Merger Waves: Theory and Evidence

Jinghua Yan*

First draft: February, 2006
This version: December 17, 2006

ABSTRACT

This paper presents a model that incorporates product market competition into the standard neoclassical framework. The model explains why value-maximizing firms conduct mergers that appear to lower shareholder value. In a Cournot setting, the model demonstrates a prisoners' dilemma for merging firms in a merger wave. Consistent with the model's implications, the paper empirically documents that horizontal mergers are followed by substantially worse performance when they occur during waves. Moreover, further empirical tests show that the empirical relation between performance and merger waves is independent of the method of payment and increasing in the acquirer's managerial ownership. These findings are difficult to reconcile with alternative interpretations from existing theories.

*The Wharton School at the University of Pennsylvania. Email: jyan@wharton.upenn.edu.
I thank Philip Bond, Gary Gorton, Wei Jiang, Jozsef Molnar, David Musto, Vinay Nair, Michael Roberts, Pavel Savor, Geoffrey Tate, Ding Wu, Julie Wulf, Jianfeng Yu, Paul Zurek, and seminar participants at Wharton, the CRSP Forum, and Lehman Brothers for helpful discussions. I am especially grateful to Rob Stambaugh and my committee members, Franklin Allen (chair), Chris Géczy, Itay Goldstein, João Gomes, Robert Holthausen, and Andrew Metrick for their comments, guidance, and encouragement.

1 Introduction

One of the most enduring puzzles in modern corporate finance is why many mergers appear to lower shareholder value.[1] The existing neoclassical theory, which assumes value maximization and market efficiency, fails to provide an explanation. By deviating from the standard neoclassical assumptions, two strands of literature have provided solutions to this puzzle. Agency theory attributes the negative post-merger stock performance to a principal-agent problem. Market timing theory attributes it to an overdue correction of mispricing. In the absence of agency costs and market inefficiencies, this paper proposes an explanation by incorporating the role of product market competition into the standard neoclassical framework. In a neoclassical setting where mergers facilitate technology transfer between firms, mergers that take place outside merger waves (hereafter, off-the-wave mergers) increase shareholder value due to the value maximization principle. However, if such horizontal mergers take place in a wave that is driven by technology shocks,[2] the improved technology of merging firms and an increasingly concentrated market structure alters the competitive landscape for non-merging rival firms. When merging firms' improvement in production efficiency is sufficiently high, stand-alone rivals in an on-going merger wave may face a declining profit margin and a shrinking market share. The merger wave thus resembles a game of prisoners' dilemma: each individual pair chooses to merge despite the fact that their combined value is less than that prior to the merger wave. Therefore, mergers that take place in a merger wave (hereafter, on-the-wave mergers) may appear to lower shareholder value. The poor performance following on-the-wave mergers can nevertheless be consistent with value maximization.

Guided by the model developed in the paper, I discover that horizontal mergers are followed by substantially worse performance when they occur during waves. Waves are identified here using the concentration or "clusteredness" of contemporaneous same-industry M&A activity. Among all horizontal mergers announced during the period from 1979 to 2004, acquirer stocks in the most clustered quintile of mergers underperform those in the least clustered quintile by 15% over one year and by 40% over two years. This relation is robust to a number of performance measures, industry classifications, and empirical approaches. Moreover, as is predicted by the theory model, the underperformance of on-the-wave mergers

[1] See Jensen and Ruback (1983) and Andrade, Mitchell, and Stafford (2001) for surveys of this literature.

[2] Several papers have found that merger waves are triggered by industry-level technology or deregulation shocks, e.g., Mitchell and Mulherin (1996), Rhodes-Kropf, Robinson, Viswanathan (2005), and Harford (2005).

is more pronounced in less concentrated or non-durable goods industries. Finally, industry rivals' performance following on-the-wave mergers is also worse than that following off-the-wave mergers.

The empirical relation between performance and clusteredness documented here could potentially be consistent with market-timing and agency theories as well. The market timing theory, exemplified by Shleifer and Vishny (2003) and Rhodes-Kropf and Viswanathan (2004), suggests that the acquirer uses its relatively overvalued stock as currency to purchase the target company's stock. Such stock market driven mergers have poor long-run stock performance due to the correction of misvaluation. A central prediction of the market timing theory is that stock deal acquirers underperform cash deal acquirers in the long run.[3] To examine the possibility of the market timing theory as an alternative explanation, I show that in the data the relation between performance and clusteredness is independent of the method of payment. Moreover, this relation is weakened when the net sales of insider shares are positive, indicating that market overvaluation perceived by company insiders does not drive down stock price following merger waves. These findings are difficult to reconcile with the misvaluation explanation provided by the market timing theory.

The agency theory of mergers, first proposed by Jensen (1986), suggests that value-destroying mergers are driven by the manager's incentive to grow the firm beyond its optimal size. More recently, Gorton, Kahl, and Rosen (2005) show that when managers have private benefits of control, fundamental shocks may trigger defensive merger waves. One of the key predictions of agency theory is that low managerial ownership in the acquirer firm leads to poor post-merger performance.[4] In the data, I show that the negative relation between stock performance and clusteredness strengthens as acquirer managerial ownership increases, which fails to support the agency theory as an alternative explanation.

The following table compares the current model with existing theories and shows the key differences in assumptions and empirical implications. The empirical results in this paper distinguish among these alternatives *only* in the context of on-the-wave versus off-the-wave mergers.

[3]See Loughran and Vijh (1997) and Rau and Vermaelen (1998) for evidence supporting this prediction.
[4]See Lewellen, Loderer, and Rosenfeld (1985) for evidence supporting this prediction.

	Assumptions		Empirical Implications	
Theory (literature)	Value maximization	Market efficiency	Characterization of merger waves	Prediction on performance
Neoclassical (JR 2002)	Yes	Yes	Fundamental shocks	Always non-negative
Agency (Jensen 1986, GKR 2005)	No	Yes	Preemptive waves	Mixed (managerial ownership: low < high)
Market Timing (SV 2003, RKV 2004)	Yes	No	Misvaluation waves	Mixed (method of payment: stock < cash)
Neoclassical with imperfect competition (this paper)	Yes	Yes	Fundamental shocks	Mixed (merger waves: on- < off-the-wave)

The rest of this paper is organized as follows: Section 2 presents a model with imperfect product market competition, which demonstrates the prisoners' dilemma faced by firms in a merger wave. Section 3 tests the empirical implications of the model and addresses a number of alternative interpretations of the results from competing theories and hypotheses. Section 4 relates the model and findings of this paper to the existing literature on mergers. Section 5 concludes.

2 The Model

The objective of this model is to demonstrate the theoretical possibility of a merger wave equilibrium where the value-maximization principle is upheld and merging firms' value is lowered. To do this, I choose a simple framework: a static Cournot equilibrium with one period of production. The current model is not the first to investigate the role of Cournot competition on horizontal mergers. Prior literature has focused on product pricing and welfare (Salant, Switzer, and Reynolds (1983) and Farrell and Shapiro (1990)), strategic coordination on merging decisions (Fauli-Oller (2000)), and endogenous dynamics of mergers (Gowrisankaran (1999)). While these papers examine firms' incentive to merge, they do not consider explicitly the main focus of this model regarding the implications for firm value, i.e., comparing firm value across two equilibria, the status quo and the merger wave.[5] In a

[5] Molnar (2006) is an exception in this regard: it examines mergers' implication on firm value in a sequential auction framework with two acquirers competing for one target. In contrast, this paper features a more general framework with randomly paired firms making simultaneous merger decisions.

Cournot setting, the improved technology of merging firms alters the competitive landscape for non-merging rival firms. In a merger wave equilibrium, each individual merger enhances the value of merging firms. Nonetheless, the value of a merged firm under the merger wave equilibrium need not be higher than the total value of two stand-alone firms under the status quo. Thus, merging firms may face a prisoners' dilemma: each individually value-maximizing pair of firms conduct mergers that appear to lower shareholder value.

Davidson and Deneckere (1985) argue that quantity setting games, such as Cournot, understate firms' incentive to merge. They advocate the use of Bertrand competition with product differentiation instead. The choice of Cournot framework in this paper is for expositional simplicity. It can be shown that in a Bertrand framework with product differentiation as in Davidson and Deneckere (1985) the main conclusions of this paper remain unchanged.[6]

2.1 A Simple Model of Mergers

Consider an existing economy with N firms, where N is even. One half of the firms ($\frac{N}{2}$) produce using high technology and the other half produce using low technology. Let C^h (C^l) denote the cost of production for high (low) technology firms ($C^h < C^l$). Each firm is endowed with K_0 units of capital stock. All firms produce identical goods and strategically set the quantity of production in a Cournot setting. Stand-alone firms of type i produce $q^i = \frac{K^i}{C^i}$ units of goods, where $i = h, l$.

At $t = 0$, high and low technology firms form $\frac{N}{2}$ identical pairs, and each pair decides whether or not to merge. If two firms remain stand-alone, their costs of production will remain unchanged. If two firms decide to merge, the combined firm's technology, denoted as C^{hl}, satisfies $\frac{1}{C^{hl}} = \frac{1}{C^h} + \delta \frac{1}{C^l}$, where $\delta \geq 0$. δ is a measure of complementarity between the two merging firms. For example, Adidas has higher overall productivity than Reebok as is reflected by their market share, but Reebok has stronger distribution channels in European and Asian markets than Adidas does. If the merged firm integrates the strength of each firm, its productivity will be higher than both stand-alone firms. Thus, gains from the merger are higher if the degree of complementarity is higher. The notion of complementarity in this model is similar to that in Rhodes-Kropf and Robinson (2006).[7] If $\delta = 0$, then we have the special case that $C^{hl} = C^h$, as in Jovanovic and Rousseau (2002). When the degree of

[6] The main conclusions also hold in a quantity-setting game with tangible assets, as in Perry and Porter (1985).

[7] The supermodularity condition in Rhodes-Kropf and Robinson (2006) would require $\delta > 1$. In this regard, the complementarity assumption in this paper, i.e., $\delta \geq 0$, is less restrictive.

complementarity is zero, the low type firm's technology will be completely replaced by the high type firm's technology.

I rule out "mergers of likes", i.e., mergers of two firms of the same type, for tractability reasons. One can relax this assumption and the key results remain.[8] In the special case $C^l = C^h$, i.e., all firms are initially identical at $t = 0$, this assumption becomes irrelevant. In addition, I rule out one firm matching with multiple firms. It takes a considerable amount of time for the acquirer to conduct a thorough due diligence on the target.[9]

Let A_i denote the decision of pair i, where $i \in \{1, 2, ..., \frac{N}{2}\}$. $A_i = 1$ if two firms merge and $A_i = 0$ otherwise. Let x denote the number of pairs that decide to merge at $t = 0$, i.e., $x = \sum_{i=1}^{\frac{N}{2}} A_i$. It follows that among the remaining $N - x$ firms in the economy, x firms produce under C^{hl}, $\frac{N}{2} - x$ firms produce under C^h, and $\frac{N}{2} - x$ firms produce under C^l. I rule out entry into the industry.

At $t = 1$, each merged or stand-alone firm adjusts its capital stock level K^i such that the marginal revenue of each additional unit of capital stock equals the external cost of capital stock (denoted as k), i.e., $\frac{dR^i}{dK^i} = k$. Without loss of generality, I normalize k to 1. Thus, firm i's profit is given by

$$\pi^i(K^i, P) = P\frac{K^i}{C^i} - (K^i - K_0),$$

where the market clearing price for firms' output is assumed to satisfy an inverse demand function of a constant elasticity form,

$$P = X^{\gamma}Q^{-\gamma},$$

where Q denotes the aggregate industry output, i.e., $Q = \sum_{i=1}^{N} q^i$, and $\frac{1}{\gamma}$ is the price elasticity of demand. For expositional simplicity, X, which describes the condition of the aggregate economy, is normalized to 1. Since the external capital stock market is frictionless in this model, i.e., there is no capital adjustment cost, the initial capital stock level K_0 becomes irrelevant to firms' incentive to merge. Without loss of generality, I assume that $K_0 = 0$.

As in the existing literature on mergers and imperfect product market competition,[10] this

[8] The empirical evidence on "mergers of likes" is mixed. Servaes (1991) finds that mergers of high M/B and low M/B have higher total returns. In contrast, Rhodes-Kropf and Robinson (2006) find that mergers typically pair together firms with similar M/B ratios.

[9] Empirically, it is uncommon for one firm to make multiple sizable acquisitions within the same year.

[10] See, for example, McCardle and Viswanathan (1994), Salant, Switzer, and Reynolds (1983), Farrell and Shapiro (1990), and Gowrisankaran (1999).

paper uses backward induction to solve the model. Production and profits are determined by a Cournot production game among the remaining firms in the industry. Thus, I characterize the optimal strategy for each pair of high-low technology firms as follows: there are x mergers in the existing economy, $A_i^*(x) = 1$ if and only if

$$\pi^{hl}(x+1) - I > \pi^h(x) + \pi^l(x), \tag{1}$$

where I denotes the integration costs (or fixed cost savings, if $I < 0$).[11]

A merger wave equilibrium is a pure-strategy equilibrium when the optimal response for the $\frac{N}{2}th$ pair conditional on all other $(\frac{N}{2} - 1)$ pairs merging is to merge:

Definition 1 *Merger wave equilibrium:* $x^* = \frac{N}{2}$ *is a pure-strategy equilibrium, i.e.,* $A(\frac{N}{2} - 1) = 1$*, where* $A_i^*(x)$ *is given by (1).*

Conventional event studies[12] draw conclusions on change in firm value by comparing the post-merger value, i.e., $\pi^{hl}(\frac{N}{2}) - I$, with the status quo value, $\pi^h(0) + \pi^l(0)$, hence the following definition on the types of merger waves:

Definition 2 *A merger wave is value-creating (value-destroying) if* $\pi^{hl}(\frac{N}{2}) - I > \pi^h(0) + \pi^l(0)$ *(*$\pi^{hl}(\frac{N}{2}) - I < \pi^h(0) + \pi^l(0)$*).*

Given the number of mergers in the existing economy (x), the solution for firm profit in a Cournot setting at $t = 1$ is standard: Firm's revenue depends on its capital stock directly, because it uses the capital stock to produce the revenue generating good, and indirectly, because the price of the good depends, partly, on the firm's production:

$$\frac{d\pi^i}{dK^i} = \frac{P}{C^i} + \frac{K^i}{C^i}\frac{dP}{dK^i} - 1 = 0.$$

Differentiating the inverse demand function $P = Q^{-\gamma}$ with respect to K^i gives,

$$\frac{dP}{dK^i} = \frac{dP}{dQ}\frac{dQ}{dK^i} = -\gamma\frac{P}{C^iQ},$$

and substituting into the previous equation together with $q^i = \frac{K^i}{C^i}$ yields,

$$1 = (1 - \gamma\frac{q^i}{Q})\frac{P}{C^i}, \tag{2}$$

[11] Integration costs are often dependent on firm size and synergies. The simplified assumption here is not crucial to the conclusion.

