PEER EFFECTS AND DIVIDEND POLICY

A Dissertation Presented

by

FANG YANG

Submitted to the College of Graduate Studies of the
Middle Tennessee State University in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

MAY, 2009

Department of Economics and Finance

APPROVAL PAGE

PEER EFFECTS AND DIVIDEND POLICY

BY

FANG YANG

**A DOCTORAL DISSERTATION SUBMITTED
IN PARTIAL FULFILLMENT OF
THE REQUIREMENT FOR THE DEGREE OF**

DOCTOR OF PHILOSOPHY / ECONOMICS

MIDDLE TENNESEE STATE UNIVERSITY

MAY, 2009

APPROVED BY, Joachim Zietz

Dr. Joachim Zietz, Committee Chair

Ellis Anthon Eff

Dr. Ellis Anthon Eff, Committee Member

Adam D. Rennhoff

Dr. Adam D. Rennhoff, Committee Member

Gregory E. Givens

Dr. Gregory E. Givens, Director, Economics Graduate Program

Charles L. Baum

Dr. Charles L. Baum, Department Chair, Economics & Finance

Michael D. Allen

Dr. Michael D. Allen, Dean, College of Graduate Studies

ACKNOWLEDGMENTS

I would like to express my grateful acknowledgements to the many individuals who have contributed to and improved the quality of my dissertation. My first thanks go to my dissertation chair, Dr. Joachim Zietz, who spent significant time reading and revising hundreds of pages of drafts. His directions of the project, valuable suggestions, and constant encouragement along the way have motivated me to accomplish the work. He, more than anyone else, has taught me how to conduct research. I also sincerely thank the other members of my dissertation committee. Dr. Anthon Eff provided invaluable help. He was always optimistic and willing to spend a large amount of time correcting my software codes. His suggestions tremendously improved the quality of the dissertation. Dr. Adam D. Rennhoff's expertise on the panel estimation techniques helped me construct the empirical methodology. Dr. Owens's suggestions on peer effects improved the quality of the work. I also want to thank Dr. Duane B. Graddy whose instruction, mentoring, and advice have meant more to me than he will ever know.

In addition to those mentioned above, I wish to thank the following individuals for their insightful suggestions and interest in my work, Dr. Stuart J. Fowler, Dr. Gregory E. Givens, Dr. Charles F. Beauchamp, John M. Nunley, Pamela D. Morris and Alan Seals.

Lastly, I thank my wonderful husband, Rong, for being my rock when rains came and beat on me. His constant love, understanding and strong support became my reason to keep working. All errors in this dissertation are mine.

ABSTRACT

PEER EFFECTS AND DIVIDEND POLICY

By

Fang Yang

This paper adds a variable capturing peer effects to the dividend regression models explaining a firm's dividend behavior. The existing literature on dividend policy primarily focuses on agency theory, tax effects, or investors' preferences to explain the observed trends in firms' dividend behavior. Peer effects are formulated by a spatial lag variable, which is constructed on the basis of equal 2-digit, 3-digit, or 4-digit Standard Industrial Classification (SIC) codes.

The estimation of peer effects has been a very difficult task because of the reflection problem as well as data availability problems (Manski, 1993). Manski indicates that it is impossible to separately identify peer effects from the other types of neighborhood effects in the linear model. This study directly confronts the reflection problem by using a lagged peer variable in the dividend model.

The hypothesis is that a firm is more likely to change its dividend policy when its peers are doing the same regardless of its own financial conditions. To test this hypothesis, peer effects are treated in analogy to spatial correlation in regional science and real estate economics. The empirical methodology uses spatial econometrics

techniques as typically employed in these fields (e.g., Anselin, 1988). Specifically, to identify peer effects, a spatial lag variable is constructed and added to the dividend regression models with traditional control variables. The models are estimated on both cross-sectional and panel data. The cross-sectional regression models are estimated on the companies of the S&P 1500 Super Composite Index and the S&P 500 index using data from the years 2003 to 2006. The panel regressions employ S&P 1500 data for the seven years from 2000 to 2006.

The cross-sectional results from the S&P 1500 sample show strong evidence of peer effects. The spatial lag variable constructed for the 2-digit Standard Industrial Classification (SIC) data is highly significant and has the expected positive sign for all four years. The results are similar for the models that employ spatial lag variables for the 3-digit and the 4-digit industries. There is also strong evidence in support of the peer effects hypothesis for the S&P 500 sample. On average, the peer effects measures have a stronger impact on the amount of dividends paid than do size and profitability, which are the traditional explanatory variables emphasized in the dividend literature.

