by TFP linearizes the relationship between  $Q$  and the propensity to merge suggests that growth potential is a more potent variable than  $Q$  in characterizing M&A activity.

Our data include all within-industry U.S. domestic mergers reported in the SDC

files from 1986 to 2005 where both acquirer and target are companies listed on U.S. stock exchanges. This gives us a dataset with 1,317 actual merger pairs among 3,051,796 potential ones. Using the data, we develop a series of quasi-bootstrap logit regressions for the probability of an actual merger on the ratios of bidder to target  $Q$ s, productivity differences, and relative growth potentials. Our empirical results bear out the main implications of the model.

The remainder of this paper is organized as follows. In Section 2 we construct a Nash bargaining model that incorporates a cost of refurbishing low-productivity capital. In Section 3 we describe our dataset and econometric approach, and then make an empirical assessment of the model. We draw our conclusions together in Section 4.

## 2 Model

The model builds upon Gort (1969) and Jovanovic and Rousseau (2002, 2008). Some firms in the economy are better positioned to take advantage of a technology shock than others and thus become more productive after the shock. We model this as two technologies that coexist after a shock but are embedded exclusively in the capital of individual firms. Thus capital is technology-specific as in Hulten (1992) and Greenwood et al. (1997, 2000). Given that high- and the low-technology firms face the same output price and high-technology firms make better use of the assets they control, they also have higher  $Q$ s than low-technology firms. As a result, the high-technology firms seek to acquire low technology-firms and convert their low-technology capital to high-technology capital, even though this conversion entails costs. At the same time, a positive shock in an industry increases the opportunity cost of operating as an inefficient producer in that industry, thereby altering the values of assets and creating incentives for transfers to more productive users through M&A.

All firms are assumed to be price-takers, to produce a homogenous output, to be endowed with technology-specific production assets, and to have the same technology (i.e., productivity and Tobin's  $Q$ ) initially. At time 0, a positive technology shock arrives that affects some firms more than others. Hence, after the shock, there are two distinct technologies, with each associated with a distinct set of assets. The productivity parameters associated with the high and low technologies are represented by  $z_h$  and  $z_l$ , with  $z_h > z_l$ , while their respective assets are denoted as  $K_h$  and  $K_l$ . Since the technology-specific assets  $K_i$  can be directly used only by firms with technology  $z_i$ , converting low-technology assets to high-technology uses comes at some cost. The cost is assumed to be a convex function of the ratio of the productivity of the high-technology firm (which becomes a potential acquirer) to the low-technology firm (which becomes a potential target). In other words, the greater the productivity ratio, the higher the cost incurred to convert capital associated with the low technology. After merging, the combined firm takes its technology and intangibles from the high

productivity firm so that the combined firm has the same  $Q$  as the acquirer did. The acquirer and target then arrive at a Nash bargaining solution to share the rents.

In our model, a firm's value is written as

$$V_i = q_i K_i, \quad i = h \text{ or } l, \quad (1)$$

where  $V_i$  is the market value of firm  $i$  and  $q_i$  is its Tobin's  $Q$ , which in turn assigns a market value to a given reproduction cost of the firm.

After merging, the combined firm's market value becomes

$$V_M = (1 - C)q_h K_l + q_h K_h, \quad (2)$$

where  $V_M$  is the market value of the combined firm and  $C(\frac{z_h}{z_l})$  is the per unit cost of converting  $K_l$  into  $K_h$  that satisfies  $C' > 0$ ,  $C'' > 0$ ,  $C(1) = C'(1) = 0$  and  $\lim_{\frac{z_h}{z_l} \rightarrow +\infty} C(\frac{z_h}{z_l}) = +\infty$ .

We use a model of negotiations to determine how the partners share the surplus generated by a merger. There are different choices for this, but the simplest is the Nash bargaining solution, which solves

$$W(V_h, V_l) = \max_{\Pi_{hM}, \Pi_{lM}} (\Pi_{hM} - V_h)^\sigma (\Pi_{lM} - V_l)^{(1-\sigma)} \quad (3)$$

$$s.t. \quad \Pi_{hM} + \Pi_{lM} = V_M,$$

where  $W$  is the joint welfare of the acquirer and its target,  $\Pi_{iM}$  is the portion of the combined entity that goes to firm  $i$ , and  $\sigma \in [0, 1]$  represents the acquirer's bargaining power in the transaction.

**Lemma 1** *In equilibrium the shares of the surplus when a high-technology firm merges with a low-technology firm are*

$$\Pi_{hM} = \sigma(V_M - V_h - V_l) + V_h, \quad (4)$$

$$\Pi_{lM} = (1 - \sigma)(V_M - V_h - V_l) + V_l. \quad (5)$$

**Proposition 2** *Assume that at time 0 the firms affected by a positive technology shock adopt the high technology immediately while the others use the low technology. If the combined market value of the total assets of a high-technology firm and a low-technology firm is higher under common control than it is separately, they will merge immediately.*

**Proof.** If the high-technology and low-technology firms merge at time  $s$ , then firm  $h$ 's market value at time 0 is

$$\begin{aligned}
V_{hs} &= \int_0^s (e^{-rt} r V_h) dt + e^{-rs} \Pi_{hM} \\
&= V_h - e^{-rs} V_h + e^{-rs} [\sigma(V_M - V_h - V_l) + V_h] \\
&= V_h + e^{-rs} \sigma(V_M - V_h - V_l) \text{ (since } \sigma(V_M - V_h - V_l) > 0) \\
&< V_h + \sigma(V_M - V_h - V_l) = V_{h0} \text{ (when } s=0).
\end{aligned}$$

Therefore,  $V_{hs} < V_{h0}$ .

With the same logic, we can prove that  $V_{ls} < V_{l0}$ .

Since both firms are worth more by merging at time 0 rather than waiting until a later time  $s$ , they merge at time 0. ■

The proposition suggests that mergers occur in waves that are driven by technological shocks. Faria (2003) and Harford (2005) offer empirical support for this view.

If the high technology and the low technology firms merge at time 0, the gain for each firm from the merger is<sup>2</sup>

$$\begin{aligned}
G_h &= \Pi_{hM} - V_h = \sigma(V_M - V_h - V_l) + V_h - V_h = \sigma(V_M - V_h - V_l), \\
G_l &= \Pi_{lM} - V_l = (1 - \sigma)(V_M - V_h - V_l) + V_l - V_l = (1 - \sigma)(V_M - V_h - V_l).
\end{aligned}$$

The two firms will merge if and only if  $G_i > 0$ , which is equivalent to  $V_M - V_h - V_l > 0$ , or

$$G_h + G_l = (1 - C)q_h K_l + q_h K_h - q_h K_h - q_l K_l > 0.$$

Rearranging, the condition can be written as

$$\frac{q_h}{q_l} > \frac{1}{1 - C\left(\frac{z_h}{z_l}\right)}. \quad (6)$$

However, since  $C$  is a convex function of  $\frac{z_h}{z_l}$ ,  $z_i$  is a productivity parameter, and  $q_i$  is a function of  $z_i$ , we cannot use this inequality directly to predict whether a merger will occur. This is because TFP reflects a firm's current performance, while Tobin's  $Q$  reflects both current and expected future performance. In other words,  $Q$  not only measures TFP, but also the value of latent intangibles.<sup>3</sup> If there were no such intangibles,  $Q$  and TFP would be proportional, which is what Jovanovic and

---

<sup>2</sup>We suppress the time indices here since mergers occur immediately after the technology shock.