[12] See, for example, Andrade, Mitchell, and Stafford (2001).

which indicates that the firm internalizes the price impact in proportion to its market share, $\frac{q^i}{Q}$. Since all firms' marginal valuations of capital equate to the cost of the capital stock, (2) holds for all i. Therefore, $(1-\gamma\frac{q^i}{Q})\frac{1}{C^i} = (1-\gamma\frac{q^j}{Q})\frac{1}{C^j}$, for any $j \in \{1, 2, ..., N-x\}$. Summing over firms yields,

$$\frac{q^i}{Q} = \frac{\bar{C}(x) - (1-\frac{\gamma}{N-x})C^i}{\gamma\bar{C}(x)}, \tag{3}$$

subject to the constraints

$$\forall x \in [0, \frac{N}{2}-1], \frac{\left(1-\frac{\gamma}{N-x}\right)C^i}{\bar{C}(x)} < 1, \tag{4}$$

where

$$\bar{C}(x) = \frac{(\frac{N}{2}-x)C^h + (\frac{N}{2}-x)C^l + xC^{hl}}{N-x}.$$

$\bar{C}$ denotes the equally weighted industry average capital requirement per unit of production. I rewrite (2) as $k = (\frac{1-\frac{\gamma}{N-x}}{\bar{C}})P$. Hence,

$$P(x) = \frac{\bar{C}(x)}{1-\frac{\gamma}{N-x}}, \tag{5}$$

and the inverse of the demand function gives

$$Q(x) = P(x)^{-\frac{1}{\gamma}} = (\frac{\bar{C}(x)}{1-\frac{\gamma}{N-x}})^{-\frac{1}{\gamma}}.$$

The market share of each firm becomes

$$q^i(x) = \frac{\bar{C}(x) - (1-\frac{\gamma}{N-x})C^i}{\gamma\bar{C}(x)}Q(x) = \frac{\bar{C}(x) - (1-\frac{\gamma}{N-x})C^i}{\gamma\bar{C}(x)}(\frac{\bar{C}(x)}{1-\frac{\gamma}{N-x}})^{-\frac{1}{\gamma}}.$$

Finally, the profit function of each firm is derived as follows,

$$\pi^i(x) = \underbrace{\bar{C}(x)^{1-\frac{1}{\gamma}}\left(1-\frac{\gamma}{N-x}\right)^{\frac{1}{\gamma}-1}}_{\text{aggregate economic condition}}\underbrace{\frac{1}{\gamma}\left(1-\frac{(1-\frac{\gamma}{N-x})C^i}{\bar{C}(x)}\right)}_{\text{market share}}\underbrace{\left(1-\frac{(1-\frac{\gamma}{N-x})C^i}{\bar{C}(x)}\right)}_{\text{profit margin}}, \tag{6}$$

where $i = h, l,$ and hl.

A firm's profit function is the product of three terms: the condition of the aggregate economy, the firm's market share, and the firm's profit margin. The latter two terms both depend on the firm's cost of production, C^i, relative to the average cost of production in the economy, $\bar{C}$, and the number of remaining firms in the economy, $N - x$. When the number of mergers increases in the economy, there are fewer firms remaining and the economy moves closer to monopoly, thus increasing the rival firms' market share and profit margin. However, when the degree of complementarity between merging firms is sufficiently high, the decrease in the number of competitors does not fully compensate for the increase in average competitiveness. In this case, a merger will lower the stand-alone value of rival firms.

2.2 Existence of Value-Destroying Merger Wave Equilibrium

Depending on the values of a few key parameters, a merger wave equilibrium may enhance or lower shareholders' value under the status quo. This section starts with the special case of $\delta = 0$ as in Jovanovic and Rousseau (2002) to examine conditions for value-creating and value-destroying merger waves.

Proposition 1 *A merger wave equilibrium* $\left(x^* = \frac{N}{2}\right)$ *always payoff dominates the status quo* ($x = 0$)*, i.e.,* $\pi^{hl}(\frac{N}{2}) - I > \pi^h(0) + \pi^l(0)$*, if (i)* $\delta = 0$ *and (ii) regularity conditions (4) hold.*

The Appendix provides a detailed proof of this proposition. Proposition 1 shows that under constant returns to scale and the standard neoclassical assumptions on technological advancement in mergers, a merger can only have positive externalities on its rivals, i.e., the stand-alone values of rivals improve as the number of mergers in the industry increases. For rival firms, the gain from having fewer competitors (lower $N - x$) always outweighs the loss from facing (on average) more technologically advanced competitors (lower $\bar{C}$). If the merger wave is an equilibrium, then each pair of firms must be better off than standing still conditional on all other pairs merging. By transitivity, merger wave equilibrium must payoff dominate the status quo. This statement holds for all values of l. The following proposition demonstrates that a non-zero degree of complementarity is not only necessary but also sufficient for a merger wave equilibrium to be payoff dominated by the status quo.

Proposition 2 *There exists a value-destroying merger wave equilibrium* $\left(x^* = \frac{N}{2}\right)$*, i.e.,* $\pi^{hl}(\frac{N}{2}) - I < \pi^h(0) + \pi^l(0)$*, if*

(i) $\delta > \underline{\delta}$*, where* $\underline{\delta}$*, the complementarity threshold, is given by* $l\left(\frac{N-\gamma}{\left(\frac{N}{2}-\gamma\right)(1+l)} - 1\right)$,

(ii) $I \in \left(\pi^{hl}(\frac{N}{2}) - \pi^h(0) - \pi^l(0),\ \pi^{hl}(\frac{N}{2}) - \pi^h(\frac{N}{2} - 1) - \pi^l(\frac{N}{2} - 1)\right)$*, and*

(iii) regularity conditions (4) hold.

The Appendix provides a detailed proof of this proposition. When the degree of complementarity between two firms is sufficiently high, i.e., $\delta > \underline{\delta}$, mergers bring about negative externalities. For rival firms, the gains from having fewer competitors (lower $N - x$) can be outweighed by facing tougher competition in the product market (lower $\bar{C}$). Therefore, in a merger-wave equilibrium, every pair of firms may be worse off than under the status quo, even though each individual pair's strategy is value maximizing. Such a merger wave would be labeled as "value-destroying" by conventional event studies. It can nevertheless be consistent with value maximization.

Proposition 2 derives the set of conditions for value-destroying merger waves centered upon two key parameters: δ and I. Since integration costs are highly idiosyncratic and unobservable, I will focus on the degree of complementarity (δ) to derive the main empirical implications of the model. It is a well established fact that merger waves are driven by technology or deregulation shocks. Such a shock would translate into a sudden increase in complementarity in this model. For example, a technology breakthrough in online payment processing leads to high complementarity between a conventional bookstore and an internet retailer. Or, a deregulation that allows telecommunication companies to operate across different states leads to an increase in complementarity between two phone companies initially operating in different states. An increase in δ has two important effects: first, it increases firms' incentive to merge, thus moving the economy from its status quo to a merger-wave equilibrium. This effect of technology shocks on mergers is similar to that in Jovanovic and Rousseau (2002). Second, Proposition 2 shows that when δ is sufficiently high, the merger wave may lower shareholder value. Thus, the model establishes an inherent link between two of the most well-known empirical facts about mergers, namely merger waves and poor post-merger performance.

2.3 Comparative Statics

I have shown that a merger wave may increase or lower shareholder value depending on the value of δ. This section examines the comparative statics on $\underline{\delta}$, the minimum threshold for a merger wave to be value-destroying, with regard to two relevant industry characteristics.

Corollary 3 *The complementarity threshold ($\underline{\delta}$) is decreasing in the initial number of firms in the economy (N). Moreover, as $N \to \infty$, $\underline{\delta} \to 0$.*

Corollary 3 states that the minimum degree of complementarity required for a merger wave to destroy private value ($\underline{\delta}$) is decreasing in the number of firms operating in the industry (N). The intuition is as follows: the externalities of a horizontal merger are two-fold. On the one hand, a merger improves rivals' value due to higher oligopoly rents. On the other hand, it lowers rivals' value due to tougher competition. The increase in oligopoly rents is decreasing in the initial number of companies, e.g., a merger wave that reduces the number of firms from 10 to 5 yields less oligopoly rents for remaining firms than one that reduces the number of firms from 2 to 1. Therefore, the threshold of technological synergy ($\underline{\delta}$) for the competition effect to dominate must also be decreasing in the initial number of companies (N). Hence, the model predicts that in a concentrated industry (low N), $\underline{\delta}$ is likely to be high and merger waves are less likely to destroy value.

Corollary 4 *The complementarity threshold ($\underline{\delta}$) is decreasing in the price elasticity of demand ($\frac{1}{\gamma}$).*

Corollary 4 states that $\underline{\delta}$, the minimum degree of complementarity required for a merger wave to destroy private value, is decreasing in the price elasticity of demand of the industry ($\frac{1}{\gamma}$). The intuition is that as price elasticity of demand increases, product market competition toughens and the competition effect will be more pronounced everything else equal. Therefore, the minimum degree of complementarity for a merger to have destructive impact on rival firms' values is lower for highly competitive industries (high $\frac{1}{\gamma}$). Hence, the model predicts that in industries with low price elasticity of demand, merger waves are less likely to destroy value.

2.4 On-the-wave and Off-the-wave Mergers

The baseline model demonstrates a prisoners' dilemma problem for merging firms in a merger wave, thus establishing a link between merger wave and post-merger performance. Due to its simplified assumptions, the baseline model cannot accommodate two stylized empirical facts: (a) some mergers take place outside a merger wave, and (b) some firms remain stand-alone during merger waves.[13] I can overcome this limitation by making the integration costs idiosyncratic, i.e.,

[13]In the baseline model, when integration costs are extremely high, an increase in δ may not trigger a merger wave in a sense that only the first few pairs of firms will merge. This result of the model is less consistent with the established fact that technology shocks trigger merger waves, probably because the threshold on integration costs is unrealistically high.

$$I^i = I + \varepsilon_I^i$$

where ε_I^i denotes the idiosyncratic variation in integration costs for the i^{th} pair of firms. The assumption is motivated by the idiosyncratic nature of both physical integration (such as computer systems) and cultural integration. Under this simple characterization, the model describes the merger activities in an industry as follows:

Under normal economic conditions, the degree of complementarity is low and the incentive to merge is low. Only pairs with extremely low realizations of integration costs ($\varepsilon_I^i \ll 0$) merge outside a merger wave. The proposed merger[14] between OfficeDepot and Staples was such an example.[15] Participants of these off-the-wave mergers are always better off than the sum of their stand-alone value under the status quo due to value maximization. This prediction has been verified by existing literature that indicates that horizontal mergers on average create value for shareholders (see, for example, Mitchell and Mulherin (1996)).

Technology innovations or deregulation shocks increase the degree of complementarity and trigger a merger wave. Proposition 2 shows that when δ is sufficiently high, a merger wave may leave each pair of firms worse off than under the status quo. Therefore, the central prediction of the model is that on-the-wave mergers may lower shareholder value. Moreover, in a value-destroying merger wave, some matched pairs have high positive idiosyncratic integration costs (ε_I^i), thus remaining stand-alone. These stand-alone firms also absorb negative externalities brought about by value-destroying merger waves of their rivals. Therefore, the model also predicts that rival firms' value may fall following horizontal merger waves.

Finally, the model also has two cross-industry predictions: Corollary 3 predicts that industries with higher concentration (lower N) are less likely to have value-destroying merger waves; Corollary 4 predicts that industries with low price elasticity of demand (higher γ) are less likely to have value-destroying merger waves.

2.5 Product Market Prices and Mergers

Although this is not the focus of this paper, the model has empirical implications on the relation between mergers and product market prices. A horizontal merger generates two countervailing effects on product market prices, a market power effect and a productive efficiency effect. When productive synergies are low, the market power effect dominates and

[14] The merger was later rejected by an anti-trust review by the Federal Trade Commission.

[15] In the data used in this paper, there are no other horizontal mergers in the same 4-digit SIC code over a 7-month period centered on the event month of this merger.

price rises. When productive synergies are high, the productive efficiency effect prevails and product market price falls. The empirical evidence on the impact of horizontal mergers on product market pricing is mixed yet consistent with my model. Using a sample of airline mergers from 1985 to 1988, Kim and Singal (1993) showed that merging firms raised airline ticket prices by 9% relative to the routes unaffected by the mergers.[16] In the sample used in this paper, the M&A activities in the transportation industry, i.e., Fama-French industry 40, are low during Kim and Singal's sample period relative to other periods, e.g., only 38 out of the total 323 mergers were announced during this 4-year window. Therefore, the increase in product price is consistent with this paper's theoretical prediction that low-synergy, off-the-wave mergers lead to higher product market price. More recently, Focarelli and Panetta (2003) argue that it takes time to realize productive synergies. Using a unique dataset of deposit rates of Italian banks, they show that deposit rates fall in the short run but rise in the long run after mergers, i.e., the market power effect dominates in the short run, but the productive efficiency effect dominates in the long run. Thus, the competition effect identified in this paper is not limited to the sample of domestic mergers studied in this paper.

3 Empirical Methods and Results

3.1 Data and Methods

From Thomson Financial's Securities Data Corporation (SDC), I obtain all domestic completed mergers or tender-offer bids from 1979 to 2004. I exclude all repurchases and leveraged buyouts.[17] I assign each acquirer and target to one of the Fama-French 48 industry groups based on their SIC codes recorded by SDC at the time of the announcement. If the acquirer and the target are in the same industry group, then the merger is identified as horizontal.

For each horizontal merger, I use the number of contemporaneous horizontal mergers in the same industry, as a measure of "clusteredness," or concentration of merger announcements. That is, I sum the number of horizontal mergers in the same industry announced during the announcement month, the previous 3 months, and the following 3 months. I then normalize this number by the total number of mergers announced in that industry over the

[16] In another related paper, Prager and Hannan (1998) find that large horizontal mergers of banks substantially reduce deposit rates, but financial companies are excluded in the sample due to heavy regulation.

[17] I do not exclude deals for which less than 100% of the target shares are acquired. To avoid doublecounting of multiple announcements of the same merger, I keep one observation per calendar year for each unique pair of acquirer and target. In the final sample, deals in which 100% of the target share was acquired account for over 90% of the observations.

entire sample period. This adjusted number of contemporaneous horizontal mergers will be the measure of clusteredness.

Harford (2005) calculates the highest 24-month concentration of mergers for each industry. He identifies merger waves if the actual 24-month peak concentration exceeds the 95th percentile of the simulated distribution. The measure of clusteredness used in this paper is different in two crucial ways due to different theoretic motivations: I only focus on horizontal mergers because two firms in the same industry are more likely to have complementarity of strength in production. In addition, my measure of clusteredness is continuous because externalities from rivals' mergers are continuous in nature.[18] [19]

I choose the window [t-3m, t+3m] because in the model firms make their merger (and ensuing investment) decisions simultaneously, i.e., during the same year. The results to be shown are robust to using a shorter or longer window, e.g., [t-1m, t+1m], [t-2m, t+2m], [t-6m, t+6m], and [t-12m, t+12m]. I normalize the number of contemporaneous mergers with the total number of mergers within each industry over the entire sample period similarly to Harford (2005). This is to tease out variability of merger activities across different industries.[20]

Table 1a lists the performance measure variables I use in this paper. I examine the stock return from the day before the announcement to 365 days after the announcement. In some tables, I also report the stock performance over two years. The choice of horizon balances two offsetting concerns. On the one hand, short-horizon event studies fail to take into account the impact from the horizontal mergers that are announced subsequently.[21] On the other hand, long-horizon performance may be driven by events unrelated to the merger clusteredness measures.

Summary statistics of the key characteristics are shown in Table 1b. The magnitude of the acquirer and the target's announcement return over a short event window, i.e., 10 days, in the data sample is close to the results reported in other recent large-sample studies, such as Andrade, Mitchell, and Stafford (2001), Rosen (2006), and Moeller, Schlingemann,

[18] Rosen (2006) uses the number of mergers over the last three years to identify on- and off-the wave mergers. I will relate my findings to his below.