For the S&P 1500 panel data, the coefficient of the peer variable also tends to be positive and statistically significant. However, the peer effects results are not as consistent across alternative models as those for the cross-section regressions.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1: Introduction

The issue of firm dividend policy has drawn much attention over the decades (Black and Scholes, 1974; Brittain, 1964; Charitou and Vafeas, 1998; DeAngelo et al., 1992; DeAngelo et al., 2004; DeAngelo et al., 2006; Denis and Osobov, 2008; Dhillon and Johnson, 1994; Eije and Megginson, 2008; Fama and French, 2001; Fama and French, 2002; Li and Zhao, 2008; Lintner, 1956; Mancinelli and Ozkan, 2006; Michael et al., 1995; Miller and Modigliani, 1961; Pettit, 1972; Skinner, 2008; Zhou and Ruland, 2006) and has remained controversial.

Dividend policy issue is important for several reasons. First, it involves substantial amounts of money and is, therefore, a core component of a firm's financial policy and its investment decisions. Second, dividend policy continues to be a puzzle for both academic researchers and firm managers. Even though many studies have been done after Black (1976), his view still represents current researchers' opinion (p. 5): "why do corporations pay dividends? Why do investors pay attention to dividends? … I claim that the answers to these questions are not obvious at all. The harder we look at the dividend picture, the more it seems like a puzzle, with pieces that just do not fit together."

Finally, the dividend payout pattern affects a firm's stock price and, therefore, the market value of a firm. Generally, the stock price will go up when a firm initiates or increases dividend payments. Aharony and Swary (1980) find that the share price changes significantly along with the announcements of dividend increases and decreases. Asquith and Mullins (1983) document that the share price reacts positively to dividend initiations while Healy and Palepu (1988) and Michael et al. (1995) report that dividend omissions are associated with a significantly negative response of the share price. For that very reason, a firm usually does not like to terminate or reduce dividend payments (Woolridge and Ghosh, 1988 and 1991).

Some determinants of firms' dividend policy are well known, [1] such as investment opportunities, earnings, size and past dividend policy. Specifically, firms with more investment opportunities usually have a lower dividend payout ratio than stable firms. A higher dividend payout ratio is common for firms with more stable earnings. Dividend policy during times of strong economic growth (late 90s) tends to be different from dividend policy during recessionary times. A smaller firm is less likely to have a high payout ratio. A firm's past dividend policy affects its future one and a firm is, therefore, much more likely to have a smooth and persistent dividend payout policy (Linter, 1956).

There are a number of conflicting theoretical and empirical models of firm dividend policy attempting to explain a firm's dividend behavior. According to the classical work of Miller and Modigliani (M&M) (1961), the dividend policy of a firm does not affect its value when capital markets are complete and perfect. Dividend policy has, therefore, no impact on shareholders' wealth.

Later research has relaxed the strong assumptions of the M&M study[2] and has offered several theories of dividend policy. Numerous studies focus on the impact of tax policy on dividend behavior (i.e., Litzenberger and Ramaswamy, 1979; Perez-Gonzalez,

---

[1] They can be found at http.//pages.stern.nyu.edu/~adamodar/New_Home_Page/lectures/dividend.html

[2] Assumption from M&M study is perfect capital markets.

2003; Poterba and Summers, 1984 and 1985). Perez-Gonzalez (2003), for example, presents evidence that dividend payouts are directly related to the tax treatment of dividends relative to capital gains. In particular, the number of dividend paying firms should decrease when the tax on dividends rises relative to the tax on capital gains. Agency theory (e.g., Easterbrook, 1984; Fluck, 1995; Gomes, 1996; Jensen, 1986; Myers, 1996; Rozeff, 1982) suggests that dividends are paid by firms in order to reduce agency costs. Because of the separation of principal and agent for a firm, Jensen (1986) argues that shareholders are better off if they receive part of the earnings as dividend payout; this way, these funds will not be invested in unprofitable projects by a firm. According to the dividend signaling models (e.g., Aharony and Swany, 1980; Asquith and Mullins, 1983; Bhattacharya, 1979; John and Williams, 1985; Miller and Rock, 1985; Ross, 1977), a firm uses dividends as a way to signal its current and future stability and earnings potential to outsiders. Shefrin and Statman (1984) argue that investors' preference for dividends follows from the behavioral theories of individual choice.

Even though researchers have devoted much effort trying to resolve the dividend puzzle, there continues to be little agreement on the question of why corporations pay dividends. As Ang (1987, p.55) states, "we have moved from a position of not enough good reasons to explain why dividends are paid to one of too many. Unfortunately, some of these may not be very good reasons, i.e., not consistent with rational behavior."