<sup>3</sup>Griliches (1981) and Cockburn & Griliches (1988) report that there is a significant relation between the market value of a firm and its unanticipated intangible capital.

Rousseau (2002) implicitly assume. In our formulation, however, the high-technology firm is better equipped to adopt a new technology than the low-technology firm, meaning that it possesses a greater latent ability to do so (e.g., perhaps more flexible management) and therefore has a higher growth potential. Since  $Q$  includes this latent ability while TFP does not,  $Q$  and TFP are not proportional even though they are positively correlated.<sup>4</sup> To capture this growth potential or latent ability to adopt new technologies, we introduce the variable  $\lambda_i$ , which is a firm's  $Q$  deflated by its TFP:

$$\lambda_i = \frac{q_i}{z_i}, \quad i = h \text{ or } l. \quad (7)$$

A higher  $\lambda_i$  implies a higher growth potential. Next we define  $\lambda = \frac{\lambda_h}{\lambda_l}$  as the relative growth potentials of two firms. When the distance between  $\lambda_h$  and  $\lambda_l$  is large,  $\lambda$  is also large.

Finally, we define the ratio of a bidder's  $Q$  to that of a potential target,  $\frac{q_h}{q_l}$ , as  $q_r$ , with  $q_r \geq 1$ , and the ratio of bidder's productivity to that of its potential target,  $\frac{z_h}{z_l}$ , as  $z_r$ , with  $z_r \geq 1$ . The total surplus from the mergers,  $G_h + G_l$ , normalized by the value of low technology firm, can then be written as

$$g(q_r) = \frac{G_h + G_l}{K_l \times q_l} = q_r(1 - C) - 1 > 0. \quad (8)$$

Thus  $g(q_r)$  measures the gain from merging as a share of the low technology firm's pre-merger value.

When  $g(q_r) > 0$ , the two firms will merge. When  $g(q_r) \leq 0$ , the two firms will not merge regardless of the ratio of their  $Q$ s.

**Proposition 3** *Given  $q_r \geq 1$  and  $\lambda > 0$ : The ratio of the potential acquirer and target  $Q$ s has an inverted U-shaped effect on the probability of a merger.*

**Proof.** Since  $\frac{z_h}{z_l} \geq 1$  and  $\frac{z_h}{z_l} = \frac{q_r}{\lambda}$ ,  $C$  can be written as a function of  $\frac{q_r}{\lambda}$  and  $q_r \geq \lambda$ .

The two firms will merge if and only if  $g(q_r) > 0$ . We can calculate

$$g(1) = 0, \quad (9)$$

$$g(\lambda) = \lambda - 1, \quad (10)$$

$$g'(q_r) = 1 - C - \frac{q_r}{\lambda} C', \quad (11)$$

$$g'(\lambda) = 1 > 0, \quad (12)$$

$$g''(q_r) = -2\frac{1}{\lambda}C' - \frac{q_r}{\lambda^2}C''. \quad (13)$$

---

<sup>4</sup>Dwyer (2001) shows that the plant-level productivity and the market value of a firm are positively related and that a manufacturing technique with high productivity acts as an intangible asset for the firm that owns it.



Since  $C' > 0$  and  $C'' > 0$  when  $q_r > \lambda$ , the condition  $g''(q_r) < 0$  also holds. Consequently,  $g(q_r)$  is a concave function of  $q_r$ .

When  $q_r \rightarrow +\infty$ ,  $C(\frac{q_r}{\lambda}) \rightarrow +\infty$  given  $\lambda$ . Therefore  $g(q_r) \rightarrow -\infty$  and  $g'(q_r) \rightarrow -\infty$  when  $q_r \rightarrow +\infty$ .

Since  $g'(\lambda) > 0$ , and since  $g'(q_r) \rightarrow -\infty$  when  $q_r \rightarrow +\infty$  and  $g''(q_r) < 0$ , there exists a  $q_r^*$ , such that  $g'(q_r^*) = 0$ . Further, for any  $q_r \in [\lambda, q_r^*)$  we have  $g'(q_r) > 0$ , and for any  $q_r \in (q_r^*, +\infty)$  we have  $g'(q_r) < 0$ . Hence  $g(q_r^*) = \max_{q_r \in [\lambda, +\infty)} g(q_r)$ . When  $g(q_r^*) \leq 0$ , no merger occurs. When  $g(q_r^*) > 0$ , the ratio of the potential acquirer and target  $Q$ s has an inverted U-shaped effect on the probability of a merger, since  $g''(q_r) < 0$ . ■

For example, when  $C = c(\frac{z_h}{z_l} - 1)^2$  we can solve for  $q_r$  as follows:

$$q_r \in \left( \frac{2\lambda}{3} - \frac{\lambda}{3} \sqrt{1 + \frac{3}{c}}, \frac{2\lambda}{3} + \frac{\lambda}{3} \sqrt{1 + \frac{3}{c}} \right) \iff g'(q_r) > 0. \quad (14)$$

Expression 14 implies that  $g(q_r)$  increases until  $q_r^* = \frac{2\lambda}{3} + \frac{\lambda}{3} \sqrt{1 + \frac{3}{c}}$ , which corresponds to the maximum value of  $g(q_r)$ .

Figure 2 illustrates the gains from merger as a function of  $q_r$  using the example in 14 for various choices of  $\lambda$  and  $c$ . The solid line shows that the ratio of the potential acquirer and target  $Q$ 's has an inverted U-shaped effect on the surplus. The dashed line shows that both the gains and the gain-maximizing value of  $q_r$  falls as the growth potentials of the two firms get closer (i.e., as  $\lambda$  falls), while the dotted line shows a similar effect when costs increase.

If  $g(q_r^*) < 0$  (i.e., it lies beneath the horizontal axis in Figure 2), no merger occurs regardless of the ratio of the two firms'  $Q$ s. For instance, if  $c$  is too large, the condition  $g(q_r^*) < 0$  will hold for all values of  $q_r$ .

If  $g(q_r^*) > 0$ , a merger occurs. There are two cases:

$$\text{Case 1: } \frac{2\lambda}{3} - \frac{\lambda}{3} \sqrt{1 + \frac{3}{c}} \leq 1 < q_r^*.$$

For any  $q_r \in [1, q_r^*)$ ,  $g'(q_r) > 0$ . This means that  $g(q_r)$  is increasing in the range of  $[1, q_r^*)$ .