[19] Harford (2005) only includes deals greater than $50 million in value. Including smaller deals reduces noise in the clusteredness measure constructed in this paper. The results shown below are robust to different thresholds on deal value.

[20] The baseline results are robust to using the number of companies from the COMPUSTAT tape in the given industry as the denominator. However, this measure can potentially be confounded by industry concentration as well as the population of public companies in the industry.

[21] Short-run stock performance is also subject to a speculation/anticipation effect, as in Song and Walkling (2000) and Song and Walkling (2004), as well as a temporary demand shift due to merger arbitrage, as in Mitchell, Pulvino, and Stafford (2004).

and Stulz (2004). In the entire sample, the average medium/long-run abnormal return of the acquirer is 4.5% over one year and 6.2% over two years. Neither is statistically distinguishable from zero. The same is true for operating performance measures. If these merger announcements are evenly distributed over the sample period of 25 years, then the mean of the clusteredness measure should be $\frac{7 \text{ months}}{(26\times12) \text{ months}} = 2.2\%$. In the data, the mean of the clusteredness measure using the Fama-French 48 industry classification is substantially higher (4.0%). This is consistent with the well-established fact that horizontal mergers happen in waves (see Andrade, Mitchell, and Stafford (2001)).

3.2 Results and Discussions

3.2.1 Merger Wave and Post-merger Performance

My model establishes an inherent link between merger waves and post-merger performance. While the model's implication applies to the combined value of acquirer and target, in the data, the majority of the targets are private and their stock and operating performance cannot be observed. Due to this data limitation, I first examine the performance on the acquirer for the entire sample of 11,366 mergers. I then perform the analysis on the combined value of acquirer and target for the subset of mergers in which the target is public.

In the univariate analysis, I sort all mergers where the acquirer is public by the measure of clusteredness into quintiles. In Table 2, I report by quintile the stock and operating performance measures and deal characteristics defined in Table 1. There is a strong decreasing trend in both stock and operating performance from the least clustered (quintile 1) to the most clustered (quintile 5). The difference between the two extreme quintiles is striking. On average, acquirer's stock of the most clustered horizontal mergers underperform that of the least clustered by 14.5% over one year and by 39.9% over two years. The univariate results in Table 2 are robust to two equal-sized subperiods: 1979-1996 and 1997-2004.

To verify the findings of the univariate analysis, I regress stock performance measures on the previously defined measure of within industry clusteredness, a stock dummy[22], the interaction between stock dummy and industry-level merger clusteredness, and the measure of overall clusteredness. The measure of overall clusteredness is defined as the number of *all* mergers during the event window [t-3m, t+3m] normalized by 10^{-4}. This measure is analogous in spirit to the measure of hot/cold markets in Rosen (2006). In all regression models used in this paper, I control for the acquirer's book to market ratio[23] and log of the

[22] Stock dummy, created by SDC, equals 1 if over 50% of the consideration is paid in stock.

[23] Acquirer's B/M ratio is winsorized at the 0.5% and 99.5% levels to lessen the impact of outliers.

acquirer's market capitalization, both measured at the end of calendar year t-1.[24] I also control for industry dummies and year dummies (not reported). Standard errors are robust to industry and year clustering. Coefficients and t-statistics are shown in Table 3a. The measure of within industry clusteredness is negatively associated with stock and operating performance after the merger. In model (2), for example, a 1% increase in clusteredness leads to a 3.6% decrease in cumulative abnormal return for acquirer's stock over 365 days. Results (not reported) remain unchanged if I use the number of mergers announced from [t-3m, announcement date-1d] as an instrument variable for merger clusteredness to invalidate concerns about endogeneity. The negative relation does not change after I control for the stock dummy (-3.564 vs. -3.560). The negative and significant coefficients on the stock dummy are consistent with prior literature that documents underperformance of stock deals. Comparing model (3) with models (1) and (2), including the stock dummy does not change R^2 whereas including the clusteredness measure increases R^2 by 0.01. Moreover, the coefficient on the interaction term between the stock dummy and the clusteredness measure is indistinguishable from zero. The significance on the overall clusteredness coefficient is also weak.[25] In addition, acquirer's size is negatively associated with post-merger performance. This is consistent with Moeller, Schlingemann, and Stulz (2004)'s findings over a shorter event window. Finally, consistent with Rau and Vermaelen (1998), value acquirers (high B/M) outperform glamor acquirers (low B/M).

Table 3b repeats the baseline regressions in Table 3a on horizontal mergers defined using the 4-digit SIC codes. A finer industry classification allows us to better identify same-industry mergers as the Fama-French 48 industry classification is relatively coarse. On the other hand, using 4-digit SIC codes reduces sample size by 40% and makes the clusteredness measure noisier as there are fewer horizontal mergers at the 4-digit SIC level.[26] Nonetheless, the baseline results still hold across all specifications. It is interesting to note that the coefficient on the stock dummy is much weaker in Table 3b than in Table 3a: stock deals underperform by 4.5% over one year in model (1) of Table 3a and by only 2.0% in model

[24] Results (not reported) show that the baseline results remain unchanged if we control for (a) the trailing one-year abnormal return of the acquirer, (b) the number of "dormant" days since the last acquisition made by a firm in the given industry, as in Song and Walkling (2004), and (c) acquirer's cash holdings prior to the merger, as in Harford (1999). The coefficient on acquirer's cash holdings is negative and significant, which is consistent with Harford (1999).

[25] This is consistent with the findings in Rosen (2006), who used the number of mergers as a measure of waves and did not find a significant relationship between clusteredness and long-run performance. This result is also consistent with the conclusion in Mitchell and Stafford (2000) that there does not exist differential post-completion performance between hot- and cold-market mergers.

[26] I eliminate the $1 million threshold on deal value in constructing merger clusteredness for the 4-digit SIC classification, as small deals can be important in a finely-defined industry.

(1) of Table 3b. In the latter case, the coefficient becomes statistically insignificant. This is consistent with market timing theory: under finer industry classification, the stock price of same-industry firms is more likely to move in tandem and the method of payment is less likely to be indicative of misvaluation.

In Table 4, I regress operating performance, i.e., change in acquirer's OPINC/assets,[27] on the clusteredness measure controlling for acquirer characteristics. The coefficient on the clusteredness measure remains significant for operating performance. In the entire sample, a 1% increase in the clusteredness measure leads to a 0.4% decrease in the change in the acquirer's OPINC/assets. Change in acquirer's operating performance may not be an adequate benchmark as pre-merger operating performance should be the weighted average of acquirer and target's OPINC/assets prior to the announcement, also commonly known as "pro forma" OPINC/assets. I calculate this adjusted change in OPINC/assets for deals where accounting information on the target is available. This substantially reduces the sample size, as most of the targets in the main sample are private. The results from model (2) in Table 4 show that the relation between operating performance and merger clusteredness is not driven by the poor operating performance of targets during merger waves: after taking into account the target's pre-merger operating performance, a 1% increase in clusteredness leads to a 0.6% reduction in post-merger operating performance. Consistent with previous literature, including Heron and Lie (2002) and Healy, Palepu, and Ruback (1992), I do not find that method of payment predicts post-merger operating performance in models (1) and (2). The magnitude and significance of the coefficient on the clusteredness measure remain unchanged if I regress the levels of post-merger performance measures on the same set of explanatory variables controlling for lagged levels (see models (3) and (4)). Finally, the results strengthen when using net income measures of operating performance,[28] as shown in models (5) and (6) of Table 4: the magnitudes of the coefficient on the clusteredness measure increase to -5.2 and -10.5, respectively. Harford (2005) documents that merger waves take place during periods of high capital liquidity. The larger magnitude of underperformance

[27] This paper measures operating performance by deflating operating income by book value of assets. The common concern that book value of assets includes non-operating assets is less relevant in this setting, as taking the difference in OPINC/assets teases out most of the firm-specific mismeasurement. Moreover, the choice of M&A accounting method (pooling vs. purchasing) is unlikely to be systematically different for on-the-wave vis-a-vis off-the-wave mergers. Other choices of deflators used in the literature include market value, as in Healy, Palepu, and Ruback (1992), and sales, as in Heron and Lie (2002) and Kaplan (1989). As Barber and Lyon (1996) point out, using sales ignores changes in the productivity of assets whereas using market value ignores time variation in the discount rate and growth prospects. The disadvantages of both measures are obvious in the setting of this paper.

[28] Change in OPINC/assets (ROE) is winsorized at the 0.5% and 99.5% (2% and 98%) levels to lessen the impact of outliers.

using ROE is probably due to the fact that acquirers take advantage of low interest rates in a merger wave and finance their acquisitions with debt.

Fama (1998) and Mitchell and Stafford (2000) argue that long-term event studies or buy-and-hold analyses may suffer from a "bad model" problem.[29] It is possible that the underperformance of clustered mergers identified here is due to the poor performance of the asset pricing model during the periods of merger waves. Moreover, Schultz (2003) presents a pseudo market timing theory that explains the discrepancy in post-IPO performance between the event-time approach and the calendar-time approach. To address both of these concerns, I employ a calender-time rolling window technique that several papers have advocated.[30] At the end of each quarter, I sort the acquirers of all horizontal mergers announced during the quarter into (three) equal-sized tertiles by the degree of clusteredness. I then form a long-short portfolio that buys acquirers' stocks from the least clustered tertile and short sells those from the most clustered tertile. I hold this portfolio for 3 months starting from 3 months after the quarter end. I skip a quarter between the formation period and holding period.[31] If a stock disappears from the CRSP tape within 6 months of the announcement, value-weighted market returns are used to replace the missing monthly returns. I then use the Fama French 4-factor model to estimate the abnormal return of this rolling window portfolio:

$$\begin{aligned} & R_{acquirers\ of\ least\ clustered,\ t} - R_{acquirers\ of\ most\ clustered,\ t} \\ = \ & \alpha + \beta \left[R_{Mt} - R_{ft} \right] + sSMB_t + hHML_t + uUMD_t + \varepsilon_t \end{aligned}$$

where R_{ft} is the one-month Treasury bill rate, R_{Mt} is the monthly return on a value-weighted market portfolio of NYSE, Amex, and NASDAQ stocks, SMB is the difference between the returns on portfolios of small and big stocks (below or above the NYSE median), HML is the difference between the returns on portfolios of high- and low-BE/ME stocks, and UMD (Up Minus Down) is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios. In Table 5, I report average

[29] Loughran and Ritter (2000) counter-argue that the calendar-time portfolio approach suffers from low-power problems. If their argument holds and the calendar-time portfolio approach indeed reduces the power of rejecting the null hypothesis, it only increases the statistical significance of the results found in this paper.

[30] See, for example, Brav, Geczy, and Gompers (1997), Fama (1998), and Mitchell and Stafford (2000).

[31] I skip a quarter here for the following two reasons: first, it allows the acquirer returns to reflect information about same-industry mergers that are announced subsequently, i.e., [t+1m, t+3m]. Second, it teases out the short-run effect on stock price by temporary demand shift due to merger arbitrage, as in Mitchell, Pulvino, and Stafford (2004). Without skipping a quarter, results (not reported) remain unchanged under value weights and become statistically weaker under equal weights.

(monthly) returns, Fama-French 4-factor alpha (monthly), and statistical significance of the alpha for 3 samples: all cash deals, deals with non-zero stock payment (the complement of all cash deals), and all deals. Table 5 shows results for two specifications: value-weighted and equal-weighted. The equal-weighted (value-weighted) long-short portfolio generates a monthly return of 1.4% (1.1%) and an α of 1.1% (1.1%), which is statistically significant at 5% (10%). Moreover, the two mutually exclusive subsamples (cash deals and stock deals) have similar performance: 1.1% vs. 1.4% under equal weights and 1.3% vs. 1.1% under value weights. Consistent with earlier evidence, cash deal acquirers on average outperform stock deal acquirers in this data sample. Nonetheless, the relation between underperformance and clusteredness is a distinct trend independent of the method of payment.

It is important to note that even though the calendar-time approach rejects the null hypothesis that there is no differential performance between on-the-wave and off-the-wave mergers, these results by no means reject market efficiency because in forming the long-short portfolios, ex post data is used to construct the clusteredness measure. Therefore, this is not an implementable trading strategy.

3.2.2 Alternative Interpretations

The baseline results indicate that on-the-wave mergers underperform off-the-wave mergers. This newly identified empirical phenomenon draws several alternative interpretations as other existing theories and hypotheses deliver predictions of the same stylized fact. In this section, I examine the possibility of each of these alternatives.

Hypothesis I: Acquirer stocks underperform following on-the-wave mergers because acquirers overpay for targets during merger waves.

It is well documented that acquirers often pay substantial premia above targets' market price. Roll (1986) first proposed that hubris may lead acquirer managers to overpay for targets. More recently, Moeller, Schlingemann, and Stulz (2004) and Malmendier and Tate (2005) provide evidence in support of the hubris hypothesis. Is it possible that acquirers' stocks perform poorly because acquirer managers overpaid for the targets during merger waves? To answer this question, I examine the combined return of acquirer and target stocks for the subsample of mergers in which targets are public. I sort the subsample into quintiles by the clusteredness measure. In Panel A of Table 6, I report for each quintile acquirer and target's combined returns, defined as a value-weighted portfolio of acquirer and target over 365 days.[32] This equals the return of a hypothetical shareholder who owns

[32]That is to hold a value-weighted portfolio of acquirer and target stocks from day t-1 to day t+10 and

both companies prior to the merger. If the stock underperformance is due to overpaying for targets, the decreasing trend shown in Table 2 should weaken or disappear when using these combined measures. Nonetheless, the difference in acquirer and target's CAPM alpha (OPINC/assets) between the two extreme quintiles is 12.6% (1.9%).[33] In addition, the regression results shown in Panel B of Table 6 confirm that acquirer and target's combined return decreases significantly as the clusteredness measure increases.[34] And the results are largely the same if the 4-digit SIC code classification is used to identify horizontal mergers (see models (5) and (6)). Finally, two additional tests (results not reported) failed to support Hypothesis I. First, if acquirer CEOs are more confident during merger waves, then one should expect the merger premium paid on the target to be higher for highly clustered mergers. In the data, however, the relation between clusteredness and merger premium is negative and insignificant. Second, the magnitude and significance of the coefficient on the clusteredness measure remain unchanged after controlling for the CEO overconfidence measure created by Malmendier and Tate (2005). The coefficient on the interaction term between the overconfidence dummy and clusteredness is indistinguishable from zero, which indicates that the relation between clusteredness and performance is independent of whether or not the acquirer's CEO is overconfident.[35]

Hypothesis II: On-the-wave mergers are followed by poor performance because in a merger wave, managers engage in unprofitable defensive mergers to retain their private benefits of control.

Gorton, Kahl, and Rosen (2005) suggest that value-destroying merger waves may take place in a rational setting in the presence of private benefits of control. Is it possible that the value-destroying merger waves are defense-driven? To measure the severity of the agency cost in each merger, I obtain the acquirer CEO's percentage ownership from Thomson Financial Insider Trading data.[36] The criteria of non-missing ownership reduces

---

to hold acquirer's stock from t+11 to t+365. This is similar in spirit to the measure for synergy gains suggested by Bradley, Desai, and Kim (1988).