In this paper, I add to this line of work by considering the possibility that a firm's dividend decision is influenced by peer-group effects. The purpose of this study is to identify whether peer effects play any role in dividend behavior.[3]

There are several reasons to suspect that the behavior of a firm that pays dividends may be influenced by the dividend decisions of other firms. The main reason for peer group effects is likely to be the desire to avoid unfriendly takeovers when the stock price reacts negatively to an unfavorable dividend signal. Pursuing a dividend policy that is too different from that of peer firms would be such an unfavorable signal. Peer group effects have recently received attention in other parts of finance and it appears that firm behavior will be influenced by similar motivations (Baker and Powell, 2000; Baker et al., 2001).

The hypothesis is that a firm is more likely to change its dividend policy when its peers are doing the same regardless of its own financial conditions. To test this hypothesis, peer effects are treated in analogy to spatial correlation in regional science and real estate economics. The empirical methodology uses spatial econometrics techniques as typically employed in these fields (e.g., Anselin 1988). Specifically, to identify peer effects, a spatial lag variable is constructed and added to the dividend regression models with traditional control variables.[4] The models are estimated on both cross-sectional and panel data. The cross-sectional regression models are estimated on the companies of the S&P 1500 Super Composite Index and the S&P 500 index using the

[3] There is no intent to develop a model that can predict the reaction of companies to changes in the tax laws that pertain to dividends. For example, the study will not address what will happen to dividends versus buy-backs if the current dividends laws expire or are changed in the future.

[4] If $\mathbf{y}$ is the vector of the dependent variable, then a spatial lag variable is computed as $\mathbf{W}^s\,\mathbf{y}$, where $\mathbf{W}^s$ is the spatial weight matrix that is created on the basis of SIC codes.

data from years 2003 to 2006. The panel regressions employ S&P 1500 data for seven years from 2000 to 2006.[5]

The scope of this study is limited to the analysis of peer effects among dividend paying firms that also paid dividends in the past year. Firms that initiate or terminate dividends in a given year are not considered because both events are often assumed to be special events in the history of a firm for which peer effects may not be the dominant motivation.

This study contributes to the empirical literature in three respects: (1) the observation period is more recent (2003-2006) compared with most previous studies; hence, the results better reflect the current environment of firms; (2) a variable capturing peer effects is added to the traditional dividend model; and (3) the study includes a broad set of firms from different industries.

Cross-sectional results from the S&P 1500 sample show strong evidence of peer effects. The spatial lag variable constructed for the 2-digit Standard Industrial Classification (SIC) data is highly significant and has the expected positive sign for all four years. The results are similar for the models that employ spatial lag variables for the 3-digit and the 4-digit industries. There is also strong evidence in support of the peer effects hypothesis for the S&P 500 sample. On average, the peer effects measures have a stronger impact on the amount of dividends paid than do size and profitability, which are the traditional explanatory variables emphasized in the dividend literature.

---

[5] The study cannot go beyond 2006 because of lack of current data. In order to estimate panel models, the inverse Mills ratio (*IMR*) needs to be calculated for each year. If 2007 data are used for the independent variables, the data for the dependent variable must be from 2008, which is not available yet. This study uses lagged independent variables to explain the dividend behavior in the current year.

For the S&P 1500 panel data, the coefficient of the peer variable also tends to be positive and statistically significant. However, the peer effects results are not as consistent across alternative models as those for the cross-section regressions.

This study is organized as follows. Chapter 2 contains the literature review. Chapter 3 specifies how peer effects for a firm's dividend decision are modeled. Chapter 4 describes the data and empirical methodology. Chapter 5 presents the results of the empirical estimates. Chapter 6 concludes.

Chapter 2: Literature Review

2.1 Literature Review on Dividend Policy

Since Black (1976) proposed the "dividend puzzle" in his classic work, a large body of research has evolved to explain why firms pay dividends. Most studies concentrate on the traditional dividend theories[6] and analyze commonly used variables in the dividend regression equations.

Most of the research on dividend policy has focused on explaining whether there is a tax effect on dividend payout. In the United States, dividends are taxed at an individual's personal income tax rate and have historically been taxed more heavily than capital gains. Therefore, individual investors would prefer capital gains to dividends and firms would decrease dividend payouts when the tax on dividends is much higher than the tax on capital gains. Poterba (2004) suggests that there is indeed a negative relationship between dividend payments and taxes on dividends. His finding is based on time series data for the U.S. from 1929 to 2002. The regression results confirm those of Fama and French (2001).

Agency theory (e.g., Easterbrook, 1984; Fluck, 1995; Gomes, 1996; Jensen, 1986; Myers, 1996; Rozeff, 1982) is another explanation of why a firm pays dividends. According to this theory, dividends are paid by firms in order to reduce agency costs. The

[6] For example, Ang (1987), Frankfurter (1999), and Lease et al. (2000) for a review of various dividend theories.

payment of dividends makes it less likely that retained earnings will be invested in unprofitable projects by a firm manager.