For any  $q_r \in [q_r^*, +\infty)$ ,  $g'(q_r) < 0$  and  $\lim_{q_r \rightarrow +\infty} g(q_r) = -\infty$ . This means that  $g(q_r)$  is decreasing in the range of  $[q_r^*, +\infty)$ . Hence, there exists a  $q_r^{**}$  such that  $g(q_r^{**}) = 0$  and  $g(q_r)$  will be negative when  $q_r > q_r^{**}$ . We thus have a range  $(1, q_r^{**})$  such that for any  $q_r \in (1, q_r^{**})$ ,  $g(q_r) > 0$  also holds, and the two firms merge.

$$\text{Case 2: } 1 < \frac{2\lambda}{3} - \frac{\lambda}{3} \sqrt{1 + \frac{3}{c\lambda}} < q_r^*.$$

Since  $g(1) = 0$  and  $g'(q_r) < 0$  for any  $q_r \in \left(1, \frac{2\lambda}{3} - \frac{\lambda}{3} \sqrt{1 + \frac{3}{c}}\right)$ ,  $g\left(\frac{2\lambda}{3} - \frac{\lambda}{3} \sqrt{1 + \frac{3}{c}}\right) < 0$ . Using the same reasoning as in case 1, we can show there exists  $q_r^{21}$  and  $q_r^{22}$  such

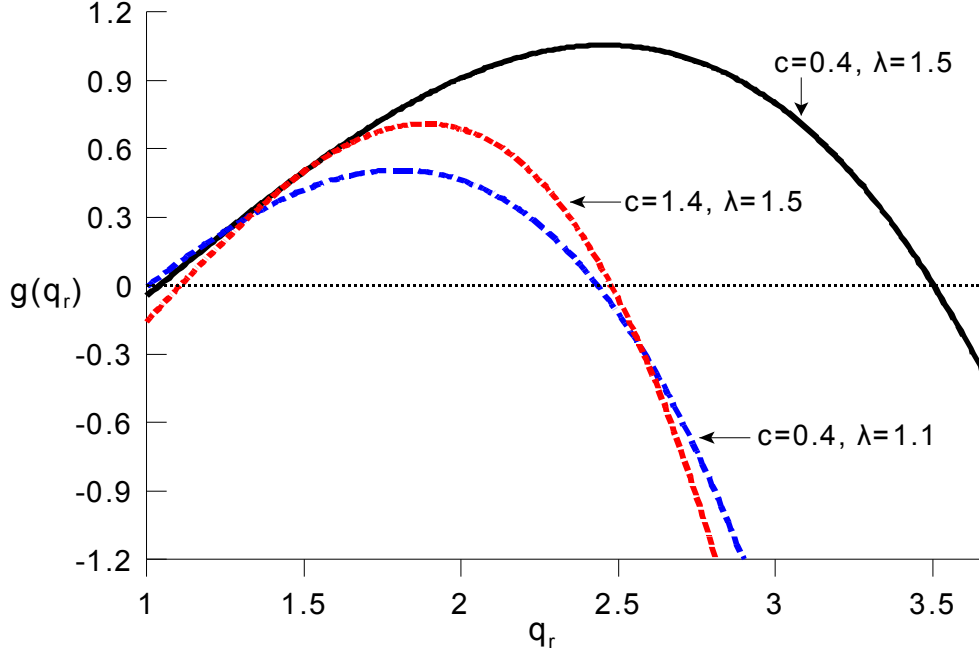


Figure 2: The gains from merging for selected values of  $c$  and  $\lambda$ .

that  $q_r^{21} \in \left(\frac{2\lambda}{3} - \frac{\lambda}{3}\sqrt{1 + \frac{3}{c}}, q_r^*\right)$ ,  $q_r^{22} \in (q_r^*, \infty)$ ,  $g(q_r^{21}) = 0$ , and  $g(q_r^{22}) = 0$ . Then, for any valid  $q_r \in (q_r^{21}, q_r^{22})$ , it follows that  $g(q_r) > 0$  and the two firms merge.

Proposition 3 demonstrates that the ratio of the acquirer to target  $Q$  is non-monotonically related to the likelihood of a merger. This inverted U-shape augments the  $Q$ -theory of mergers, which suggests high market-to-book firms simply acquire those with low market-to-book values. Thus, in our model, an acquirer may not purchase the lowest  $Q$  firm that it can find, but rather a firm with a lower  $Q$ . The model also implies that the probability of being involved in a merger depends on  $\lambda$ . A firm with high growth potential (perhaps due to a greater ability to adopt new technologies) is likely to acquire other firms. A firm with lower growth potential (perhaps due to less ability to adopt new technologies) will be in an inferior position in future competition and more likely to be acquired by high productivity firms. Thus, when the ratio of the growth potential of acquirer to target is high, the probability of a merger is high. The following proposition states how mergers are affected by the ratio of the potential acquirer to target firms' growth potential:

**Proposition 4** *Given  $q_r \geq 1$ , and  $\lambda > 0$  : the probability of a merger is positively and linearly related to  $\lambda$ .*

**Proof.** The maximum value of function  $g$  can be simplified as follows.

Define  $G = \max g(q_r) = q_r^* - q_r^* C(\frac{q_r^*}{\lambda}) - 1$ . A larger value of  $G$  implies a larger likelihood of a merger, since there is a greater probability that  $g(q_r)$  exceeds zero.

The effect of  $\lambda$  on  $G$  is

$$\frac{dG}{d\lambda} = \frac{\partial G}{\partial q_r^*} \frac{dq_r^*}{d\lambda} + \frac{(q_r^*)^2}{\lambda^2} C', \quad (15)$$

$$\begin{aligned} \frac{d^2 G}{d\lambda^2} &= \frac{\partial G}{\partial q_r^*} \frac{d^2 q_r^*}{d\lambda^2} + \frac{\partial^2 G}{\partial (q_r^*)^2} \left( \frac{dq_r^*}{d\lambda} \right)^2 + \frac{2q_r^*}{\lambda^2} C' \frac{dq_r^*}{d\lambda} + \frac{(q_r^*)^2}{\lambda^3} C'' \frac{dq_r^*}{d\lambda} \\ &\quad + \frac{2q_r^*}{\lambda^2} C' \frac{dq_r^*}{d\lambda} + \frac{(q_r^*)^2}{\lambda^3} C'' \frac{dq_r^*}{d\lambda} - \frac{2(q_r^*)^2}{\lambda^3} C' - \frac{(q_r^*)^3}{\lambda^4} C'' \\ &= \frac{\partial G}{\partial q_r^*} \frac{d^2 q_r^*}{d\lambda^2} + \frac{\partial^2 G}{\partial (q_r^*)^2} \left( \frac{dq_r^*}{d\lambda} \right)^2 - 2 \frac{\partial^2 G}{\partial (q_r^*)^2} \frac{q_r^*}{\lambda} \frac{dq_r^*}{d\lambda} + \frac{\partial^2 G}{\partial (q_r^*)^2} \left( \frac{q_r^*}{\lambda} \right)^2, \end{aligned} \quad (16)$$

since

$$\begin{aligned} \frac{\partial G(q_r^*)}{\partial q_r^*} &= 1 - C - \frac{q_r^*}{\lambda} C', \\ \frac{\partial^2 G}{\partial (q_r^*)^2} &= -2 \frac{1}{\lambda} C' - \frac{q_r^*}{\lambda^2} C''. \end{aligned}$$

By the envelope theorem,  $\frac{dG}{d\lambda} = \frac{(q_r^*)^2}{\lambda^2} C' > 0$ . And since

$$g'(q_r^*) = 1 - C \left( \frac{q_r^*}{\lambda} \right) - \frac{q_r^*}{\lambda} C' \left( \frac{q_r^*}{\lambda} \right) = 0,$$

we have

$$\begin{aligned} \frac{dg'(q_r^*)}{d\lambda} &= 0 = \left( -\frac{1}{\lambda} C' - \frac{q_r^*}{\lambda^2} C'' \right) \left( \frac{dq_r^*}{d\lambda} - \frac{q_r^*}{\lambda} \right) \\ &= \frac{\partial^2 G}{\partial (q_r^*)^2} \left( \frac{dq_r^*}{d\lambda} - \frac{q_r^*}{\lambda} \right). \end{aligned}$$

Since  $\frac{\partial^2 G}{\partial (q_r^*)^2} < 0$ ,  $\frac{dq_r^*}{d\lambda} = \frac{q_r^*}{\lambda}$  holds. Substituting  $\frac{dq_r^*}{d\lambda} = \frac{q_r^*}{\lambda}$  into 16, we get  $\frac{d^2 G}{d\lambda^2} = 0$ .