[33] To rule out the possibility of sample selection bias, results (not reported) show that the negative relation between acquirer's CAR and clusteredness for the overall sample (used in Table 2 and Table 3a) remains unchanged for the subsample of deals with public targets (used in Table 6).

[34] Another way to control for the premium paid on the target is to examine the acquirer's long-run return excluding the annnouncement event window [t-1d, t+10d]. Regressing acquirer's cumulative abnormal returns over [t+10d, t+365d] yields a negative and significant coefficient on the clusteredness measure (untabulated).

[35] Using the "longholder", i.e., those who hold their vested stock options until the year of expiration, measure reduces the sample size to 353. Industry dummies are not included for this specification due to the substantially smaller sample size.

[36] The remaining sample contains only CEOs who traded their shares during the event window. Although

the sample by 40%. I stratify the remaining sample by the percentage of ownership by the acquirer CEO. Panel A of Table 7 shows the univariate results for the subsample where percentage ownership is greater than 0.5%, the median value for the percentage of ownership in the sample. The pattern of relative underperformance by clustered mergers becomes even stronger than in Table 2.[37] The difference in one-year stock performance rises from 14.5% to 23.0%. This indicates that high managerial ownership exacerbates the underperformance of clustered mergers. In Panel B of Table 7, I use regression analysis to examine the effect of CEO ownership on the relation between merger performance and clusteredness. The coefficient on the clusteredness measure is negative and significant for all four performance measures. The positive coefficient on shares held by the CEO in model (1) and (2) implies that higher acquirer managerial ownership leads to better acquisitions for shareholders. This is consistent with the findings in Lewellen, Loderer, and Rosenfeld (1985) and supports the agency theory of mergers. However, the coefficient on the interaction term carries a negative sign for both stock performance measures and is significant for acquirer's 2-year abnormal return. The sign of the coefficient is negative yet indistinguishable from zero for the operating performance measure. Therefore, even though the regression results support agency theory in general, they do not indicate that agency theory explains the relative underperformance of clustered mergers as Gorton, Kahl, and Rosen (2005) suggest.

Hypothesis III: On-the-wave mergers are followed by poor stock performance because merger waves are triggered by over-valuation waves.

Shleifer and Vishny (2003) and Rhodes-Kropf and Viswanathan (2004) predict that mergers announced during overvaluation waves tend to have poor post-merger performance. This alternative hypothesis may potentially drive the baseline results in two distinct ways. The measure of clusteredness and the method of payment are positively correlated in the merger sample: Table 2 shows that stock deals account for 26% of the most clustered quintile vis-a-vis 15% of the least clustered quintile of horizontal mergers. More importantly, merger waves may be driven by market overvaluation in the industry. Therefore, even cash deals

not prohibited by the SEC, insider sales around merger announcement are less common due to firms' self-imposed restrictions. This sample selection issue should not lead to any bias one way or the other. Indeed, results (not reported) verify that the main deal characteristics are the same for the subsample with non-missing CEO ownership data from the TFN. To lessen the impact of outliers, I winsorize the percentage shares held at 0.5% and 99.5% levels. The results without winsorization (not reported) are stronger. Alternatively, I obtain CEO ownership from COMPUSTAT ExecutiveComp. The results (not reported) are qualitatively similar with weaker statistical significance due to much smaller sample size.

[37]It is also stronger than the results (not reported) using a subsample with non-missing CEO ownership data from TFN.

in a merger wave may be followed by poor stock performance.[38] To address these two issues, I first stratify the entire sample by the method of payment and public status of the target and focus on deals in which 100% of the payment is made in cash and the target is private. These mergers are not subject to the premise of market timing theory: acquirer's overvaluation or target's undervaluation.[39] [40] To better prevent the clusteredness measure from being influenced by valuation waves, I use the concentration of all-cash deals instead of that of all deals to measure clusteredness. The clusteredness of all-cash deals is defined as the number of all-cash deals[41] over the event window [t-3m, t+3m] divided by the total number of all-cash deals over the entire sample period. The choice of methodology here ensures that both the sample of mergers and the clusteredness measure are not related to stock market valuation. The univariate results sorted on the new clusteredness measure are shown in Panel A of Table 8. Consistent with the previous findings, these all-cash and private-target deals on average have better post-announcement performance than the overall sample: the average CAR over 365 days is 7.5% in this subsample compared to 4.5% in the overall sample. Moreover, acquirer's stocks in all 5 quintiles have positive CAR over 365 days. The difference in stock performance between the two extreme quintiles remains economically significant. Acquirer's CAR over 365 days decreases from 14.5% for the least clustered quintile to 6.3% for the most clustered quintile (although not uniformly). While the difference for the subsample of cash deals (8.2%) is smaller than the difference for the whole sample shown in Table 2 (12.5%), it stays economically and statistically significant. The univariate results indicate that the market timing theory cannot entirely explain away the trend documented in Table 2. The regression approach confirms the finding. The point estimate of the coefficient for clusteredness in model (1) in Panel B of Table 8 (-2.07) is substantially smaller in magnitude than that in model (2) of Table 3a (-3.6). It is possible that the relation between stock performance and clusteredness is partially driven by stock

[38] Market timing theory predicts that mergers announced during stock market overvaluation waves should be stock deals. The argument here allows the possibility that such deals are paid in cash for other reasons.

[39] If one assumes that private companies can be misvalued and clustering is due to undervaluation, the market timing theory as in Shleifer and Vishny (2003) implies a positive relation between clusteredness and performancce: A merger wave driven by undervaluation of the target stocks leads to overperformance of the acquirer stocks. This theoretical possibility therefore predicts the opposite.

[40] Harford (1999) finds that cash-rich firms tend to make value-destroying acquisitions. It is possible that mergers on the wave involve acquirers with better access to financing in capital market. To rule out this alternative interpretation, I also exclude all acquirers who issued public debt or equity from day t-365 to day t+365.

[41] All-cash deals satisfy the following two conditions: (a) the entire consideration is paid in the form of cash; (b) the acquirer did not issue any public stocks or bonds from one year prior to the merger announcement to one year afterwards.

market misvaluation.[42] Nonetheless, the negative relation remains strong for the subset of deals outside the scope of the market timing theory: for each 1% increase in the clusteredness measure, acquirer's stock performance over 365 days drops by 1.52% (see model (2)). In contrast, the relation between operating performance and clusteredness hardly weakens in the subset of cash deals: the magnitude of the coefficient in model (3) of Table 8 (-0.33) is reduced by less than a quarter from model (1) of Table 4 (-0.39). These results indicate that while market timing theory is supported in the data, it can at best explain away a fraction of the relation between the clusteredness and post-merger performance.

Consistent with existing evidence (see Harford (2005)), the clusteredness measures using cash deals and overall deals are highly correlated in the sample (correlation of 0.81 using the Fama-French 48 industry classification). Is it possible that the relation between clusteredness of cash deals and performance shown in Table 8 is simply driven by misvaluation waves, an omitted variable in the regression? Or, to what extent is the clusteredness of cash deals a proxy for the clusteredness of stock deals, hence the market misvaluation? To address this potential omitted variable problem, I create a clusteredness measure using only the stock deals, i.e., the complement of the previously defined all-cash deals. If I regress the post-merger stock performance on both clusteredness measures, the market timing theory predicts that only the high clusteredness of stock deals leads to poor post-merger performance, whereas this paper's theory model predicts that both coefficients have negative signs, because both measures proxy for the degree of competition effect. Panel C of Table 8 shows the horse-race regression results. In model (4), the clusteredness measures of both cash and stock deals have a significant and negative effect on post-merger stock performance. They weaken each other's explanatory power (compared to the coefficients in models (2) and (3)) possibly because the two measures are highly correlated (correlation of 0.61). The results using the 4-digit SIC codes, as shown in models (5) through (8), indicate that the coefficients on both clusteredness measures are significant and negative. In contrast to using the Fama-French 48 industry classification, the two variables do not weaken each other's explanatory power, as the correlation between the two clusteredness measures becomes much lower (0.00) in this finer industry classification. Overall, the results indicate that both clusteredness measures have a similar impact on post-merger performance, which is less consistent with the market timing theory.

To further study the possibility of the market timing theory as an alternative explanation, I look for proxies for perceived misvaluation by acquirer managers. It is well established

[42] An alternative interpretation of the weakened coefficient is that the cash deal clusteredness measure is noisier due to fewer cash deals.

that insider trading can forecast returns.[43] Firms whose shares have been intensively sold (bought) by insiders tend to underperform (overperform) benchmarks in subsequent periods. Following Jenter (2005), I use insider trading as a window into managers' perception of their stocks being over- or under-valued. I use data collected by Thomson Financial from the required SEC insider trading filings. For each merger, I sum all (split-adjusted) open market transactions for all insiders[44] over [t-365d, t-1d], with sales entering positively and purchases entering negatively. I create a dummy variable for net insider sales that equals one if the net sales of all insiders during the year prior to the merger announcement is positive and zero otherwise. Broadly speaking, the insider sales dummy indicates whether or not the acquirer stock is perceived to be overvalued. To examine whether the relation between clusteredness and post-merger performance is due to market timing, I regress the post-merger stock performance on the dummy of insider sales, merger clusteredness, and the interaction term of the two. In models (1) and (2) in Table 9, the negative coefficient on the dummy of net insider sales is consistent with earlier findings that insider trades are informative. It also validates the choice of proxy for market misvaluation. If the market timing theory explains the baseline results, the negative relation between performance and clusteredness should be more pronounced for overvalued acquirers, as indicated by the insider sales dummy. However, the coefficient on the interaction term is positive in both models and significant for the 2-year abnormal return. This indicates that the relation between performance and clusteredness is not explained by the misvaluation of acquirer stock during merger waves. Finally, by using the 4-digit SIC industry classification, I reach the same conclusion (see models (3) and (4)). Therefore, evidence from insider trading does not provide support for the hypothesis that the market timing theory explains the baseline findings.

Hypothesis IV: Negative industry-wide shocks drive merger waves, hence the underperformance following on-the-wave mergers.

One last alternative interpretation is that merger waves are procyclical, which implies that horizontal merger waves occur when the industry starts declining. This alternative neoclassical hypothesis suggests that the negative relation between post-merger performance and clusteredness is due to industry-wide demand shocks, rather than the competition effect as this paper's model suggests. To what extent are the baseline results driven by the variation

[43] See Seyhun (1998) for a comprehensive review of this literature and a discussion of SEC rules, filing requirements, and available data.

[44] The results remain the same if (a) only transactions by officers and directors are included or (b) the absolute level of insider trades (normalized by the number of shares outstanding) is used. The results are statistically weaker if only CEO trades are included. This is largely due to a substantial reduction in sample size.

in industry condition vis-a-vis competition among firms? To empirically disentangle these two potentially complementary effects is challenging, because industry-wide demand shocks are not directly observable.

Gomes and Livdan (2004) and Maksimovic and Phillips (2002) suggest that it is optimal for firms to make diversifying mergers when there is a negative shock to the incumbent industry. Therefore, one should observe diversifying mergers following negative industry-wide demand shocks. In the spirit of this neoclassical theory prediction, I create additional clusteredness measures using diversifying mergers to proxy for variation in industry condition over time. The clusteredness measure for deals where the acquirer (target) is in the given industry is defined as the number of diversifying mergers during the event window [t-3m, t+3m] that involve firms in the given industry as acquirer (target) divided by the total number of diversifying mergers over the sample period (1979-2004) that involve firms in the given industry as acquirer (target).[45] Broadly speaking, if the clusteredness of diversifying mergers is high, the industry condition is poor. In Table 10, I regress post-merger performance on horizontal merger clusteredness, diversifying merger clusteredness (acquirer), and diversifying merger clusteredness (target), controlling for other variables. The coefficient on diversifying merger clusteredness (where target is in this industry) is negative and significant. This indicates that the clusteredness of diversifying mergers explains the post-merger performance of horizontal mergers. Moreover, after controlling for the two measures of cross-industry clusteredness, the coefficients of horizontal merger clusteredness are reduced by half in magnitude (from -3.6 in model (1) to -2.0 in model (2) and from -4.9 in model (3) to -2.3 in model (4)), suggesting that a fraction of the explanatory power of the clusteredness measure can potentially be attributed to variation in industry condition. However, the coefficient on the horizontal merger clusteredness remains economically and statistically significant. The same results (not reported) hold if we use the 4-digit SIC industry classification.

The results indicate that (a) the clusteredness measures of cross-industry mergers explain post-merger stock performance for horizontal mergers, probably because they proxy for changes in industry condition as the existing literature suggests; and (b) after controlling for the changes in industry condition, the original clusteredness measure still explains (albeit to a lesser degree) post-merger performance. To the extent that diversifying mergers may be driven by efficiency-enhancing horizontal merger waves, my regression results are understating the explanatory power of horizontal merger clusteredness. Therefore, it is conservative to conclude that half of the explanatory power of horizontal merger clusteredness remains

[45] All mergers that involve financial or utility firms are excluded.

after controlling for variation in industry condition.

3.2.3 Post-merger Performance of Rival Firms

The theory model also predicts that merger waves may lower the value of stand-alone industry rivals, thus leading to poor performance of the entire industry. To test this prediction, I examine the relation between clusteredness and market and industry performance following the merger. Model (1) in Table 11 shows that value-weighted market index does not underperform following clustered mergers relative to unclustered mergers, which indicates that market-wide misvaluation is unlikely to be the driving force for these horizontal merger waves. Model (2) shows that industry returns are lower following clustered mergers. A 1% increase in clusteredness measure decreases industry return over 365 days by 2.7%. This is consistent with both my model and the market timing theory. Nonetheless, if industry-specific misvaluation drives the horizontal merger waves, as Rhodes-Kropf and Viswanathan (2004) suggest, the industry returns should be lower following stock deals than cash deals. However, model (2) in Table 11 does not support that argument: the coefficient on the stock dummy is negative but insignificant. I obtain similar results by using abnormal returns over 730 days.

If both the acquirer as well as the entire industry lose value following merger waves, does the acquirer underperform relative to the industry? The answer, according to models (4) and (5), is yes. An increase of 1% in clusteredness is associated with a 1.7% decrease in acquirer return relative to the industry. While this result is consistent with the market timing theory, it can also be consistent with an extension of this paper's theory model. First of all, the Fama-French 48 industry portfolios include companies that are not direct competitors of the merging firms in the product market. Therefore, the underperformance of industry portfolios may understate the negative externalities on non-merging rival firms during a merger wave. More importantly, the externalities of a horizontal merger wave are not evenly distributed across all rival firms. In the model extension discussed in the Appendix, firms that participate in mergers are also the firms that absorb the most negative externalities from rivals' mergers. Consistent with this extension, the magnitude of relative underperformance is further reduced when I control by a size and book to market matched firm as in Barber and Lyon (1997).[46] In model (7) where I adjust the acquirer's return by size

[46] Barber and Lyon (1997) suggest that test statistics using cumulative abnormal returns, i.e., using the market as the reference portfolio, are subject to a new listing bias. In this case, the positive new-listing bias is likely to be stronger for mergers on the wave, because empirically stock issues and mergers are positively correlated. Therefore, correcting such a bias may only strengthen the negative relation between

and B/M matched-firm's return, acquirer's buy and hold abnormal return has an insignificant loading on the clusteredness measure.[47] Note that the coefficient on the stock dummy remains significant, which is consistent with the market timing theory that idiosyncratic overvaluation of acquirer stock predicts negative post-merger stock performance.