The assumption of asymmetric information on the future profitability of firms is the starting point of researchers who suggest that signaling theory may explain dividend payments. The best known signaling models are those of Bhattacharya (1979), John and Williams (1985), and Miller and Rock (1985). A firm is expected to increase its dividend to send out a positive signal. The models predict a positive association between the degree of asymmetric information and dividends.

Much of the traditional empirical work on dividend behavior starts with the Lintner (1956) model, which is developed based on a survey of 28 well established U.S. industrial firms. Lintner's model assumes that a firm partially adjusts to the target dividend level; thus the change in dividends from one year to the next is a function of the level of dividends in the previous year and the level of profits in the current year. He concludes that a firm's dividend payment is affected by the dividends of the previous year and current year earnings.

Over time, other models of dividend policy[7] have been developed by researchers. Empirical work, such as Fama and Babiak (1968), confirms that lagged dividends and current profits are the most important determinants of dividend changes. Their result is consistent with Lintner's finding. Hence, profits and the past pattern of dividends have long been regarded as the major determinants when making dividend policy decisions.

[7] Other types of models of dividend policy have been developed by Alli et al. (1993), Lauenstein (1987) and Rozeff (1982).

Besides profits and past year dividends, other factors, such as risk (year to year variability of earnings), have been identified as important determinants in determining a firms' dividend policy (Pruitt and Gitman, 1991). A firm with relatively stable earnings is usually able to predict future earnings and thus is more likely to pay higher dividends than a firm with fluctuating earnings. The negative relationship between risk and dividend payment is confirmed in other studies, including Lloyd et al. (1985), Rozeff (1982).

Fama and French (2002)'s study on dividends shows that the target dividend payout varies across firms as a function of profitability, investment opportunities, target leverage, and other driving forces. In order to mitigate any possible endogeneity problems, the authors use lagged values of these explanatory variables. The dependent variable is defined as dividends divided by assets. They run regressions for each year from 1965 to 1999 and use averages across years to draw their inferences. They conclude that the payout ratio is negatively related to investment opportunities and positively related to profitability and firm size. Specifically, firms with more investments have lower target dividend payouts, more profitable firms have higher target payouts, and smaller firms have lower dividend payouts.

DeAngelo et al. (2004) examine the dividend trends over the period1978–2000.[8] They find that aggregate real dividends have increased from 1978 to 2000, even though, as Fama and Fench (2001) report, the number of firms paying has declined by more than

[8] Fama and French (2001) examine trends for the years 1978–1998. DeAngelo et al. (2004) use two more years of data that became available only after Fama and French's study.

50%. DeAngelo et al. (2004) conclude that the top payers[9] have increased dividends while many small dividends payers have become nonpayers. The increase in dividend payments by the top payers more than offset the decrease in the number of dividend payers. Their study presents clear evidence that firms' dividends have become highly concentrated over the past twenty years.

DeAngelo et al. (2006) use data 1973-2002 to assess whether the probability a firm paying dividends is positively related to the ratio of retained earnings to total equity. They consistently observe that firms with a higher ratio are more likely to pay dividends, controlling for firm size, profitability, investment opportunities.[10] The study also finds the probability a firm is paying dividends is tied to its size, profitability, and growth. These results are consistent with the findings of Fama and French (2001).

Skinner (2008) examines the relationship between total payouts (dividends and repurchases) and earnings. The author finds that the significant relation between earnings and dividends (Fama and Babiak, 1968) has weakened over time.

Shiller (1984) argues that the best model explaining dividend policy will include variables that measure behavioral and socioeconomic influences on managers. Shiller (1990) restates that the behavior of other managers and social norms influence managers in determining dividend policy.

Over time, the factors that are considered as being important in making dividend decisions have increased substantially in the literature. Several surveys try to identify the

---

[9] Most top payers are in the Dow Jones Industrial Average (DJIA), such as Exxon Mobil and General Electric. Their data indicates that the top 25 payers are responsible for 54.9% of aggregate industrial dividends in 2000.

[10] DeAngelo et al. (2006) use the market-to-book ratio, the sales growth rate, and the asset growth rate as measures of investment opportunity

factors that drive dividend policies. Surveys conducted by Baker et al. (1985) and Farrelly et al. (1986) include 562 New York Stock Exchange (NYSE) firms which paid dividends in year 1983. Based on the responses of the managers of 318 firms from the utility, manufacturing, and wholesale/retail industries the authors conclude that expected future earnings and the pattern of past dividends are the crucial factors in determining firms' dividend polices. The survey results also suggest that dividend smoothing was the first concern for managers when they set up the dividend policies. The managers believed that dividend policy would affect the stock price.