Since  $\frac{\partial G}{\partial \lambda} > 0$  and  $\frac{d^2 G}{d\lambda^2} = 0$ ,  $G$  is a linear and increasing function of  $\lambda$ . This means that the likelihood of a merger rises with  $\lambda$ . ■

Proposition 4 highlights the factors that affect mergers. Differences in growth potentials have a positive effect on the propensity to merge. Propositions 3 and 4 show that while  $Q$  is not a linear factor affecting mergers,  $\lambda$  is. Because the conversion cost, a convex function of the relative productivities of the merging firms, drives the high productivity firm to purchase a firm with lower but not the lowest productivity, and because  $Q$  positively relates to TFP, a high- $Q$  firm buys a firm with a lower yet

not the lowest  $Q$ . At the same time,  $\lambda$ , which is  $Q$  deflated by TFP, is positively and linearly related to the likelihood of a merger. Consequently,  $\lambda$  is a more potent variable than  $Q$  in characterizing the factors that drive merger activity.

In light of our theoretical results on who merges with whom and which factors affect mergers, we next test whether the data on U.S. mergers over the past two decades are consistent with the theory.

### 3 Empirics

In this section we begin by checking the basic model assumptions and their implications for the data. Next, we investigate the relation between  $Q$  and  $z$  with regression analysis. Then, using a set of quasi-bootstrap logit models, we test whether the propensity to merge relates to  $q_r$ ,  $z_r$ , and  $\lambda$  in a manner that is consistent with the model.

#### 3.1 Data

We restrict the empirics to data on domestic mergers available from Thompson’s *SDC Mergers and Acquisitions Database* for exchange-listed U.S. firms. To compare characteristics of firms that actually merge with those that could have but did not, we build a dataset in five steps:

1. Obtain year-end data from 1986 to 2005 for the 22,888 firms listed on the 2006 version of Standard and Poor’s *Compustat* database, excluding firms with less than two years of balance sheet data, and classify them according to 12 broad industry groupings defined by Fama and French.<sup>5</sup> We then construct Tobin’s  $Q$  as the market-to-book ratio of a firm’s outstanding financial securities, proxy for TFP with a firm’s sales-to-assets ratio, and store its total assets as a measure of size for later use.<sup>6</sup> We refer to this dataset of listed firms and their accounting data as the “ $Q$ ” file.

---

<sup>5</sup>The 4-digit SIC codes upon which Fama and French base their 12-industries is available at: [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

<sup>6</sup>Following Jovanovic and Rousseau (2008), we measure the numerator of  $Q$  as the the value of a firm’s common equity at current share prices (the product of items 24 and 25), to which we add the book values of preferred stock (item 130) and short- and long-term debt (items 34 and 9). We use book values of preferred stock and debt in the numerator because prices of preferred stock are not available on Compustat and we do not have information on issue dates for debt from which we might better estimate market value. We note that book values of these components are reasonable approximations of market values if interest rates do not vary too much. We compute the denominator of  $Q$  in the same way except that we use the book value of common equity (item 60) rather than its market value. Our micro-based measures of Tobin’s  $Q$  therefore focus primarily on the value of a firm’s outstanding securities, and implicitly assume that the proceeds from these issues are fully applied to the formation of capital, both physical, human, and intangible. We eliminated  $Q$ ’s for

2. Obtain a list of announcements of U.S. domestic mergers from the SDC database covering the period from 1987 through the end of 2006, excluding repurchases and leveraged buyouts. To avoid double-counting multiple announcements of the same merger, we work with only one observation per calendar year for each acquirer-target pair. This delivers 58,576 acquirer-target pairs.
3. Attempt to match each target firm with its Compustat data in the “Q” file from the end of the fiscal year preceding the merger announcement. Compustat data were recovered for 5,847 of the 58,576 targets in this way. The recovery rate is only 10 percent because the SDC files contain information on merger announcements for both listed and unlisted firms while the vast majority of Compustat firms are exchange-listed. We refer to the resulting dataset as the “SDC targets”. At the same time, we flag all SDC acquirers that can be found in the “Q” file. Out of the 5,847 targets that we could find in Compustat, we were able to match them with 17,635 acquirers.<sup>7</sup>
4. Construct a set of Cartesian products for the “Q” file (with acquirers flagged) and the “SDC targets” file (from step 3), for each of 12 Fama-French industries in each of our 20 years. By stacking all of these products we end up with a database of all potential within-industry merger pairs – 3,051,796 in total for the 20-year sample.
5. Identify whether each pair represents an actual within-industry merger announcement (i.e., what we will henceforth call an “actual merger”) or simply a pseudo-merger (i.e., a pair of firms that could have announced a merger but did not). The final dataset includes 1,317 actual within-industry mergers.

In other words, our final data set contains each observation from the SDC target file paired with every observation in the same year and industry from the Compustat file. We consider only intra-industry mergers because our model applies to cases where a target’s technology is substituted with a better one, and if two firms are not in the same industry, their assets are more likely to be complements than substitutes. This choice also keeps the empirics manageable.

---

firms with negative values for net common equity since they imply negative market-to-book ratios, as well as observations with market-to-book ratios in excess of 100 since many of these are likely to be serious data errors. The sales to asset ratio (*Compustat* item 12 divided by item 6) serves as a proxy for TFP because individual firm output is not available on *Compustat* to form the numerator as implied by the basic *AK* model. Firm size is measured by total assets (*Compustat* item 6).

<sup>7</sup>The number of matched acquirers exceeds the number of targets because mergers are often announced but not completed due to competitive bids or antitrust concerns. For our purposes, a merger announcement is as relevant as an actual merger because it suggests that the firms’ analysts and shareholders considered it a potentially synergistic transaction.

### 3.2 Basic Implications of the Model for Data

Using our sample of 1,317 actual mergers, we first investigate whether acquirers typically buy firms with lower market-to-book ratios, while at the same time avoiding firms with the lowest market-to-book ratios in their sector. The left panel in Table 1 reports the average  $q_r$  (i.e., ratio of acquirer to target  $Q$ ) for the actual mergers and pseudo-mergers within each industry. All eleven averages listed in column 1 exceed unity, which is consistent with the basic  $Q$ -theory of mergers. On the other hand, the t-statistics in column 3 for the null hypothesis that actual mergers have the same  $q_r$  as the pseudo-mergers are all negative and are statistically significant for nine of the sectors. This indicates that on average the relative  $Q$ s of actual merging pairs are lower than those of the pseudo-pairs. That is to say, a high- $Q$  firm purchases a firm with a lower but not the lowest  $Q$ .