This paper proposes a new motivation for the firm-match buy-and-hold abnormal return method in Barber and Lyon (1997). Controlling for firm-characteristics such as size and book to market is one way to provide a proxy for the stand-alone value *conditional* on the mergers of rival firms. Current empirical literature draws conclusions on mergers' impact on firm value by examining each individual merger's performance. Instead of comparing post-merger value and pre-merger value *conditional* on rivals' actions, the conventional event studies have been examining the *unconditional* changes in value measured by stock and operating performance following the merger announcement. While this method is adequate for periods when few mergers happen, it becomes problematic during merger waves. As this paper's model suggests, during a horizontal merger wave, stand-alone value of two firms conditional on rivals' merger actions can be rather different than the unconditional firm value. More specifically, when a large number of productivity-enhancing mergers take place within an industry, stand-alone firms' values may decrease significantly from their pre-wave levels. Therefore, even though mergers appear to be value-destroying using conventional empirical methods, it is not necessary that the value-maximization principle should be violated. Using the industry-size-B/M-matched control firm method, the test fails to reject the hypothesis that mergers on the wave destroy shareholder value *conditionally* relative to off-the-wave mergers.

3.2.4 Cross-industry Comparison

In the baseline model, there is only one industry. In a world with multiple industries with different characteristics (assumed to be exogenous), the model predicts that (ex ante) highly concentrated industries are less subject to value-destroying merger waves driven by competition. As Corollary 3 states, ceteris paribus, the complementarity threshold for merger waves to be value-destroying is higher for concentrated industries with fewer numbers of competitors. The Bureau of Census conducts quinquennial plant-level surveys on manufacturing firms in the US, where I obtain the number of companies and industry concentration measured by HHI for each 4-digit SIC manufacturing industry. For each horizontal merger,

clusteredness and post-merger performance.

[47] The same results (not reported) hold for horizontal mergers defined using the 4-digit SIC industry classification.

the most recent HHI reported prior to the announcement is used as the ex ante industry concentration measure.

To examine the effect of industry concentration on the relation between clusteredness and performance, I regress the stock performance measures on the clusteredness measure, concentration measure, and the interaction term between the two. This paper's model predicts a positive sign on the interaction term. The results in Table 12 indeed verify the prediction. Like in previous tests, the coefficient on merger clusteredness is significant and negative. The coefficient on the interaction term is significantly positive. The results indicate that while merger waves on average lower shareholder value, firms in highly concentrated industries have relatively better performance following the merger waves than those in competitive industries.[48]

The model also predicts that merger waves are less likely to lower shareholder value when the price elasticity of demand is low. While the price elasticity of demand ($\frac{1}{\gamma}$) is not easily observable, the industrial organization literature provides us with a theoretically motivated yet convenient empirical proxy: Klemperer (1987a, b) shows that the noncooperative equilibrium in oligopoly with consumer switching costs resembles the collusive outcome in an otherwise identical market without consumer switching costs. The implication is that the switching costs in effect decrease the price elasticity of demand in the industry. Therefore, it follows from Corollary 4 that $\underline{\delta}$ is higher for industries with high consumer switching costs, ceteris paribus.[49] Following Cremers, Nair, and Peyer (2006), I use the 2 digit SIC classification to characterize industries into relationship and non-relationship industries. I label industries that operate in the service sector or the durable goods sector as relationship industries. Generally speaking, industries are relationship-based if they not only produce goods but also provide services through ongoing relationships with consumers or clients. As Klemperer (1987a, b) suggests, relationship industries, i.e., industries with high consumer switching costs, have lower price elasticity of demand. In the Cournot setting of this paper, low price elasticity of demand leads to a high minimum threshold of productive synergies for a merger wave to lower private values of shareholders. Therefore, the model predicts that the relation between clusteredness and performance is likely to be weakened in relationship

[48] The results are qualitatively similar with weaker statistical significance if the combined return of acquirer and target stocks is used, due to a reduction in sample size.

[49] Klemperer (1987a, b) uses a Bertrand framework to demonstrate the effect of consumer switching cost on price elastiticity of demand. The key intuition is similar in the Cournot setting used in this paper.

Klemperer (1987a, b) also argues that in addition to the tacit collusion, consumer switching costs also lead to ferocious competition for market share before consumers attach themselves to suppliers. Since this paper studies the change in firm value following merger events, the pre-merger firm value already takes this pre-attachment competition effect into account.

industries.

To test this prediction, I regress performance on the relationship industry dummy, the clusteredness measure, and the interaction term between the two. The results are shown in Table 13. The theory model predicts that the coefficient on the interaction term is positive. Indeed, the interaction term has a significant and positive coefficient. This implies that while horizontal merger waves tend to lower shareholder value, they do so to a lesser degree in relationship industries.[50]

4 Related Literature

Mitchell and Mulherin (1996) and Andrade, Mitchell, and Stafford (2001) show that horizontal merger waves have been a consistent feature during the past century. There is also an extensive literature that examines the post-merger stock and operating performance to answer a variety of questions. This paper is the first to empirically document the relation between clusteredness and post-merger performance of horizontal mergers. Harford (2005) and Rosen (2006) both studied the relationship between merger wave and post-merger performance. Both found inconclusive results on the performance of on-the-wave mergers relative to off-the-wave mergers. In contrast, I identify a negative relation between clusteredness and post-merger performance within the subset of horizontal mergers. Horizontal mergers are unique due to synergies in production technology, a key ingredient in this paper's theory model. It is therefore of little surprise that this paper draws different conclusions by focusing on horizontal mergers.

Using a shorter event window, Eckbo (1983), Song and Walkling (1998), Fee and Thomas (2004), and Shahrur (2005) all empirically document positive abnormal returns on rival firm stocks upon announcement of horizontal mergers. By using a longer event window, this paper finds that rivals' reaction to horizontal mergers varies depending on the concentration of contemporaneous mergers. Merger waves that are driven by technological synergies are likely to lower industry rivals' value due to productive efficiency gains.[51] While the existing studies have not found negative post-merger performance on rivals in support of the productive efficiency gains hypothesis, this paper complements this literature on both theoretical and empirical fronts: Off-the-wave mergers driven by fixed cost savings are predicted

[50] The results are robust to (a) excluding business services (Fama-French Industry #34) from the non-relationship industries and/or including financials (Fama-French Industry #44-47) in the relationship industries and (b) using the combined return on acquirer and target.

[51] This is consistent with Banerjee and Eckard (1998).

to have positive externalities on rivals, which attenuate or offset the negative externalities arising from on-the-wave mergers driven by technological synergies. By sorting on clusteredness, this paper empirically identifies mergers that are caused by production efficiency gains. Moreover, the existing productive efficiency hypothesis has failed to reconcile with the positive comovement between merging firms and rivals observed in the data.[52] This paper's model provides an explanation why high production synergies can lead to poor post-merger returns for both the merging firms and the rivals.

In the debate whether there is differential stock performance following different types of mergers, Loughran and Vijh (1997) find that stock deal acquirers underperform cash deal acquirers whereas Rau and Vermaelen (1998) find that glamor acquirers (high M/B) underperform value acquirers (low M/B). In contrast, Mitchell and Stafford (2000) counter-argues that the long-run abnormal performance goes away after using the correct statistical inference or using the value-weighted calendar-time portfolio approach suggested by Fama (1998). This paper empirically identifies and theoretically rationalizes a pattern in post-merger performance related to clusteredness. The difference in performance between on-the-wave and off-the-wave mergers is robust to using both event study and the calendar-portfolio approach that Mitchell and Stafford (2000) and Fama (1998) advocated.

In the debate whether poor post-merger performance implies violation of value maximization, agency theory (Gorton, Kahl, and Rosen (2005)) and behavioral hypotheses (Moeller, Schlingemann, and Stulz (2004) and Malmendier and Tate (2005)) argue that the value maximization principle is violated, whereas the market timing theory (Shleifer and Vishny (2003) and Rhodes-Kropf and Viswanathan (2004)) suggests that negative post-merger stock performance need not be interpreted as violation of value maximization if mergers are driven by misvaluation in the stock market. More recently, Savor (2006) documents poor stock performance of failed stock mergers relative to successful stock mergers, henceforth arguing that stock mergers, such as AOL-Time Warner, create value for the acquirer's shareholders by taking advantage of mispricing on the market. In contrast, this paper takes a new approach to reconcile negative post-merger performance with value maximization. Introducing imperfect product market competition allows neoclassical theory to explain why poor stock and operating performance follows horizontal merger waves. I further show in the empirical section that this specific pattern of underperformance does not seem to be driven by market timing theory. The method of payment does not explain the post-merger operating

[52] The anticipation hypothesis, as in Song and Walkling (1998, 2004), gives a solution to this puzzle. However, it is unlikely that the short-lived anticipation effect can help explain the positive comovement between merging firms and rivals' long-run performance, as documented in this paper.

performance, whereas the measure of clusteredness does. In addition, while the method of payment does predict the post-merger stock performance, it does not strengthen or weaken the relation between clusteredness and post-merger stock performance.

5 Conclusion

While neoclassical theory has had success in explaining what drives mergers, it has had little success in reconciling the mixed empirical evidence on post-merger stock and operating performance. By introducing imperfect product market competition into the existing neoclassical framework, this paper enriches its predictions. More specifically, the value-maximization assumption and negative post-merger stock performance need not be contradictory with each other in a horizontal merger wave. Empirically, this paper shows that on-the-wave horizontal mergers are followed by poor stock and operating performance. This newly identified relation between performance and clusteredness cannot be explained away by existing theories in the literature. The relation is consistent with the product market competition explanation, which is further supported by the following additional evidence: (a) merger waves not only lower value of merging firms but also lead to poor performance of industry rivals; (b) merger waves in highly concentrated industries destroy less shareholder value; and (c) merger waves in relationship industries destroy less shareholder value.

This paper focuses on horizontal mergers, which comprise half of all the mergers in the SDC universe. The exclusion of diversifying mergers is due to the lack of complementarity between firms operating in unrelated industries. The link between merger waves and post-merger performance for diversifying mergers remains an open question for future research.

6 Appendix

6.1 Proof for Proposition 1:

Proof. To simplify notation, I normalize C^h to 1 and let $l = \frac{C^l}{C^h}$, $\theta = \frac{C^{hl}}{C^h}$, (then δ becomes $l(\frac{1}{\theta} - 1)$), $C = C^h$, and $\rho(x) = \frac{(N-x-\gamma)}{((\frac{N}{2}-x)(1+l)+x\theta)}$. The profit function simplifies to

$$\pi^h = \frac{1}{\gamma} C^{1-\frac{1}{\gamma}} \rho(x)^{\frac{1}{\gamma}-1} (1-\rho(x))^2$$

$$\pi^l = \frac{1}{\gamma} C^{1-\frac{1}{\gamma}} \rho(x)^{\frac{1}{\gamma}-1} \left(1 - l\rho(x)\right)^2$$

$$\pi^{hl} = X C^{1-\frac{1}{\gamma}} \rho(x)^{\frac{1}{\gamma}-1} \left(1 - \theta\rho(x)\right)^2$$

subject to the following regularity conditions,[53]

$$\forall x \in [0, \frac{N}{2} - 1]: \; \rho(x) l < 1 \tag{7}$$

From (6), we have

$$\pi^h = \frac{1}{\gamma} C^{1-\frac{1}{\gamma}} \rho(x)^{1-\frac{1}{\gamma}} \left(1 - \rho(x)\right)^2$$

$$\begin{aligned}
\frac{\partial \pi^h}{\partial x} &= \frac{1}{\gamma} C^{1-\frac{1}{\gamma}} \frac{\partial \left[\rho(x)^{\frac{1}{\gamma}-1} \left(1 - \rho(x)\right)^2\right]}{\partial x} \\
&= \frac{1}{\gamma} C^{1-\frac{1}{\gamma}} \frac{\partial \rho}{\partial x} (1-\rho) \rho^{\frac{1}{\gamma}-2} \\
&\quad \left[\left(\frac{1}{\gamma} - 1\right)(1-\rho) + 2(-1)\rho\right] \\
&= \frac{1}{\gamma} C^{1-\frac{1}{\gamma}} \frac{(1+l)(\frac{N}{2} - \gamma) - \theta(N-\gamma)}{((\frac{N}{2} - x)(1+l) + x\theta)^2} (1-\rho) \rho^{\frac{1}{\gamma}-2} \\
&\quad \left[\frac{1}{\gamma} - 1 - \rho - \frac{1}{\gamma}\rho\right]
\end{aligned}$$

Note that

1. $\theta = 1$, $\rho(0) l = \frac{(N-x-\gamma)l}{((\frac{N}{2}-x)(1+l)+x\theta)} = \frac{(N-\gamma)l}{(\frac{N}{2})(1+l)} < 1 \Longrightarrow (\frac{N}{2} - \gamma) l - \frac{N}{2} < 0 \Longrightarrow (1+l)(\frac{N}{2} - \gamma) - (N - \gamma) < 0 \Longrightarrow \frac{\partial \rho}{\partial x} < 0$ for all x.

2. One can show that $\frac{1}{\gamma} - 1 - \rho\left(\frac{N}{2} - 1\right) - \frac{1}{\gamma}\rho\left(\frac{N}{2} - 1\right) < 0$ for any l that satisfies (4) when $\theta = 1$ as long as $N \geq 2$.

Let $g(l) = \frac{1}{\gamma} - 1 - \rho\left(\frac{N}{2} - 1\right) - \frac{1}{\gamma}\rho\left(\frac{N}{2} - 1\right) = -(1 + \frac{1}{\gamma}) \frac{\frac{N}{2}+1-\gamma}{(1+l)\frac{N}{2}} - 1 + \frac{1}{\gamma}$. $g(l)$ is monotone increasing in l.

$\rho(0) l < 1 \Rightarrow l < \frac{\frac{N}{2}}{\frac{N}{2} - \gamma}$. Therefore, $g(l) < g\left(\frac{\frac{N}{2}}{\frac{N}{2}-\gamma}\right) = \left(\frac{N}{2}\right)^{-1}(1 + \gamma - N) < 0$ for any γ, if $N \geq 2$.

3. Since $\rho(x)$ is decreasing in x, $\frac{1}{\gamma} - 1 - \rho(x) - \frac{1}{\gamma}\rho(x) < 0$ for all x.

Hence, $\frac{\partial \pi^h}{\partial x} > 0$ for all x. Analogously, $\frac{\partial \pi^l}{\partial x} > 0$ for all x. $\Longrightarrow \pi^h(\frac{N}{2} - 1) + \pi^l(\frac{N}{2} - 1) > \pi^h(0) + \pi^l(0) \Longrightarrow \pi^{hl}(\frac{N}{2}) - I > \pi^h(0) + \pi^l(0)$. ■

[53] In the case where $\rho(x) l > 1$ for some x, low technology firms would become unprofitable and will be forced to exit. This set of conditions simplifies the proof of the propositions and is by no means crucial.

Note:

In the baseline model, there are only two types of firms, high and low. If each firm has different technologies, e.g., $\{C_i\}_{i=1}^N$, Proposition 1 still holds. In other words, in a Cournot setting, a merger that does not change production technologies of remaining firms always benefit the rivals.