Based on the responses from the financial managers of the 1,000 largest firms in the U.S., Pruitt and Gitman (1991) conclude that profits from the previous and the current year are important factors influencing dividend payments. Baker and Powell (2000) conduct a survey of NYSE-listed firms and find that dividend determinants are industry specific and that expected future earnings is the major determinant. In their surveys of chief financial officers (CFOs) of firms listed on the NYSE and NASDAQ, Baker and Powell (2000) and Baker et al. (2001) show that the "desire to conform to industry dividend practice" is an important factor influencing dividend policy. Both surveys report that about 45% of the respondents view this factor as of moderate-to-high importance.

The findings of Brav et al. (2005) are consistent with the evidence reported by Baker and Powell (2000) and Baker et al. (2001). Based on a recent survey of 384 corporate financial executives, they find peer-group effects for managers setting their dividends. Respondents from dividend-paying firms report that they consider the dividend policies of other industry members as an important influential factor in their own dividend decisions.

2.2 Literature Review on Peer Effects

In recent years, there is renewed interest among economists in peer effects. This renewed interest is reflected by a large number of studies that examine peer effects, neighborhood effects, and other social interaction effects. Peer effects are one type of neighborhood effects,[11] which Manski (1993 and 2000) calls endogenous effects in his studies. Manski identifies three effects associated with the similarity of observed behavior in a group, endogenous effects, exogenous effects, and correlated effects. According to this classification, an individual may be influenced by the behavior and characteristics of his/her peers, common individual factors, and similar environments.

Endogenous effects refer to how an individual's behavior is affected by the behavioral choices of his/her peers. For example, one might argue that a student's educational achievement is directly influenced by the achievements of his/her friends. In contrast, exogenous effects are present if an individual's behavior depends on the exogenous characteristics of his/her peers. In this case, a student's achievement may be influenced by the educational achievements of the parents of his/her peers. Unlike exogenous effects, endogenous effects generate a "social multiplier" (Cooper and John, 1988; Manski, 1993), that is, the effect of a policy intervention targeting on an individual will be amplified through its direct and indirect effect via the social interactions among peers.[12] Thus the presence of a social multiplier is very helpful for policy implementation. Exogenous effects do not have a social multiplier (Brock and Durlauf, 2001a). Because

---

[11] By Manski's (1993, 2000) categorization, neighborhood effects include endogenous, exogenous and correlated effects. Peer effects, contagion effects, and epidemic effects are other names of endogenous effects.

[12] See Glaeser, Sacerdote, and Scheinkman (1996) for a detailed discussion.

these two effects are by their nature quite different, they also have different policy implications. Thus, it is of great interest to policy makers to separately identify them.[13]

The term "correlated effect" is used by Manski (1993) to identify the case where an individual behaves similar to others in a group because he/she has similar individual characteristics, of which some are observed and some are unobserved. Correlated effects are also likely present in a group due to nonrandom group selection[14] or common shocks. For example, a high achieving student might choose to associate with high achieving classmates, or they are taught by the same teachers. Thus the observed group behavior may reflect correlated effects.[15] Similar to exogenous effects, correlated effects are not social effects and do not generate a social multiplier because the behavior is not impacted by exposure to peers.

Brock (1993) develops general statistical models to show how peer effects may explain volatility and market volume changes in financial markets. These models have been employed in recent work (e.g., Cont and Bouchard, 2000; Focardi et al., 2002) to explore the role of peer effects on the stock market crash. In the paper of Krauth (2003), peer groups are defined by the individuals who are directly and indirectly connected. The author modifies a standard job matching model by embedding this peer group information structure and shows how the composition of a peer group has a large effect on unemployment.

Theoretical models of peer effects have been developed inter alia by Akerlof and Kranton (2000), Becker and Murphy (2001), and Brock and Durlauf (2001b). Relative to

[13] The importance of distinguishing these two effects is explained in Moffitt (2001).
[14] This is also referred to as selection bias or a sorting process.
[15] For example, Wilson's (1987) research on the impact of poverty concentration explores correlated effects.

the advances in theory, the empirical studies of peer effects have moved quite slowly due to the multiple identification problems typically encountered in practice, including the reflection problem and the omitted variables bias.

The reflection problem is identified in Manski (1993). It refers to the problem of distinguishing between endogenous and exogenous (contextual) [16] effects. Correctly separating these two types of social effects from one another is necessary for evaluating the social net benefits that result from any behavioral intervention policy. Policy intervention on one individual's behavior may have the effect of changing the behavior of many individuals. Consider for example a homework help program. If individual academic achievement is positively affected by the average achievement of the class, then the homework help program not only directly helps those individuals who are in the program, but also indirectly others in the same class who are not participating. Although policy intervention is of little interest in the context of dividend decisions, an attempt to identify peer effects may be a key aspect to understanding what drives dividends.