Table 1. Summary statistics for  $q_r$ ,  $z_r$ , and  $\lambda$  in 11 Fama-French sectors.

industry	$q_r$			$z_r$			$\lambda$		
	actual	pseudo	t(M-N)	actual	pseudo	t(M-N)	actual	pseudo	t(M-N)
	merge	merge		merge	merge		merge	merge	
	mean	mean		mean	mean		mean	mean	
all industries	1.093	2.313	-1.61	1.584	1.680	-1.32	1.448	1.089	7.95***
cons nondrbls	1.661	2.260	-2.77***	1.023	1.331	-2.11**	1.512	1.117	2.45**
cons durables	1.719	1.916	-0.74	0.891	1.149	-1.39	2.034	1.226	2.55**
manufacturing	1.554	1.925	-2.58**	1.010	1.260	-2.63***	1.418	1.098	3.28***
energy	1.253	1.679	-3.44***	1.241	1.785	-1.70*	1.082	1.389	-1.10
chemicals	1.401	2.115	-1.32	1.035	1.199	-1.05	1.327	1.116	1.00
computers, etc.	2.081	3.347	-5.44***	1.262	2.007	-1.85**	1.674	1.060	6.28***
telephone/TV	1.126	2.163	-4.33***	1.120	4.763	-0.79	1.227	0.913	1.60
utilities	1.096	1.215	-2.36**	1.012	1.247	-1.47	1.188	1.172	0.11
wholesale	1.656	2.376	-3.48***	1.023	1.410	-2.54**	1.599	1.174	2.97***
finance	1.145	1.841	-6.29***	0.972	3.766	-1.31	1.275	1.117	1.74*
other	1.465	2.596	-3.40***	1.033	4.335	-0.53	1.368	1.007	2.30**

Note: Columns labeled t(M-N) report t statistics for differences in means of  $q_r$ ,  $z_r$ , and  $\lambda_r$  across groups. \*, \*\*, and \*\*\* denote statistical significance at the 10, 5 and 1 percent levels respectively..

The center panel of Table 1 reports the same statistics for  $z_r$  (i.e., the ratio of acquirer to target TFP). In this case the  $z$ 's of the actual acquirers exceed those of their targets in nine of eleven industries, suggesting that on average the acquirer is also more efficient than its target. At the same time, the average  $z_r$  associated with pseudo-mergers exceeds that of actual mergers in all sectors, and these differences are statistically significant for 5 of the 11 industries. This means that on average the

Table 2. Pooled regressions of  $q_{i,t}$  on  $z_{i,t}$ , 1986-2005.

	OLS(1)	OLS(2)	OLS(3)	IV(1)	IV(2)
constant	0.565*** (0.002)	0.519*** (0.006)	0.643*** (0.011)	0.445*** (0.006)	0.581*** (0.011)
$\log(z_{i,t})$	0.076*** (0.002)	0.072*** (0.003)	0.077*** (0.003)	0.101*** (0.004)	0.111*** (0.004)
industry effect		Yes	Yes	Yes	Yes
year effect			Yes		Yes
no. obs.	86,863	86,863	86,863	40,573	40,573
R <sup>2</sup>	0.151	0.114	0.153	0.100	0.123

Note: The dependent variable is  $\log(q_{i,t})$ . The IV regressions use five lags of  $z_i$  as instruments. Standard errors are in parentheses. \*\*\* denotes statistical significance at the 1 percent level

productivity difference between the actual merging pairs is less than that of the pseudo-merging pairs, which again is consistent with our model.

The statistics in the right panel of Table 1 support the model's main implication: the mean  $\lambda$  (i.e., the ratio of acquirer  $q/z$  to target  $q/z$ ) for actual mergers exceeds unity in all 11 sectors, and exceeds the mean  $\lambda$  for non-merging pairs in 10 of 11 sectors with the differences statistically significant for 7 of them. The reversal of the signs of the t-statistics for  $\lambda$  from those obtained for  $q_r$  and  $z_r$  suggests that our deflation of  $q$  with TFP isolates the component of a firm's value (i.e., intangibles) that really provides it with the motivation to seek out a low- $\lambda$  target. In other words, the relationship between mergers and  $\lambda$  is more nearly linear than that between mergers and  $q$  or  $z$ .

Table 2 reports results from regressions that consider the relation between  $q_{i,t}$  and  $z_{i,t}$  using data pooled across industries from 1986 to 2005. Column 1 reports the baseline OLS regression. The estimated coefficient of  $\log z_{i,t}$  is positive and statistically significant at the one percent level. Columns 2 and 3 show that this result is robust to the inclusion of fixed effects for industries and time. Given that omitted variables may be affecting both  $q_{i,t}$  and  $z_{i,t}$  contemporaneously and that the  $z_{i,t}$  variable is strongly first-order autocorrelated ( $\rho = .88$ ), the two-stage least squares regressions in columns 4 and 5 aim to reduce the effect of endogeneity by including five annual lags of  $z_i$  as instruments. In both IV regressions the correlation coefficient between

$\log q_{i,t}$  and  $\log z_{i,t}$  is about 10 percent – only slightly larger than obtained with OLS. Thus, Table 2 indicates that  $Q$  and TFP are positively correlated, which implies that current productivity may implicitly affect beliefs about a firm’s future prospects. At the same time, the correlation between  $q_{i,t}$  and  $z_{i,t}$  is far below unity and therefore cannot be treated as equivalent indicators.

### 3.3 $q_r$ , $z_r$ , $\lambda$ , and the Propensity to Merge

We use a series of logit regressions to examine the factors that influence the propensity for any two firms in the same sector to merge in any calendar year. Because our panel is unbalanced (pseudo mergers exceed actual ones by a factor of more than 2300), we develop a quasi-bootstrap procedure to obtain convergent and efficient estimates. This procedure involves four steps:

1. To balance the panel of 1,317 actual mergers, we randomly select 1,317 pseudo-mergers with replacement from our sample of non-merging pairs, making sure that the number of actual and pseudo-mergers from each industry are the same (as in Rhodes-Kropf and Robinson, 2007).
2. Combining the actual and pseudo-merger pairs into a single dataset, execute a logit regression of the binary merger variable on the explanatory variables of interest.
3. Repeat steps 1 and 2 for 100 times and report the mean of the estimated coefficients.
4. Correct our estimates of the constant terms from Step 3 and use the bootstrap to test the statistical significance of each of our logit estimates. A correction is required because our “matched-pairs” design leads to a choice-based sample bias. The Appendix describes the corrected quasi-bootstrap procedure in detail and shows its superior efficiency properties compared to those of single balanced logit regression.