To show this mathematically, we know from the model set-up that in the standard Cournot setting, $\frac{q^i}{Q_2} = \frac{\bar{C}-(1-\frac{\gamma}{N-x})C^i}{\gamma\bar{C}} = \frac{1}{\gamma} - \frac{C^i(N-x-\gamma)}{\gamma(N-x)\bar{C}}$. Let $j, k(C_j > C_k)$ be the pair of companies that merged. All we need to show is that $\frac{(N-x-\gamma)}{(N-x)\bar{C}} < \frac{(N-x-1-\gamma)}{(N-x-1)\bar{C}'}$, where $\bar{C}'$ denotes the average cost of production after the merger. This is equivalent to showing that $\frac{(N-x-\gamma)C^j}{(N-x)\bar{C}'} < 1$, which holds trivially from regularity condition (4).

6.2 Proof for Proposition 2:

Proof. From the previous proof, we know

$$\begin{aligned}\frac{\partial \pi^h}{\partial x} &= \frac{1}{\gamma}C^{1-\frac{1}{\gamma}}\frac{(1+l)(\frac{N}{2}-\gamma)-\theta(N-\gamma)}{((\frac{N}{2}-x)(1+l)+x\theta)^2}(1-\rho)\rho^{\frac{1}{\gamma}-2} \\ &\quad \left[\frac{1}{\gamma}-1-\rho-\frac{1}{\gamma}\rho\right]\end{aligned}$$

Note that

1. $\theta < \frac{\left(\frac{N}{2}-\gamma\right)(1+l)}{N-\gamma}, \Longrightarrow (1+l)(\frac{N}{2}-\gamma)-(N-\gamma)\theta > 0 \Longrightarrow \frac{\partial\rho}{\partial x} > 0$

2. One can show $\frac{1}{\gamma}-1-\rho(0)-\frac{1}{\gamma}\rho(0) < 0$ for any l that satisfies (4) as long as $N \geq 2$.

Let $h(l) = \frac{1}{\gamma}-1-\rho(0)-\frac{1}{\gamma}\rho(0) = -(1+\frac{1}{\gamma})\frac{N-\gamma}{(1+l)\frac{N}{2}}-1+\frac{1}{\gamma}$. $h(l)$ is monotone increasing in l.

$\rho(0)l < 1 \Rightarrow l < \frac{\frac{N}{2}}{\frac{N}{2}-\gamma}$. Therefore, $h(l) < h\left(\frac{\frac{N}{2}}{\frac{N}{2}-\gamma}\right) = \left(\frac{N}{2}\right)^{-1}(1+\gamma-N) < 0$ for any γ, if $N \geq 2$.

3. Since $\rho(x)$ is increasing in x, $\frac{1}{\gamma}-1-\rho(x)-\frac{1}{\gamma}\rho(x) < 0$ for all x.

Hence, $\frac{\partial\pi^h}{\partial x} < 0$ for all x. Analogously, $\frac{\partial\pi^l}{\partial x} < 0$ for all x. $\Longrightarrow \pi^h(\frac{N}{2}-1)+\pi^l(\frac{N}{2}-1) < \pi^h(0)+\pi^l(0)$.

Therefore, there exists $I \in \left(\pi^{hl}(\frac{N}{2})-\pi^h(0)-\pi^l(0),\ \pi^{hl}(\frac{N}{2})-\pi^h(\frac{N}{2}-1)-\pi^l(\frac{N}{2}-1)\right)$ such that $\pi^{hl}(\frac{N}{2})-I < \pi^h(0)+\pi^l(0)$ but $\pi^{hl}(\frac{N}{2})-I > \pi^h(\frac{N}{2}-1)+\pi^l(\frac{N}{2}-1)$.

Finally, we need to show that for any l that satisfies $\rho(0)l < 1$, $\exists\theta$, such that $\theta < \frac{\left(\frac{N}{2}-\gamma\right)(1+l)}{N-\gamma}$ and $\rho(\frac{N}{2}-1)l < 1$ can both hold.

$\rho(\frac{N}{2}-1)l < 1 \Rightarrow \theta > \frac{(\frac{N}{2}-\gamma)l-1}{\frac{N}{2}-1}$. Therefore, we only need to show that $\frac{(\frac{N}{2}-\gamma)l-1}{\frac{N}{2}-1} <$

$\frac{(\frac{N}{2}-\gamma)(1+l)}{N-\gamma}$ for any l that satisfies $\rho(0)l < 1$.

Let $f(l) \equiv \frac{(\frac{N}{2}-\gamma)(1+l)}{N-\gamma} - \frac{(\frac{N}{2}-\gamma)l-1}{\frac{N}{2}-1}$. $f(l)$ is monotone decreasing in l as $N > 2$.

$\rho(0)l < 1 \Rightarrow l < \frac{\frac{N}{2}}{\frac{N}{2}-\gamma}$. $\Rightarrow f(l) > f(\frac{\frac{N}{2}}{\frac{N}{2}-\gamma}) = 0$. ■

6.3 Extensions

6.3.1 Mergers of Likes

In the baseline model, mergers of likes, i.e., mergers of two firms of the same technology type, are ruled out for technical simplicity. In reality, technology may not be the only determinant for the pairing of merging firms. Other factors, such as the personal relations between company managers, may also play an important role in the pairing. Therefore, mergers of likes are quite possible. In fact, Rhodes-Kropf and Robinson (2006) document that high M/B firms typical acquire high M/B targets. In this subsection, we examine the scenario where firms of the same technology type are allowed to pair and merge with one another. As in the baseline model, there are two types of firms at time 0, high technology (C^h) and low technology (C^l). Allowing merger of likes introduces 3 possible pairings of firms: high-low, high-high, and low-low. To compare the inclination to merge for different pairs of firms, a few additional assumptions are needed: If two high- (low-) type firms merge, the combined firm will have a cost of production of $C^{hh}(C^{ll})$. Without loss of generality, I normalize $C^h = 1$, then let $l = \frac{C^l}{C^h}$ and $\theta = \frac{C^{hh}}{C^h}$, where $l > 1$ and $\theta < 1$. I also assume that the degree of complementarity between two high types is the same as that of two low types, e.g., $\theta = \frac{C^{ll}}{C^l}$. For expositional simplicity, I assume integration costs to be a constant fraction of firm profit under status quo, e.g., $I^l = \iota\pi^l$ and $I^h = \iota\pi^h$.[54]

Lemma 1 *For any θ, there exists $\iota \in \{\underline{\iota}, \bar{\iota}\}$, such that two high types do not merge, e.g., $\pi^{hh} - 2\pi^h - I^h < 0$, and two low types merge, e.g., $\pi^{ll} - 2\pi^l - I^l > 0$, if $\gamma = 1$.*

The above lemma states that ceteris paribus, low-type pairs are more likely to merge than high-type pairs because the proportional value increase from a merger is greater for low-types.[55] Moreover, low type firms also absorb more negative externalities from merger waves:

Lemma 2 *If $\delta > \underline{\delta}$, $\frac{\partial\pi^l/\pi^l}{\partial x} < \frac{\partial\pi^h/\pi^h}{\partial x}$.*

[54] One can also assume $I^l = I^h$. Additional conditions will be needed and the technical complexity will considerably increase. The simplied assumptions below make the main intuitions more transparent.

[55] The proof for this lemma is available upon request.

The above Lemma states that when a merger wave brings about negative externalities on rivals ($\delta > \underline{\delta}$), low-technology firms suffer a larger fraction of value loss than high-technology firms. The intuition is that low technology firms' market share and profit margin shrink more than high technology firms do, because the merger waves move the low types closer to being unprofitable. In the extreme case, when a merger wave moves the low types to making almost zero profit, then their value loss becomes 100%.

The above two lemmas show that during a value-destroying merger wave, low-type firms (a) suffer a greater amount of value loss due to competition and (b) are more likely to merge with each other everything else equal. This leads to the following ancillary prediction: During a merger wave, merging firms underperform their stand-alone rivals, which is consistent with the empirical results in Table 11.

References

[1] G. Andrade, M. Mitchell, and E. Stafford. New evidence and perspectives on mergers. *Journal of Economic Perspectives*, 15:103–120, 2001.

[2] A. Banerjee and E. W. Echard. Are mega-mergers anticompetitive? Evidence from the first great merger wave. *The RAND Journal of Economics*, 29:803–827, 1998.

[3] B. M. Barber and J. D. Lyon. Detecting abnormal operating performance: The empirical power and specification of test statistics. *Journal of Financial Economics*, 41:359–399, 1996.

[4] B. M. Barber and J. D. Lyon. Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics*, 43:341–372, 1997.

[5] M. Bradley, A. Desai, and E. H. Kim. Synergistic gains from corporate acquisitions and their division between the stockholders of target and acquiring firms. *Journal of Financial Economics*, 21:3–40, 1988.

[6] A. Brav, C. Geczy, and P. Gompers. Is the abnormal return following equity issuances anomalous? *Journal of Financial Economics*, 56:209–249, 2000.

[7] K. M. Cremers, V. Nair, and U. Peyer. Weak shareholder rights: A product market rationale. Working Paper, University of Pennsylvania, Mar. 2006.

[8] R. Deneckere and C. Davidson. The incentive to form coalitions with Bertrand competition. *The RAND Journal of Economics*, 16:473–486, 1985.

[9] B. E. Eckbo. Horizontal mergers, collusion, and stockholder wealth. *Journal of Financial Economics*, 11:241–273, 1983.

[10] E. Fama. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49:283–306, 1998.

[11] J. Farrell and C. Shapiro. Horizontal mergers: An equilibrium analysis. *American Economic Review*, 80:107–126, 1990.

[12] R. Fauli-Oller. Takeover waves. *Journal of Economics and Management Strategy*, 9:189–210, 2000.

[13] C. E. Fee and S. Thomas. Sources of gains in horizontal mergers: Evidence from customer, supplier, and rival firms. *Journal of Financial Economics*, 74:423–460, 2004.

[14] D. Focarelli and F. Panetta. Are mergers beneficial to consumers? Evidence from the market for bank deposits. *American Economic Review*, 93:1152–1172, 2003.

[15] J. Gomes and D. Livdan. Optimal diversification: Reconciling theory and evidence. *Journal of Finance*, 59:507–535, 2004.

[16] G. Gorton, M. Kahl, and R. Rosen. Eat or be eaten: A theory of mergers and merger waves. Working Paper, University of Pennsylvania, Apr. 2005.

[17] G. Gowrisankaran. A dynamic model of endogenous horizontal mergers. *The RAND Journal of Economics*, 30:56–83, 1999.

[18] J. Harford. Corporate cash reserves and acquisitions. *Journal of Finance*, 53:1969–1997, 1999.

[19] J. Harford. What drives merger waves? *Journal of Financial Economics*, 77:529–560, 2005.

[20] P. M. Healy, K. G. Palepu, and R. S. Ruback. Does corporate performance improve after mergers? *Journal of Financial Economics*, 31:135–175, 1992.

[21] R. Heron and E. Lie. Operating performance and the method of payment in takeovers. *Journal of Financial and Quantitative Analysis*, 37:137–156, 2002.

[22] M. C. Jensen. Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review*, 76:323–329, 1986.

[23] M. C. Jensen and R. S. Ruback. The market for corporate control: The scientific evidence. *Journal of Financial Economics*, 11:5–50, 1983.

[24] D. Jenter. Market timing and managerial portfolio decisions. *Journal of Finance*, 60:1903–1950, 2005.

[25] B. Jovanovic and P. Rousseau. The Q-theory of mergers. *American Economic Review*, 92:198–204, 2002.

[26] S. Kaplan. The effects of management buyouts on operating performance and value. *Journal of Financial Economics*, 24:581–618, 1989.

[27] E. H. Kim and V. Singal. Mergers and market power: Evidence from the airline industry. *American Economic Review*, 83:1–24, 1993.

[28] P. Klemperer. The competitiveness of markets with switching costs. *The RAND Journal of Economics*, 18:138–150, 1987.

[29] P. Klemperer. Markets with consumer switching costs. *Quarterly Jounral of Economics*, 102:375–394, 1987.

[30] W. Lewellen, C. Loderer, and A. Rosenfeld. Merger decisions and executive stock ownership in acquiring firms. *Journal of Accounting and Economics*, 7:209–231, 1985.

[31] T. Loughran and J. Ritter. Uniformly least powerful tests of market efficiency. *Journal of Financial Economics*, 55:361–389, 2000.

[32] T. Loughran and A. Vijh. Do long-term shareholders benefit from corporate acquisitions? *Journal of Finance*, 52:1765–1790, 1997.

[33] V. Maksimovic and G. Phillips. Do conglomerate firms allocate resources inefficiently across industries? Theory and evidence. *Journal of Finance*, 57:721–767, 2002.

[34] U. Malmendier and G. Tate. Who makes acquisitions? CEO overconfidence and the market's reaction. Working Paper, University of Pennsylvania, Oct. 2005.

[35] K. F. McCardle and S. Viswanathan. The direct entry versus takeover decision and stock price performance around takeovers. *Journal of Business*, 67:1–43, 1994.

[36] M. Mitchell, T. Pulvino, and E. Stafford. Price pressure around mergers. *Journal of Finance*, 59:31–63, 2004.

[37] M. L. Mitchell and J. H. Mulherin. The impact of industry shocks on takeover and restructuring activity. *Journal of Financial Economics*, 41:193–229, 1996.

[38] M. L. Mitchell and E. Stafford. Managerial decisions and long term stock price performance. *Journal of Business*, 73:287–329, 2000.

[39] S. B. Moeller, F. P. Schlingemann, and R. M. Stulz. Firm size and the gains from acquisitions. *Journal of Financial Economics*, 73:201–228, 2004.

[40] J. Molnar. Preemptive horizontal mergers: Theory and evidence. Working Paper, Univeersity of Mannheim, Feb. 2006.

[41] M. Perry and R. Porter. Oligopoly and the incentive for horizontal merger. *American Economic Review*, 75:219–227, 1985.

[42] R. Prager and T. H. Hannan. Do substantial horizontal mergers generate significant price effects? *Journal of Industrial Economics*, 46:433–452, 1998.

[43] M. Rhodes-Kropf and D. T. Robinson. The market for mergers and the boundaries of the firm. Working Paper, Columbia University, July 2006.

[44] M. Rhodes-Kropf and S. Viswanathan. Market valuation and merger waves. *Journal of Finance*, 59:2685–2718, 2004.

[45] R. Roll. The hubris hypothesis of corporate takeovers. *Journal of Business*, 59:197–216, 1986.

[46] R. Rosen. Merger momentum and investor sentiment: The stock market reaction to merger announcements. *Journal of Business*, 79:987–1017, 2006.

[47] S. W. Salant, S. Switzer, and R. J. Reynolds. Losses from horizontal merger: The effects of an exogenous change in industry structure on Cournot-Nash equilbrium. *Quarterly Jounral of Economics*, 98:185–199, 1983.

[48] P. Savor. Do stock mergers create value for acquirers? Working Paper, University of Pennsylvania, Jan. 2006.

[49] P. Schultz. Pseudo market timing and the long-run underperformance of IPOs. *Journal of Finance*, 58:483–517, 2003.

[50] H. Servaes. Tobin's q and the gains from takeovers. *Journal of Finance*, 46:409–419, 1991.

[51] H. N. Seyhun. *Investment Intelligence from Insider Trading.* MIT Press, Cambridge, MA, 1998.

[52] H. Shahrur. Industry structure and horizontal takeovers: Analysis of wealth effects on rivals, suppliers, and corporate customers. *Journal of Financial Economics*, 76:61–98, 2005.

[53] A. Shleifer and R. Vishny. Stock market driven acquisitions. *Journal of Financial Economics*, 70:295–311, 2003.