Manski (1993) specifies a linear model in which three effects (endogenous, exogenous and correlated effects) are hypothesized to explain the observed similarity in the behavior of a group of individuals. In a linear model, an individual's predicted behavior is a linear function of the mean behavior of the group (endogenous effects), the average of the exogenous characteristics of the group members (exogenous effects), and individual characteristics (correlated effects). The author demonstrates that the identification of endogenous effects is impossible in such a linear model because the group average behavior is determined by the individual characteristics of the group

[16] The sociological literature calls it a contextual effect.

members. It is impossible to infer whether the group mean behavior causes the change of the individual behavior or vice versa.

Manski (2000) indicates that identification is eased if the peer effect occurs with a lag or is specified in a nonlinear manner, which includes binary response variable models[17]. Others, including Brock and Durlauf (2001a) and Moffitt (2001), investigate other alternatives for resolving the identification problem. So far, then, four solutions have been proposed to resolve the reflection problem (Brock and Durlauf, 2001a; Manski, 2000; Moffitt, 2001). First, lagged values of the group mean behavior are employed rather than contemporaneous ones (Manski, 2000). It may be more realistic to assume a time lag exists before an individual reacts to changes in the behavior of peers.

Second, when individual behavior is not a linear function of the group mean behavior, this nonlinear aspect may allow one to identify the estimated peer effects (Manski, 2000). Brock and Durlauf (2001a) point out that many of the behaviors we study are nonlinear. For example, homeownership is such a behavior because one either owns a house or not. The decision to buy a house is a function of one having a house, sales price, income, and other factors.

Third, when the individual behavior is assumed to vary with a distributional characteristic other than the mean behavior of the group, for example the median, then the reflection problem no longer exists. A fourth alternative to avoiding the reflection problem is to use an instrumental variable. Several more recent studies (Evans et al., 1992; Gaviria and Raphael, 2001; Hoxby, 2000; Ioannides and Zabel, 2002a and 2002b; Rivkin, 2001) have used instrumental variable estimation to correct for the reflection

[17] In a binary response model, the dependent variable is a binary random variable.

problem. These studies attempt to use credible instrumental variables for peer choices to achieve consistent estimation of the endogenous effects. Fertig (2003) incorporates both endogenous and exogenous effects and uses nonlinearity to identify both effects. Other examples include Drewianka (2003), Ioannides and Zabel (2002a, 2002b), Minkin (2002), and Sirakaya (2003). These studies find that both endogenous and exogenous effects are present.

It is possible that peer effects and correlated effects operate simultaneously. Thus, the separation of peer effects from correlated effects raises another identification problem. Similarity in behavior between individuals may result from similar individual characteristics or environmental exposure rather than the effects of exposure to peers. An individual often likes associating with those who have similar attributes as his/her. This commonality results from a process of self-selection rather than from a peer effect. Correlated effects likely arise when the issue of "self-selection" is present. Evans et al. (1992) and Rivkin (1997) have considered this important issue in their studies on peer effects. Evans et al. (1992) find that peer effects no longer exist once selection bias is controlled for. Rivkin (1997) criticizes the type of instrumental variable used in Evans et al. (1992) and argues that an experimental or quasi-experimental approach may correct for the selection bias. In the context of dividend behavior, Fama and French (2001) argue that newly listed firms are less likely to pay dividends because most of these firms tend to have a key common characteristic, significant growth opportunity combined with low profitability. Peer effect may likely play no role for these firms.

Some new identification strategies for peer effects include the development of programs controlling for randomness of group assignment. Empirical researchers, such as

Katz et al. (2001), Kremer and Levy (2001) and Sacerdote (2001) correct for self-selection bias by randomly assigning individuals to peer groups so that correlation effects are avoided. If there are no exogenous effects in such a setting, the coefficient on the peer variable can be interpreted as an endogenous or peer effect.

The obvious solution to solving the identification issue is to include a detailed set of individual characteristics as control variables in the estimation (Haurin et al. 2003). However, Weinberg et al. (2002) find that estimates of peer effects are still biased upward even though a rich set of control variables is included. This problem arises because it is extremely difficulty to account for unobserved individual characteristics in peer effects estimation.[18] Therefore, omitted variable bias is likely present and can cause a serious estimation problem. Several recent papers have acknowledged this difficulty and explained that their estimates could be upper bounds on the impact that peers have on individual behavior. Estimation results on peer effects should be more interesting if the identified upper bound is small.

Advanced econometric techniques permit one to use panel data to obtain fixed effects or first difference estimators (Haurin et al. 2003) to address the omitted variable bias. Aaronson (1998), Weinberg et al. (2002) have used this approach. Using experimental data is another approach and has been implemented in a number of studies (Katz et al. 2001; Ludwig et al. 2001; Rosenbaum and Harris, 2001).