Table 3 presents corrected estimates with bootstrap standard errors for specifications that control for firms’ size differentials (i.e.,  $k$ ) while also allowing  $q_r$ ,  $z_r$ , and  $\lambda$  to enter separately with both linear and quadratic terms. For example, in the first column of the table the coefficient on  $\log k$  is 0.889 and that of  $(\log k)^2$  is -0.115, with both statistically significant at the one percent level. This indicates that the probability of a merger is a nonlinear function of the size differential between acquirer and target, though the quadratic term is not very large. The estimated coefficients for  $\log q_r$  and  $(\log q_r)^2$ , on the other hand, are 0.612 and -0.405 respectively and both are also statistically significant at one percent level. This indicates that the distance between acquirer’s  $Q$  and target’s  $Q$  has a strong inverted U-shaped effect on the probability of the two firms merging. Column 2 shows this finding to be robust to



Table 3. Quasi-bootstrap logit regressions.

	(1)	(2)	(3)	(4)	(5)	(6)
constant	-8.162*** (0.041)	-7.977*** (0.053)	-8.042*** (0.044)	-7.849** (0.059)	-8.047*** (0.045)	-7.868** (0.058)
log k	0.889*** (0.043)	0.900*** (0.045)	0.875*** (0.043)	0.883*** (0.044)	0.857*** (0.043)	0.868*** (0.044)
(log k) <sup>2</sup>	-0.115*** (0.009)	-0.116*** (0.009)	-0.112*** (0.009)	-0.113*** (0.009)	-0.112*** (0.008)	-0.113*** (0.009)
log q <sub>r</sub>	0.612*** (0.058)	0.613*** (0.058)				
(log q <sub>r</sub> ) <sup>2</sup>	-0.405*** (0.043)	-0.425*** (0.046)				
log z <sub>r</sub>			-0.066 (0.059)	-0.072 (0.061)		
(log z <sub>r</sub> ) <sup>2</sup>			-0.367*** (0.058)	-0.378*** (0.061)		
log λ					0.382*** (0.059)	0.391*** (0.060)
(log λ) <sup>2</sup>					-0.300*** (0.047)	-0.308*** (0.048)
Industry effects	No	Yes	No	Yes	No	Yes
Obs.	2634	2634	2634	2634	2634	2634

Note: Coefficients are means from the quasi-bootstrap estimation with bootstrap standard errors in parentheses. \*\* and \*\*\* represent statistical significance at the 5 and 1 percent levels, respectively.

the inclusion of dummy variables for industries. We conclude that the data strongly support a pattern of ‘high buys low but not the lowest’ in terms of  $Q$ .

Columns 3 and 4 in Table 3 indicate that the productivity difference ( $z_r$ ) between an acquirer and its target also has a strong nonlinear effect on the propensity to merge

that is robust to the inclusion of industry fixed effects. We also observe that  $(\log z_r)^2$  and  $(\log q_r)^2$  have very similar effects on the probability of two firms being involved in a merger because the estimated coefficients on these variables in columns 1 and 3 are very close (the absolute difference is only 0.038). To illustrate, we simulate the effects of  $\log q_r$ , and  $\log z_r$  on the likelihood of a merger using the estimated coefficients listed in columns 1 and 3 and present them, along with the difference between the two curves in Figure 3.<sup>8</sup> The figure shows that  $\log q_r$  and  $\log z_r$  have an inverted U-shaped effect on mergers, but their difference (i.e.,  $\log \lambda$ ) affects mergers positively and nearly linearly.<sup>9</sup>

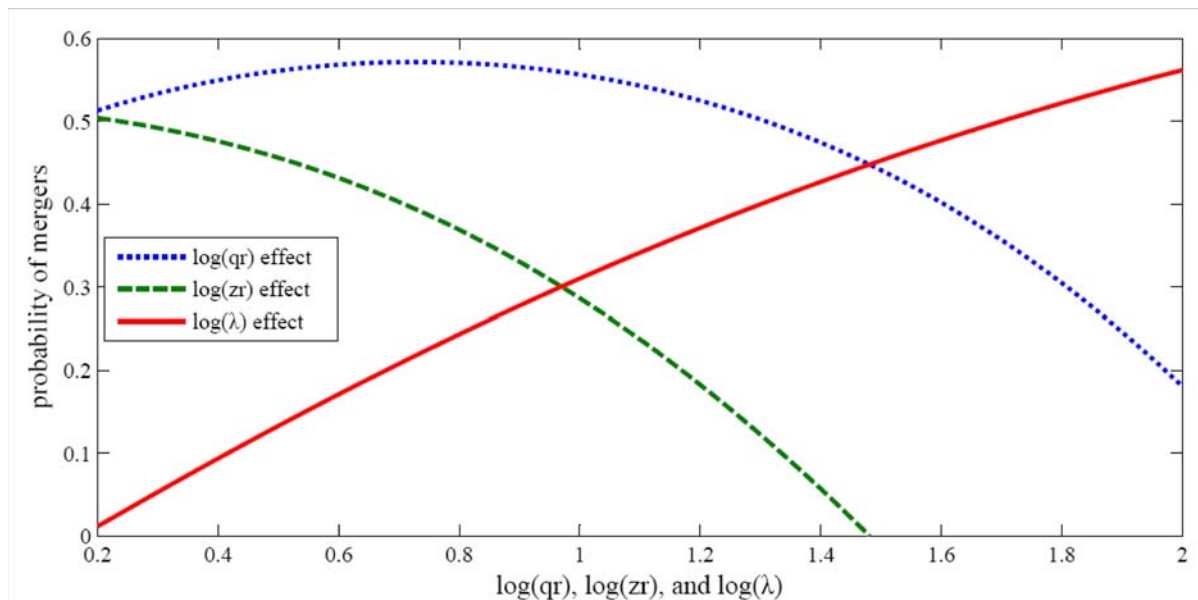


Figure 3:  $q_r$ ,  $z_r$ , and  $\lambda$ 's effect on the probability of mergers

Columns 5 and 6 in Table 3 report specifications in which we directly estimate the effects of  $\log \lambda$  and  $(\log \lambda)^2$  on the probability of a merger. In this case, we get a large positive estimate on  $\log \lambda$  but a negative one for  $(\log \lambda)^2$  with both statistically significant at the one percent level. While this suggests that there is a nonlinear effect of  $\lambda$  on the propensity to merge, we note that the coefficients on  $(\log \lambda)^2$  are much smaller than those obtained on  $(\log q_r)^2$  and  $(\log z_r)^2$ . In other words, the effect of  $\log \lambda$  on mergers has much less curvature than that of  $\log q_r$  and  $\log z_r$ , meaning that deflating  $q_r$  with  $z_r$  to form  $\lambda$  does indeed tend to “linearize” the relationship between the propensity to merge,  $q_r$ , and  $z_r$ , especially in the range of  $\lambda$  values that dominate our data (i.e.,  $\lambda < 3$ ). In terms of  $\lambda$ , then, the merger pattern is more closely one in

<sup>8</sup>Since the actual likelihood of a merger is close to zero, we simulate the probability of a merger with uncorrected coefficients and set  $\log k = 1.5$ .

<sup>9</sup>Indeed, the “ $q_r$ ” effect in Figure 3 is the empirical analog of Figure 2.

which firms with high growth potential tend to acquire firms with low potential, or simply put, the ‘high buys low’ pattern that the model predicts.