[54] M. H. Song and R. A. Walkling. Abnormal returns to rivals of acquisition targets: A test of the acquisition probability hypothesis. *Journal of Financial Economics*, 55:143–171, 2000.

[55] M. H. Song and R. A. Walkling. Anticipation, acquisitions and the bidder return puzzle. Working Paper, Drexel University, July 2004.

Table 1a – Variable Definitions

Variable Name	Definition
Overall merger clusteredness	Number of *all* mergers from month t-3 to month t+3 divided by 10,000; t is the announcement month.
Industry merger clusteredness	Number of horizontal mergers from month t-3 to month t+3 divided by total number of mergers in that industry over the sample period; t is the announcement month.
Stock dummy (SDC defined)	Stock dummy is 1 if value of stock consideration is more than 50% of the total consideration of the deal and 0 otherwise.
All cash dummy	All cash dummy is 1 if value of cash consideration is 100% of the total consideration of the deal or if the acquirer is private and 0 otherwise.
%shares held by acquirer's CEO	Number of shares owned by the CEO (reported through insider trading filings dated closest to the merger announcement date) divided by the total number of shares outstanding of the acquirer firm as of calendar year t-1.
Acquirer's return [-1d, +365d]	Acquirer's cumulative return from day t-1 to day t+365; t is the announcement date.
Acquirer's CAR [-1d, +365d]	Acquirer's cumulative abnormal return from day t-1 to day t+365, where abnormal returns are estimated using the market model. β's are estimated from day t-250 to day t-11. t is the announcement date. A minimum of 100 observations are required to estimate β's.
Market return [-1d, +365d]	CRSP value-weighted index returns daily from day t-1 to day t+365; t is the announcement date.
Industry return [-1d, +365d]	Fama-French 48 industry portfolio value-weighted returns daily (for the acquirer and target's industry) from day t-1 to day t+365; t is the announcement date.
Acquirer's industry-adjusted return [-1d, +365d]	Acquirer's cumulative return [-1d, +365d] - Industry return [-1d, +365d].
Acquirer's size and B/M-adjusted return [-1d, +365d]	Acquirer's cumulative return [-1d, +365d] - Size&B/M-matched firm return [-1d, +365d]; size&B/M-matched firm is the same-industry firm with market cap within 70-130% of the acquirer and closest in B/M at December year t-1; if the matched firm's return is missing, it is replaced with market return.
Acquirer & target's return [-1d, +365d]	Return of a value-weighted portfolio of acquirer and target stocks formed at t-1d and held until t+365d; t is the announcement date.
Acquirer's abnormal return [-1d, +365d]	Acquirer's cumulative abnormal return from day t-1 to day t+365, where abnormal returns are estimated using the market model. β's are estimated from day t-250 to day t-11. t is the announcement date. A minimum of 100 observations are required to estimate the β's of acquirer and target.
Acquirer's change in OPINC/assets	Acquirer's OPINC/assets for the *first full* fiscal year after the deal's effective date - acquirer's OPINC/assets for the *last* fiscal year before the deal's announcement date. OPINC/assets is calculated as Operating Income (DATA13) divided by the sum of Depr., Depl. & Amort. (DATA 196) and Total Assets (DATA 6).
Acquirer & target's change in OPINC/assets	Acquirer's OPINC/assets for the *first full* fiscal year after the deal's effective date - "pro forma" OPINC/assets of acquirer and target; "pro forma" OPINC/assets = asset-weighted average of acquirer and target's OPINC/assets for the *last* fiscal year prior to the deal's announcement date.

Table 1b – Summary Statistics

The initial sample is obtained from the SDC Mergers & Acquisitions database with US Targets, non-LBO, non-repurchase, completed status, and dated from 1979 to 2004. I then identify horizontal mergers and define the clusteredness of horizontal mergers.

The final sample consists of deals that satisfy the following conditions: (a) public acquirers, (b) deal value (SDC reported) > $1m; (c) announced from 4/1/1979 to 9/30/2004; (d) non-financials and non-utilities.

Variable Name	Mean	Median	# of Obs
Acquirer's return [-1d, 10d]	2.1%	0.8%	10960
Acquirer's abnormal return [-1d, 10d]	1.8%	1.0%	10266
Acquirer's return [-1d, 365d]	13.5%	2.8%	10977
Industry return [-1d, 365d]	13.0%	12.7%	10977
Market return [-1d, 365d]	13.0%	15.5%	10977
Acquirer's abnormal return [-1d, 365d]	4.5%	5.4%	10271
Acquirer's return [-1d, 730d]	20.8%	0.6%	10984
Industry return [-1d, 730d]	26.2%	29.0%	10984
Market return [-1d, 730d]	25.1%	21.9%	10982
Acquirer's abnormal return [-1d, 730d]	6.2%	9.6%	10271
Acquirer's CAPM beta	0.95	0.85	10283
Acquirer's ROA (1-year)	-6.6%	2.2%	9615
Acquirer's change in ROA (1-year)	-5.8%	-1.1%	9501
Acquirer's OPINC/assets	6.6%	9.7%	9545
Acquirer's change OPINC/assets	-1.5%	-0.7%	9419
Acquirer's size (in $bn)	3.6	0.3	9829
Acquirer's B/M	0.5	0.4	9873
Target's return [-1d, 10d]	18.7%	13.8%	1605
Target's abnormal return [-1d, 10d]	18.2%	14.0%	1603
Target's CAPM beta	0.78	0.72	1650
Target's size (in $bn)	0.9	0.1	1414
Target's B/M	0.7	0.5	1388
deal value (in $m)	257.6	24.1	11366
Industry merger clusteredness (FF 48)	4.0%	3.8%	11366
Industry merger clusteredness (4-digit SIC)	5.9%	4.4%	6784
%shares held by acquirer's CEO	4.9%	0.4%	6862
Dummy of insider sales [t-365d, t-1d]	81.8%		11366
Dummy of relationship industry	23.3%		11366

Table 2 – Post-merger Performance and Clusteredness (Univariate Analysis)

All horizontal mergers in the sample are sorted into quintiles by the measure of clusteredness (*num3adj*). The table shows the mean value of the following variables for deals in each quintile: whether the method of payment is predominantly in stock (>50%), acquirer's post-announcement return over 365 days, acquirer's post-announcement return over 730 days, acquirer's abnormal return over 365 days (using the market model, [t-250d, t-11d] as the estimation window), and acquirer's abnormal return over 730 days.

Difference between the quintile 1 and quintile 5 and the Z-statistics of the difference in mean are calculated.

Quintile	Number of observations	Merger clustered-ness	% stock deals	Acquirer's return [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's return [-1d, +730d]	Acquirer's CAR [-1d, +730d]
1	2270	1.4%	15.2%	21.0%	11.3%	33.7%	15.4%
2	2280	2.7%	16.5%	20.2%	10.3%	39.8%	15.5%
3	2275	3.8%	18.4%	11.6%	2.4%	22.9%	6.1%
4	2276	5.0%	20.3%	8.7%	-0.4%	14.6%	-0.3%
5	2265	7.5%	25.7%	6.5%	-1.3%	-6.2%	-6.1%
1-5				14.5%	12.5%	39.9%	21.4%
Z-statistic				4.9	6.1	10.7	7.7

Table 3a – Post-merger Stock Performance and Clusteredness (Regression Analysis)

All regressions are run on the entire sample of horizontal mergers under the Fama-French 48 industry classification. All models control for industry (Fama-French 48 Industry) fixed effect and year fixed effect. All standard errors are robust to industry and year clustering. *Overall merger clusteredness* is the total number of mergers during the event window [t-3m, t+3m] in the SDC dataset universe adjusted by 10^{-4}. Acquirer's B/M and market capitalization are measured at the end of calendar year t-1. All standard errors are robust to industry and year clustering.

	(1)	(2)	(3)	(4)	(5)	(6)
	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +730d]
Dummy (stock deal)	-0.046		-0.046	-0.052	-0.048	-0.147
	(1.85)*		(1.86)*	(0.81)	(1.92)*	(1.91)*
acquirer's B/M	0.119	0.119	0.116	0.116	0.115	0.219
	(5.21)***	(5.36)***	(5.20)***	(5.21)***	(5.17)***	(6.06)***
log(acquirer's mktcp)	-0.042	-0.041	-0.041	-0.041	-0.041	-0.065
	(9.76)***	(9.76)***	(9.71)***	(9.63)***	(9.62)***	(8.80)***
Industry merger clusteredness		-3.560	-3.564	-3.597	-3.189	-5.084
(FF 48 industry level)		(2.96)***	(2.97)***	(3.69)***	(2.73)***	(4.44)***
D (stock deal) x Industry merger clusteredness				0.132		0.964
				(0.07)		(0.48)
Overall merger clusteredness					-1.120	-0.187
					(2.18)**	(0.43)
Observations	9185	9185	9185	9185	9185	8805
R-squared	0.06	0.07	0.07	0.07	0.07	0.09

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3b – Post-merger Stock Performance and Clusteredness (4-digit SIC level)

All regressions are run on the entire sample of horizontal mergers under the 4-digit SIC classification. All models control for industry (4-digit SIC code) fixed effect, year fixed effect, and industry-year clustering in errors. *Overall merger clusteredness* is the total number of mergers during the event window [t-3m, t+3m] in the SDC dataset universe adjusted by 10^{-4}. Acquirer's B/M and market capitalization are measured at the end of calendar year t-1. All standard errors are robust to industry and year clustering.

	(1)	(2)	(3)	(4)	(5)	(6)
	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +730d]
Dummy (stock deal)	-0.020		-0.023	0.010	-0.025	-0.117
	(0.82)		(0.95)	(0.25)	(1.03)	(1.98)**
acquirer's B/M	0.123	0.120	0.120	0.119	0.117	0.237
	(4.03)***	(3.97)***	(3.94)***	(3.94)***	(3.85)***	(5.27)***
log(acquirer's mktcp)	-0.043	-0.043	-0.043	-0.043	-0.043	-0.070
	(7.43)***	(7.50)***	(7.46)***	(7.45)***	(7.41)***	(7.86)***
Industry merger clusteredness		-1.829	-1.838	-1.704	-1.749	-2.720
(4-digit SIC level)		(3.58)***	(3.60)***	(3.47)***	(3.42)***	(3.80)***
D (stock deal) x Industry merger clusteredness				-0.642		0.724
				(0.84)		(0.76)
Overall merger clusteredness (num3all)					-1.053	0.334
					(1.77)*	(0.55)
Observations	5444	5444	5444	5444	5444	5210
R-squared	0.17	0.17	0.17	0.17	0.18	0.21

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4 – Post-merger Operating Performance and Clusteredness

Operating performance measures are defined as follows: OPINC/assets = DATA13/(DATA6+DATA196); ROE=DATA18/DATA60. Change in operating performance is measured as the difference between first full fiscal year after merger completion and the last fiscal year prior to merger announcement. *Overall merger clusteredness* is the total number of mergers during the event window [t-3m, t+3m] in the SDC dataset universe adjusted by 10^{-4}. All models control for industry and year fixed effect. All standard errors are robust to industry and year clustering.

	(1)	(2)	(3)	(4)	(5)	(6)
	Acquirer's change in OPINC/assets	Acquirer & target's change in OPINC/assets	Acquirer's post-merger OPINC/assets	Acquirer's post-merger OPINC/assets	Acquirer's change in ROE	Acquirer & target's change in ROE
Industry merger clusteredness	-0.395	-0.567	-0.844	-0.914	-5.186	-10.513
(FF 48 industry level)	(2.33)**	(1.91)*	(4.79)***	(2.98)***	(4.01)***	(3.96)***
Overall merger clusteredness	0.120	0.070	0.127	0.095	0.089	-0.268
	(1.62)	(0.51)	(1.77)*	(0.64)	(0.14)	(0.22)
Dummy (stock deal)	0.000	-0.001	-0.013	-0.016	-0.075	-0.088
	(0.06)	(0.13)	(3.05)***	(2.39)**	(1.70)*	(1.37)
acquirer's B/M	0.019	0.034	0.010	0.021	-0.035	0.131
	(4.04)***	(1.44)	(2.57)**	(1.75)*	(0.88)	(1.65)*
log(acquirer's mktcp)	-0.002	-0.003	0.008	0.008	-0.015	0.003
	(1.13)	(1.27)	(7.27)***	(4.36)***	(1.71)*	(0.20)
Acquirer's lagged OPINC/assets			0.55	0.56		
			(19.92)***	(12.40)***		
Target's lagged OPINC/assets				0.02		
				(1.49)		
Observations	8346	1195	8346	1195	8417	1218
R-squared	0.04	0.09	0.43	0.50	0.04	0.11

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5 – Calendar-time Portfolio

At the end of each quarter, acquirers of all horizontal mergers announced during the quarter are sorted into tertiles by the degree of clusteredness (*num3adj*). A long-short portfolio that buys the least clustered tertile (#1) and short sells the most clustered tertile (#3) is held for 3 months starting from 3 months after the quarter end. If a stock disappears from the CRSP within 6 months of the announcement, value-weighted market returns are used to replace the missing monthly returns. Returns of these calendar-time portfolios from January 1985 to December 2004 are examined.[1]

Equal-weighted

Calendar-time portfolio	monthly return	monthly alpha	t(alpha)
Stock deals:			
Acquirers of most clustered deals (portfolio #3)	-0.1%	-0.9%	-1.14
Acquirers of least clustered deals (portfolio #1)	1.1%	0.2%	0.29
Long least clustered and short most clustered (#1-#3)	1.1%	0.9%	0.94
All cash deals:			
Acquirers of most clustered deals (portfolio #3)	0.0%	-0.4%	-1.03
Acquirers of least clustered deals (portfolio #1)	1.4%	0.6%	1.98
Long least clustered and short most clustered (#1-#3)	1.4%	0.9%	2.00
All deals:			
Acquirers of most clustered deals (portfolio #3)	0.0%	-0.4%	-1.24
Acquirers of least clustered deals (portfolio #1)	1.4%	0.7%	2.25
Long least clustered and short most clustered (#1-#3)	1.4%	1.1%	2.36

Value-weighted

Calendar-time portfolio	monthly return	monthly alpha	t(alpha)
Stock deals:			
Acquirers of most clustered deals (portfolio #3)	-0.5%	-1.2%	-1.40
Acquirers of least clustered deals (portfolio #1)	0.9%	0.5%	0.67
Long least clustered and short most clustered (#1-#3)	1.3%	1.5%	1.37
All cash deals:			
Acquirers of most clustered deals (portfolio #3)	-0.1%	-0.3%	-0.64
Acquirers of least clustered deals (portfolio #1)	1.0%	0.3%	0.65
Long least clustered and short most clustered (#1-#3)	1.1%	0.5%	0.89
All deals:			
Acquirers of most clustered deals (portfolio #3)	0.0%	-0.4%	-1.06
Acquirers of least clustered deals (portfolio #1)	1.0%	0.7%	1.63
Long least clustered and short most clustered (#1-#3)	1.1%	1.1%	1.90

[1] SDC merger data starts in 1979. January 1985 is chosen as the beginning of calendar-time portfolio due to lack of data points prior to 1984.

Table 6 – Acquirer and Target's Combined Post-merger Performance and Clusteredness

Acquirer and target's combined return is defined as the return to a value-weighted portfolio of acquirer and target stocks over [t-1d, t+10d] compounded to the return of the acquirer stock over [t+10d, t+365d].[2] All models control for industry and year fixed effect. All standard errors are robust to industry and year clustering.