There is little consensus on the best identification approach because each of the above strategies has limitations (Krauth, 2006). The instrumental variables method requires one to find instruments that are truly exogenous and that are also relevant to

[18] Aaronson (1998) and Plotnick and Hoffman (1999) use siblings data to control for unobserved characteristics.

the peer outcomes. This is likely to be very difficult in the group context. Random group assignment is good at avoiding correlated effects by eliminating any selection bias, but it is only applicable in a few special settings, such as government assisted housing because authorization has to be given to conduct the group assignment.

Lack of a detailed dataset and the determination of the functional form through which peer effects arise[19] are also important questions faced by researches. Unlike most of the previous studies, Boozer and Cacciola (2001) attempt to overcome the data limitation problem by using experimental data from the Project Star program in the state of Tennessee. The authors study the effects of educational achievements of students previously enrolled in a smaller class on the educational outcomes of their classmates in subsequent years. They find strong evidence of peer effects.

A number of authors have studied peer effects on student educational performance. Henderson et al. (1978) is one of the early most important studies in the economics literature. In their study, peer group effects are measured by the mean IQ of classmates. The authors find strong evidence that peer effects exist, and that these effects follow a concave functional form. The nonlinearity of peer effects is especially interesting. The concavity suggests that the marginal effect is decreasing with an increase in the level of mean classroom IQ. The authors conclude that improving mean IQ score is not an efficient way to maximize average achievement in a classroom because the increase in a student achievement would be slowed as the classroom mean IQ is further improved. Hanushek et al. (2003) conduct similar research and employ mean test scores for the same grade as a measure of peer effects. The authors find that

---

[19] Peer effects could exist in a linear or nonlinear manner.

a student's test score performance is improved by the mean test score performance. Some evidence in the study shows that nonlinearity of peer effects is present.

Hoxby and Terry (1999) attempt to explain the growing dispersion in wage inequality among college graduates. Their findings suggest that about forty percent of increasing wage inequality can be attributed to peer effects. Most studies on peer effects have focused on educational achievements, labor market success, and the behavior of disadvantaged youth rather than behavior related to financial markets. The majority of studies suggest that peer effects do matter. Effectively no quantitative study exists on the importance of peer effects or its operating manner[20] for a firm's dividend decision.

[20] Operating manner refers to whether peer effects exist in the linear or nonlinear fashion.

Chapter 3: Modeling Peer Effects

No work other than surveys of managers appears to exist to try to identify the existence of peer influence on dividend behavior of a firm. This study provides a first attempt to add peer effects to a traditional model of dividend behavior. To avoid attracting the attention of stock analysts, companies would want to avoid rapid changes in dividends. Consistent with previous studies, most of the dividend models estimated in this study assume that a firm's dividend payments are a function of the traditional control variables, such as firm size, profitability etc. The peer effects are incorporated through an additional variable (*PeerD*). The resulting regression equation is

$$D_t = f(PeerD_{t-1,} \text{control variables})$$

The peer effects variable enters into the above dividend regression with a time lag as $PeerD_{t-1.}$ By making dividends a function of the dividend behavior of a company's peers, the coefficient on the peer variable should ideally identify the peer effects. In particular, a positive and statistically significant variable would be expected if peer effects are effective.

A core problem of all attempts to measure peer effects is whether the peer effect is statistically identified. This is known as the reflection problem in the social interaction literature (Manski, 1993). As described by Manski (1993), a reflection problem refers to the failure of identification, in particular to the problem of distinguishing between endogenous and exogenous (contextual)[21] effects. A reflection problem arises in the present context to the extent that the dividend behavior of a firm and that of its peers are

[21] It is called a contextual effect in the sociological literature.

determined simultaneously. When a lagged as opposed to a contemporaneous peer variable is employed, the reflection problem can be avoided because the lagged peer variable is predetermined. This is the first option mentioned by Manski (2000) to help circumvent the reflection problem. Therefore, the above dividend regression model with a lagged peer variable is free of reflection problem.

However, using a lagged peer outcome may underestimate the peer influence because some concurrent peer effects can not be captured by the lagged peer variable (Hanushek et al., 2003). One alternative is to use a contemporaneous peer variable and find another way to circumvent the reflection problem. The third option mentioned by Manski (2000) to deal with the reflection problem suggests itself in this context. When an individual behavioral response is assumed to vary not in response to the mean behavior of the group but in response to some other metric, such as the median, the reflection problem is resolved. In the context of the dividend regression model, it is assumed that a firm's dividend behavior is affected by an average of peer behavioral outcomes in the same SIC industry group, where all firms in the SIC group are weighted equally.