## 4 Conclusion

We build a model of mergers in which capital is firm specific and a cost is required to convert a target’s capital into a form usable by the acquirer. The model predicts that (1) the ratio of an acquirer’s  $Q$  to that of a potential target has an inverted U-shaped effect on the probability of the two firms merging, meaning that an acquirer may not purchase the lowest- $Q$  firm that it can find but rather a firm with a lower  $Q$ , and (2) the likelihood of a merger is positively and linearly related to the relative growth potential of the acquirer and its target.

The model can be viewed an integration of the ‘high buys low’ result of the  $Q$ -theory of mergers with the ‘like buys like’ pattern uncovered by Rhodes-Kropf, Robinson and Viswanathan (2005) and Rhodes-Kropf and Viswanathan (2007). Though we interpret the non-linearity of the relationship between mergers and Tobin’s  $Q$  as arising from the firm-specificity of capital and the associated adjustment costs, this also includes human capital (e.g., management) that must adjust to bring the target’s assets under common control.

Using data for mergers among U.S. firms available from the SDC dataset from 1986 to 2005, we found that our measure of relative growth potential,  $\lambda$ , does deliver a merger pattern of ‘high buys low’. A series of quasi-bootstrap logit regressions for the probability of a merger reveal inverted U-shaped relationships between mergers and the ratios of both bidder-target  $Q$ s and TFP levels, while the relation between mergers and relative growth potentials of the possible partners has far less curvature. We consider these empirical results to offer strong support for the model.

## 5 Appendix

*Efficiency of the Quasi-bootstrap Logit Regressions.*—Assume the set  $N$  includes all observations on the dependent variable  $Y_{i0} = 0$ ,  $i = 1, 2, \dots, n_0$ , and the set  $M$  includes all observations on the dependent variable  $Y_{j1} = 1$ ,  $j = 1, 2, \dots, n_1$ . Assume also that  $n_0 \gg n_1$ , meaning that  $n_0$  is dozens to thousands of times more than  $n_1$ .<sup>10</sup> The fraction of ones in the population,  $\tau$ , equals  $\frac{n_1}{n_0+n_1}$ . Given a set of regressors  $x_i$ , the object to estimate is  $P(Y_{i1} = 1 | x_i)$  (i.e., the full conditional distribution of  $Y$ ). We assume that the underlying distribution of the dependent variable is logit so that  $P(Y_{i1} = 1 | x_i)$  can be expressed as

---

<sup>10</sup>Here, 1 represents the actual merger pairs and 0 represents the pseudo-merger pairs. Therefore,  $N_1 = 1,317$  and  $N_0 = 3,054,479$ .

$$P(Y_{i1} = 1|x_i) = \frac{1}{1 + e^{-x_i'\beta}}, \quad (\text{A.1})$$

where  $\beta$  is the true parameters for the sample.<sup>11</sup>

We construct a new set  $A_t$ ,  $t = 1, 2, \dots, T$  that contains  $n_1$  observations randomly selected with replacement from  $N$ . We then run a logit regression using all the observations from  $A_t$  and  $M$ ,  $t = 1, 2, \dots, T$ . From this procedure, we obtain  $T$  estimates of  $\beta$ , which are  $\{\hat{\beta}_t, t = 1, 2, \dots, T\}$ .

For some  $T \times T$  weighting matrix  $W > 0$ , let

$$J_T(\beta) = T \begin{pmatrix} (\hat{\beta}_1 - \beta)/T \\ \vdots \\ (\hat{\beta}_T - \beta)/T \end{pmatrix}' W \begin{pmatrix} (\hat{\beta}_1 - \beta)/T \\ \vdots \\ (\hat{\beta}_T - \beta)/T \end{pmatrix}. \quad (\text{A.2})$$

According to the minimum distance method (MDM), the estimate should minimize  $J_T(\beta)$ . When  $W$  is the identity matrix, the solution of  $\beta$  minimizing  $J_T(\beta)$  is the mean of  $\{\hat{\beta}_t, t = 1, 2, \dots, T\}$ . We define  $\bar{\hat{\beta}}$  as the MDM estimate,

$$\bar{\hat{\beta}} = \frac{\sum_{t=1}^T \hat{\beta}_t}{T}. \quad (\text{A.3})$$

Since all the estimates  $\{\hat{\beta}_t, t = 1, 2, \dots, T\}$  share some dependent variables  $M$ ,  $\{\hat{\beta}_t, t = 1, 2, \dots, T\}$  are not independent from each other. The estimate  $\bar{\hat{\beta}}$  therefore is not the most efficient. To get a more efficient estimate, we set  $W$  as the variance-covariance of  $\{\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_T\}$ , which is unknown and can not be constructed easily. However,  $\bar{\hat{\beta}}$  still has the following asymptotic properties.

**Claim 5** *The asymptotic properties of  $\bar{\hat{\beta}}$  are:*

- (1)  $\bar{\hat{\beta}} \rightarrow_p \beta$ .
- (2) Under  $H_0 : \beta = 0$ ,  $\sqrt{T}(\bar{\hat{\beta}}) \rightarrow_d N(0, \sigma^2)$ , where  $\sigma^2$  is unknown.
- (3)  $\bar{\hat{\beta}}$  is more efficient than using a single draw of matched pairs in a probit regression.

Since  $\sigma^2$  is unknown, we can not directly test the statistical significance of  $\bar{\hat{\beta}}$ . Instead, we use the four-step bootstrap described below to obtain the standard errors and  $P$  values of  $\bar{\hat{\beta}}$ :

---

<sup>11</sup>We adopt the logistic regression for our probability models and their ‘‘matched-pairs’’ data design because even though the estimates are biased they can be easily corrected under logistic assumption.

Step 1. Draw with replacement  $n_1$  observations from  $M$ .

Step 2. Draw with replacement  $n_1$  observations from  $A_t$  and combine them with the sample we obtain in Step 1.

Step 3. Run logit regressions using each combined sample from Step 2.

Step 4. Repeat Steps 1-3  $B$  times.<sup>12</sup>

Hence, we define set  $N_{tb}$  that includes  $n_1$  observations randomly drawn with replacement from each  $A_t$ , and set  $M_{tb}$  that includes  $n_1$  observations randomly drawn with replacement from  $M$ , where  $t = 1, 2, \dots, T$  and  $b = 1, 2, \dots, B$ . Next, we run a logit regression using all the observations from sets  $N_{tb}$  and  $M_{tb}$  and denote the estimate as  $\hat{\beta}_{tb}^*$ .

We define that

$$\bar{\beta}_b^* = \frac{\sum_{t=1}^T \hat{\beta}_{tb}^*}{T}, \quad b = 1, 2, \dots, B, \quad (\text{A.4})$$

which is constructed in the same way as  $\bar{\beta}$ . The standard deviations of  $\{\bar{\beta}_b^*, b = 1, 2, \dots, B\}$  are thus the standard errors of  $\bar{\beta}$ . Since the mean of  $\bar{\beta}_b^*$  is  $\bar{\beta}$  and the null hypothesis  $H_0$  is that  $\beta = 0$ , we take  $\bar{\beta}_b^* - \bar{\beta}$  as the simulated test statistics. There are two cases for which to construct the empirical distribution function (EDF) based on the one-sided test.