Panel A: Univariate Analysis

Quintile	# of obs	Merger clustered-ness	Acquirer & target's return [-1d, +365d]	Acquirer & target's CAPM-alpha [-1d, +365d]	Acquirer & target's change in OPINC/assets
1	369	1.2%	22.9%	8.8%	-0.6%
2	370	2.5%	27.0%	13.2%	0.0%
3	371	3.6%	13.5%	2.7%	0.8%
4	369	4.7%	11.4%	-2.7%	-2.3%
5	371	7.1%	1.9%	-3.8%	-2.5%
1-5			21.0%	12.6%	1.9%
Z-statistic			4.5	2.9	2.2

Panel B: Regression Analysis

	(1) Acquirer & target's CAPM-alpha [-1d, +365d]	(2) Acquirer & target's CAPM-alpha [-1d, +365d]	(3) Acquirer & target's CAPM-alpha [-1d, +365d]	(4) Acquirer & target's CAPM-alpha [-1d, +730d]	(5) Acquirer & target's CAPM-alpha [-1d, +365d]	(6) Acquirer & target's CAPM-alpha [-1d, +730d]
Industry merger clusteredness	-3.650	-3.622	-2.987	-4.788		
(FF 48 industry level)	(2.26)**	(2.32)**	(1.93)*	(2.70)***		
acquirer's B/M ratio	0.076	0.058	0.054	0.206	0.104	0.286
	(1.20)	(0.95)	(0.90)	(1.80)*	(0.97)	(1.94)*
log(acquirer's marketcap)	-0.020	-0.025	-0.026	-0.033	-0.013	-0.048
	(1.64)	(2.13)**	(2.23)**	(1.51)	(0.79)	(0.86)
Dummy (stock deal)		-0.136	-0.136	-0.303	-0.079	-0.164
		(3.90)***	(3.91)***	(4.51)***	(1.53)	(1.46)
Overall merger clusteredness			-1.680	-1.645		
			(1.84)*	(1.47)		
Industry merger clusteredness					-2.652	-6.295
(4-digit SIC level)					(2.68)***	(1.95)*
Observations	1338	1338	1338	1302	773	773
R-squared	0.11	0.12	0.13	0.09	0.32	0.23

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

[2] The market-model beta for the merged company is the value-weighted beta of each individual firm estimated over [t-250d, t-11d]. A minimum of 100 days of valid observations of stock returns is required.

Table 7 – Clusteredness, Post-merger Performance, and Managerial Ownership

The acquirer CEO's percentage ownership is obtained from Thomson Financial Insider Trading data.

All models control for acquirer's size, acquirer's B/M, method of payment dummy, and industry and year dummies. All standard errors are robust to industry and year clustering.

Panel A: Univariate Analysis (above median[3] percentage ownership)

Quintile	# of obs	Merger clustered-ness	Acquirer's return [-1d, +365d]	Acquirer's CAR [-1d, +365d]	Acquirer's return [-1d, +730d]	Acquirer's CAR [-1d, +730d]	Acquirer's change in OPINC/asset
1	631	1.8%	28.0%	17.2%	48.0%	23.6%	-1.2%
2	633	3.1%	23.6%	9.3%	54.2%	19.0%	-0.1%
3	630	4.1%	9.1%	4.1%	18.4%	6.5%	-0.7%
4	632	5.4%	12.9%	2.4%	13.2%	-2.7%	-1.4%
5	634	7.9%	0.7%	-5.9%	-12.1%	-7.2%	-6.8%
1-5			27.3%	23.0%	60.1%	30.9%	5.6%
Z-statistic			4.4	5.8	9.2	5.8	4.7

Panel B: Regression Analysis

	(1) Acquirer's CAR [-1d, +365d]	(2) Acquirer's CAR [-1d, +730d]	(3) Acquirer's change in OPINC/assets
Industry merger clusteredness	-3.570	-4.954	-0.527
(FF 48 industry level)	(2.57)**	(3.55)***	(3.05)***
%Shares Held by CEO	0.319	0.556	-0.030
	(2.15)**	(2.41)**	(0.96)
%Shares Held by CEO	-2.824	-6.764	0.359
x Industry merger clusteredness	(0.91)	(1.66)*	(0.49)
Observations	6459	6459	5935
R-squared	0.07	0.10	0.05

Robust t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

[3] In the sample, the median CEO ownership is 0.5%.

Table 8 –Post-merger Performance and Clusteredness (cash deals only)

Industry merger clusteredness (cash deals only) is defined as the number of all-cash deals[4] over the event window [t-3m, t+3m] over the total number of all-cash deals over the entire sample period.

The sample in Panel A includes all horizontal mergers that satisfy the following conditions: (a) the entire consideration is paid in the form of cash; (b) the target company is private; (c) the acquirer did not issue equities or bonds over [t-365d, t+365d].

The sample in Panel B includes all-cash deals in the main sample. All models control for acquirer's size, acquirer's B/M, acquirer's cash holdings, and industry and year dummies. All standard errors are robust to industry and year clustering.

Panel A: Univariate Analysis

Quintile	# of obs	cash-deal clusteredness	Acquirer's CAR [-1d, +365d]	Acquirer's CAR [-1d, +730d]	Acquirer's change in OPINC/Assets
1	899	1.6%	14.5%	21.4%	-1.1%
2	901	2.7%	13.5%	23.5%	-0.7%
3	893	3.8%	2.2%	9.4%	0.0%
4	902	4.9%	1.0%	7.0%	-2.2%
5	899	7.2%	6.3%	3.6%	-3.0%
1-5			8.2%	17.8%	1.9%
Z-statistic			2.9	4.6	3.3

Panel B: Regression Analysis

	(1) Acquirer's CAR [-1d, +365d]	(2) Acquirer's CAR [-1d, +365d]	(3) Acquirer's change in OPINC/assets
Industry merger clusteredness	-2.076		
(FF48 industry level)	(1.81)*		
Industry merger clusteredness		-1.518	-0.335
(cash deals only)		(1.78)*	(2.06)**
Observations	2424	2424	2171
R-squared	0.11	0.11	0.06

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

[4] All-cash deals satisfy the following two conditions: (a) the entire consideration is paid in the form of cash; (b) the acquirer did not issue any public stocks or bonds from one year prior to the merger announcement to one year afterwards.

Panel C: Performance and Clusteredness Measures (Cash vs. Stock)

In models (1) through (4), clusteredness of all-cash deals (non-all-cash deals) is defined as the number of all-cash (non all-cash) deals[5] over the event window [t-3m, t+3m] over the total number of all-cash (non all-cash) deals over the entire sample period using Fama-French 48 industry classification. In models (5) through (8), clusteredness of all-cash deals (non-all-cash deals) is defined as the number of all-cash (non all-cash) deals over the event window [t-3m, t+3m] over the total number of all-cash (non all-cash) deals over the entire sample period using 4-digit SIC industry classification.

	(1) Acquirer's CAR [-1d, +365d]	(2) Acquirer's CAR [-1d, +365d]	(3) Acquirer's CAR [-1d, +365d]	(4) Acquirer's CAR [-1d, +365d]	(5) Acquirer's CAR [-1d, +365d]	(6) Acquirer's CAR [-1d, +365d]	(7) Acquirer's CAR [-1d, +365d]	(8) Acquirer's CAR [-1d, +365d]
Merger clusteredness	-3.913							
(FF 48 industry level)	(3.36)***							
Merger clusteredness		-2.816		-1.905				
(FF 48 & all-cash deals only)		(2.78)***		(2.24)**				
Merger clusteredness			-3.010	-2.261				
(FF 48 & non-all-cash deals only)			(3.23)***	(2.67)***				
Merger clusteredness					-1.927			
(4-digit SIC level)					(4.00)***			
Merger clusteredness						-0.323		-0.468
(4-digit SIC & all-cash deals only)						(2.30)**		(3.10)***
Merger clusteredness							-1.269	-1.58
(4-digit SIC & non-all-cash deals only)							(3.77)***	(4.15)***
Observations	10271	9596	10183	9508	6109	6019	6106	6016
R-squared	0.04	0.04	0.04	0.04	0.14	0.12	0.14	0.12

Robust t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

[5] All-cash deals satisfy the following two conditions: (a) the entire consideration is paid in the form of cash; (b) the acquirer did not issue any public stocks or bonds from one year prior to the merger announcement to one year afterwards.

Table 9 – Post-merger Performance, Clusteredness, and Insider Trading

Insider trading data are collected by Thomson Financial from the required SEC insider trading filings. For each merger, I sum all (split-adjusted) open market transactions for all insiders over [t-365d, t-1d], with sales entering positively and purchases entering negatively. I create the variable insider netsales dummy (*Dnetsales*) that equals one if the net purchase of all insiders during the year prior to the merger announcement is positive and zero otherwise.

All models control for acquirer's size, acquirer's B/M, and industry and year dummies. All standard errors are robust to industry and year clustering.

	(1) Acquirer's CAR [-1d, +365d]	(2) Acquirer's CAR [-1d, +730d]	(3) Acquirer's CAR [-1d, +365d]	(4) Acquirer's CAR [-1d, +730d]
Industry merger clusteredness	-4.581	-8.37		
(FF 48 industry level)	(3.16)***	(4.60)***		
Dummy (insider sales)	-0.044	-0.138	-0.055	-0.05
	(1.22)	(2.16)**	(1.24)	(0.85)
Dummy (insider sales)	1.264	4.325		
x Industry merger clusteredness	(1.40)	(2.50)**		
Dummy (stock deals)	-0.047	-0.108	-0.023	-0.081
	(1.89)*	(3.48)***	(0.95)	(2.09)**
Industry merger clusteredness			-2.585	-3.24
(4-digit SIC level)			(3.22)***	(3.33)***
Dummy (insider sales)			0.954	0.896
x Industry merger clusteredness (4-digit SIC level)			(1.38)	(0.94)
Observations	9185	8805	5444	5210
R-squared	0.07	0.10	0.17	0.21

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10 – Post-merger Performance and the Clusteredness of Same-industry and Diversifying Mergers (Fama-French 48 Industry)

The clusteredness measure for deals where the acquirer (target) is in the given industry is defined as the number of diversifying mergers (using Fama-French 48 industry classification) during the event window [t-3m, t+3m] that involve firms in the given industry as acquirer (target) divided by the total number of diversifying mergers over the sample period (1979-2004) that involve firms in the given industry as acquirer (target).[6]

All models control for industry and year fixed effect. All standard errors are robust to industry and year clustering.

	(1) Acquirer's CAR [-1d, +365d]	(2) Acquirer's CAR [-1d, +365d]	(3) Acquirer's CAR [-1d, +730d]	(4) Acquirer's CAR [-1d, +730d]
Horizontal merger clusteredness	-3.564	-2.021	-4.905	-2.340
(FF 48 industry level)	(2.97)***	(1.94)*	(4.33)***	(1.81)*
acquirer's B/M	0.116	0.114	0.219	0.217
	(5.20)***	(5.18)***	(6.07)***	(6.09)***
log(acquirer's mktcp)	-0.041	-0.041	-0.064	-0.064
	(9.71)***	(9.63)***	(8.76)***	(8.62)***
Dummy (stock deal)	-0.046	-0.046	-0.105	-0.104
	(1.86)*	(1.83)*	(3.36)***	(3.34)***
Diversifying merger clusteredness		-0.908		-2.125
(where acquirer is in this industry)		(0.89)		(1.44)
Diversifying merger clusteredness		-3.069		-4.522
(where target is in this industry)		(2.85)***		(2.86)***
Observations	9185	9185	8805	8805
R-squared	0.07	0.07	0.09	0.10

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

[6] All mergers that involve financial or utility firms are excluded.

Table 11 – Market, Industry, Characteristics-Adjusted Stock Performance

Market return is the CRSP value-weighted market index. Industry return is the value-weighted return on the Fama-French 48 industry portfolio. Acquirer's abnormal B&H return is defined as acquirer return (compounded) minus matched-firm's return (compounded) over the same period. If matched firm's return is missing on certain days, value-weighted market index return is used.

All models control for industry and year fixed effect. All standard errors are robust to industry and year clustering.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Market return [-1d, +365d]	Industry return [-1d, +365d]	Industry return [-1d, +730d]	Acquirer's industry-adjusted return [-1d, +365d]	Acquirer's industry-adjusted return [-1d, +730d]	Acquirer's abnormal B&H return [-1d, +365d], size	Acquirer's abnormal B&H return [-1d, +365d], size & B/M
Horizontal merger clusteredness	-0.199	-2.739	-5.708	-1.676	-1.649	-2.978	-0.362
(FF 48 industry level)	(0.98)	(4.00)***	(6.01)***	(2.05)**	(1.83)*	(3.19)***	(0.17)
Dummy (stock deal)	-0.002	-0.008	-0.012	-0.068	-0.142	-0.099	-0.065
	(0.52)	(1.16)	(1.20)	(3.31)***	(3.85)***	(3.82)***	(2.09)**
Observations	10977	10758	10224	10758	10224	10977	10977
R-squared	0.66	0.38	0.48	0.04	0.03	0.02	0.01

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 12 – Cross-Industry Comparison: Concentration

The Herfindahl-index is the industry concentration measure: $HHI = \sum_{i=1}^{N} (\Pi_i)^2$, where Πi is the market share of company i and n is the number of firms in the industry. The HHI index is obtained from Census Bureau's quinquennial plant-level surveys on manufacturing firms in the US. Both models control for industry (Fama-French 48) and year fixed effect. Both standard errors are robust to industry and year clustering.

	(1) Acquirer's CAR [-1d, +365d]	(2) Acquirer's CAR [-1d, +730d]
Industry concentration (HHI x 10^{-3})	-0.099	-0.074
	(1.58)	(0.86)
HHI x Industry merger clusteredness	3.767	4.644
	(2.19)**	(2.15)**
Industry merger clusteredness	-3.809	-4.945
(FF 48 industry level)	(2.71)***	(2.63)***
Dummy (stock deal)	0.010	-0.069
	(0.36)	(1.75)*
acquirer's B/M	0.093	0.213
	(2.58)***	(3.76)***
log(acquirer's mktcp)	-0.042	-0.068
	(5.63)***	(6.84)***
Observations	3048	2901
R-squared	0.10	0.13

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 13 – Cross-Industry Comparison: Relationship

The following two-digit-SIC industries are classified as relationship industries: 15-17, 34-39, 42, 47, 50-51, 55, 60-65, 67, 75-76, 87. The variable *relationship dummy* takes a value of one if the company operates in one of those two-digit-SIC industries and zero otherwise. Both models control for industry and year fixed effect. Both standard errors are robust to industry and year clustering.

	(1) Acquirer's CAR [-1d, +365d]	(2) Acquirer's CAR [-1d, +730d]
Industry merger clusteredness	-4.488	-6.007
	(3.39)***	(4.98)***
Dummy (relationship industry)	-0.170	-0.226
	(2.53)**	(2.65)***
Dummy (relationship industry)	3.746	4.497
x Industry merger clusteredness	(2.33)**	(2.51)**
acquirer's B/M ratio	0.116	0.219
	(5.23)***	(6.08)***
log(acquirer's marketcap)	-0.042	-0.066
	(9.97)***	(8.89)***
Dummy (stock deal)	-0.047	-0.106
	(1.91)*	(3.42)***
Observations	9185	9470
R-squared	0.07	0.10

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%