If the alternative hypothesis  $H_1$  is  $\beta > 0$ , then the EDF is

$$\hat{F}^*(\bar{\beta}) = \frac{1}{B} \sum_{b=1}^B I(\bar{\beta}_b^* - \bar{\beta} \leq \bar{\beta}). \quad (\text{A.5})$$

Our estimate of the true  $P$  value is therefore

$$\hat{p}^*(\bar{\beta}) = 1 - \hat{F}^*(\bar{\beta}) = 1 - \frac{1}{B} \sum_{b=1}^B I(\bar{\beta}_b^* - \bar{\beta} < \bar{\beta}) = \frac{1}{B} \sum_{b=1}^B I(\bar{\beta}_b^* > 2\bar{\beta}). \quad (\text{A.6})$$

The last equality in equation A.6 means that the true  $P$  value is approximated by the proportion of simulations in which  $\bar{\beta}_b^*$  is greater than  $2\bar{\beta}$ . For example, if  $B = 599$ , and 25 of all the  $\bar{\beta}_b^*$  are greater than  $2\bar{\beta}$ , then  $\hat{p}^*(\bar{\beta}) = 25/599 = 0.042$ . As a result, in this example we would reject the null hypothesis that  $\beta = 0$  at 5 percent statistic significant level.

If the alternative hypothesis  $H_1$  is  $\beta < 0$ , then the EDF is

$$\hat{F}^*(\bar{\beta}) = \frac{1}{B} \sum_{b=1}^B I(\bar{\beta}_b^* - \bar{\beta} \geq \bar{\beta}). \quad (\text{A.7})$$

---

<sup>12</sup>According to Davidson and MacKinnon (2004), if we will perform a bootstrap test at level  $\alpha$ , then  $B$  should be chosen to satisfy the condition that  $\alpha(B + 1)$  is an integer.

Our estimate of the true  $P$  value is

$$\hat{p}^*(\bar{\beta}) = 1 - \hat{F}^*(\bar{\beta}) = 1 - \frac{1}{B} \sum_{b=1}^B I(\hat{\beta}_b^* - \bar{\beta} \geq \bar{\beta}) = \frac{1}{B} \sum_{b=1}^B I(\hat{\beta}_b^* < 2\bar{\beta}). \quad (\text{A.8})$$

If  $B$  is infinitely large, the EDF converges to the true conditional distribution function. Consequently, our procedure would yield an exact test and the outcome of the test would be the same as the  $P$  value computed using the conditional distribution function of  $\bar{\beta}$ .

*Correction of the Estimates.*—Though the bootstrap has been adopted in the above procedure, it still represents a “matched-pairs” design, which results in a sample proportion of merged pairs of 0.50. This type of sampling typically implies that the proportion of merged pairs in the sample is much larger than the proportion of such pairs in the population of all pairs (merged and non-merged). This design causes a “choice-based sample bias” for the constant and the coefficients in the standard logit models, in turn meaning that the probabilities being assessed in such models are also biased. Hence, it is necessary to correct the estimates from the above quasi-bootstrap logit regressions.

Since the fraction of ones in the population,  $\tau$ , is known and equals  $\frac{n_1}{n_0+n_1}$ , we can adopt the prior correction for the logit model (see King and Zeng, 2001). For each logit regression above, the constant item  $\hat{\beta}_{t0}^*$  can be corrected by subtracting out the bias factor,  $\ln \left[ \left( \frac{1-\tau}{\tau} \right) \right]$ , and other parameters are statistically consistent<sup>13</sup>. The final corrected estimate  $\bar{\beta}_1^c$  is the same as  $\bar{\beta}_1$  and the final corrected estimate  $\bar{\beta}_0^c$  for the constant item  $\bar{\beta}_0$  is

$$\bar{\beta}_0^c - \ln \left[ \left( \frac{1-\tau}{\tau} \right) \right] \quad (\text{A.9})$$

because  $\bar{\beta}_1 = \frac{\sum_{t=1}^T \hat{\beta}_{t1}^*}{T}$  and  $\bar{\beta}_0 = \frac{\sum_{t=1}^T \hat{\beta}_{t0}^*}{T}$ . As a result, we can state the following corollary.

**Corollary 6** *The corrected estimate  $(\bar{\beta}_0^c, \bar{\beta}_1^c)$  is unbiased and has the same asymptotic properties as  $\bar{\beta}$ .*

## References

- [1] Andrade, G., M. Mitchell and E. Stafford, 2001, "New Evidence and Perspectives on Mergers." *Journal of Economic Perspectives*, 15(2), 103-120.

---

<sup>13</sup>Here,  $y = 1/2$ , so  $\frac{y}{1-y} = 1$ .

- [2] Cockburn, I. and Z. Griliches, 1988, "Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents." *American Economic Review*, 78 (2), 419-423.
- [3] Davidson, R. and J. G. MacKinnon, 2004, "Econometric Theory and Methods." Oxford University Press.
- [4] Dwyer, D. W., 2001, "Plant-level Productivity and The Market Value of A Firm." Working Paper.
- [5] Faria, A., 2002, "Mergers and the Market for Organization Capital." University of Chicago, Mimeo.
- [6] Gort, M., 1969, "An Economic Disturbance Theory of Mergers." *Quarterly Journal of Economics*, 83, 624-642.
- [7] Greenwood, J., Z. Hercowitz and P. Krusell, 1997, "Long-Run Implications of Investment-Specific Technological Change." *The American Economic Review*, 87(3), 342-362.
- [8] Greenwood, J., Z. Hercowitz and P. Krusell, 2000, "The Role of Investment-Specific Technological Change in The Business Cycle." *European Economic Review*, 44, 91-115.
- [9] Griliches, Z., 1981, "Market Value, R&D and Patents," *Economics Letters* 7, 183-187.
- [10] Harford, J., 2005, "What Drives Merger Waves?" *Journal of Financial Economics*, 77, 529-560.
- [11] Hulten, C. R., 1992, "Growth Accounting When Technical Change Is Embodied in Capital." *The American Economic Review*, 82(4), 964-980.
- [12] King, G. and L. Zeng, 2001, "Logistic Regression in Rare Events Data." *The Society for Political Methodology*, 137-163.
- [13] Jovanovic, B., and P. Rousseau, 2002, "The *Q*-Theory of Mergers." *The American Economic Review*, (Papers and Proceedings), 92(2), 198-204.
- [14] Jovanovic, B., and P. Rousseau, 2008, "Mergers as Reallocation." *Review of Economics and Statistics*, 90(4), 765-776.
- [15] Mitchell, M.L. and J.H. Mulherin, 1996, "The Impact of Industry Shocks on Takeover and Restructuring Activity." *Journal of Financial Economics*, 41, 193-229.

- [16] Rhodes-Kropf, M. and D. Robinson, 2007, "The Market for Mergers and the Boundaries of the Firm." *Journal of Finance*, forthcoming.
- [17] Rhodes-Kropf, M., D. Robinson, and S. Viswanathan, 2005, "Valuation Waves and Merger Activity: the Empirical Evidence." *Journal of Financial Economics*, 77, 561-603.
- [18] Servaes, H., 1991, "Tobin's  $Q$  and the Gains from Takeovers." *The Journal of Finance*, 46(1), 409-419.
- [19] Shleifer, A. and R. Vishny, 2003, "Stock Market Driven Acquisitions." *Journal of Financial Economics*, 70, 295-311.