If a key determinant of dividends, such as earnings, changes for all companies due to a change in the macroeconomic environment, and each company reacts to this change by changing its dividend, then a peer effect may be apparent when in fact it is non-existent. Thus, the exhibited similarity in the dividend behavior among companies is due to some common observed characteristics and similar environments faced by the companies. Manski (1993) calls this a correlated effect. A somewhat imperfect way to take into account the correlated effects is to include among the regressor variables identified as "other variables" in the above equation variables that capture events or developments that affect all companies. By the force of the Frisch-Waugh-Lovell theorem, including such variables should remove their effect from the peer variable. Including variables on the right-hand side that are highly sensitive to changes in the economic environment and central for a firm's dividend decision is likely to accomplish the same. The paper follows this route and tries to include variables such as earnings among the "other variables".[22]

In general, it is difficult to decide which variables to include among the set of "other variables" to control for the firm's dividend decision. This issue points to a weakness of the approach followed in this study. Although a rich set of explicit controls, such as firm or group characteristics, are included in the model, it is unlikely that all the determinants related to the dividend behavior are included in the estimation. Thus, omitted variables bias may still be present. The commonly used statistical techniques for correcting for an omitted variable bias include adjusting for group fixed effects (e.g. Aaronson, 1998; Brock and Durlauf, 2001a), employing a first difference estimator[23] or using experimental data (e.g. Sacerdote, 2001; Zimmerman, 2003).

Based on the survey results reported in the literature on the determinants of dividend behavior, it appears reasonable to assume that peer effects can be found primarily within an industry rather than across some other dimension, such as capitalization regardless of industry affiliation. The survey results also suggest that companies set dividends to conform to the expectations of professional stock market

[22] As an addition to placing variables, such as earnings, on the right-hand side of the dividend equation, one could think of replacing the variable that represents peer dividend behavior with a series that consists of the residuals of a regression of peer dividends on peer earnings, where peer earnings would be constructed with the same weight matrix as peer dividends.
[23] A panel data set is required so that unobserved heterogeneity can be eliminated by differencing.

analysts to avoid downgrades and a falling stock price. It is reasonable to assume that stock market analysts evaluate companies by industry. An individual company within an industry is typically compared to an industry standard. Industry averages serve as benchmarks (Akerlof, 2007; Shiller, 1984 and 1990). If one accepts this reasoning, the question arises how to define the industry that identifies the peer group of a particular company. In what follows, peer groups are defined by Standard Industry Classification (SIC) Codes. More specifically, firms are allocated into several major SIC industry groups.[24] That means that a firm from the mineral industry,[25] for example, can never be in the same peer group as a firm from manufacturing.[26] All firms from the utilities group (SIC, 4900–4949) and financial sector group (SIC, 6000–6999) are excluded because the dividend policies of these firms are influenced by regulation (DeAngelo et al., 2004; DeAngelo et al., 2006; Fama and French, 2001; Skinner, 2008).

The major SIC industry groups are very wide and are likely to lump companies together that would not ordinarily be considered peers by stock analysts or by company managers. The purpose of using wide definitions for the peer groups is that little prior information is required and little chance exists that peer groups are excluded that cut across more detailed SIC codes. However, beginning with a very wide definition of what could constitute a peer group is likely to affect the ability to claim that a significant coefficient truly identifies peer effects.

Generally, peer effects arise when a firm's dividend behavior is affected by its peer group's dividend decisions. Because of the role proximity plays in the concept of

---

[24] Standard Industrial Codes (SIC) come from http.//www.census.gov/epcd/naics/nsic2ndx.htm#S1.
[25] SIC code for the mineral industry is from 1000-1400.
[26] SIC code for the Manufacturing is from 2000-3900.

peer effects, a spatial component should be included. In the present context, however, spatial proximity is not defined by geographic proximity but by proximity in terms of SIC. Akerlof (1997) discusses a similar measure of proximity, distance in "social space". Applying these ideas to the present study, an individual firm is more significantly affected by the dividend decisions of firms that belong to the same category of SICs. This set-up is consistent with the findings of Rozeff (1982), who notes that similar dividend behaviors in a given industry.

Consider the following simple example of a "social" weight matrix (**W**) for the case with only three firms (n = 3),

$$w = \begin{bmatrix} 1 & w_{12} & w_{13} \\ w_{21} & 1 & w_{23} \\ w_{31} & w_{32} & 1 \end{bmatrix}.$$

W is a symmetric matrix with ones on the main diagonal, and its size is n x n. For instance, if the first two digits of the SIC number of firms 1 and 2 are identical, $w_{12} = w_{21} = 1$, otherwise, the elements are zeros. In order to create a spatial lag variable, the weight matrix **W** has to be standardized. The standardized weight matrix $\mathbf{W}^s$ is calculated by dividing each row element by the row sum after replacing the ones on the main diagonal with zeros. Thus, the elements in each row of $\mathbf{W}^s$ sum to unity,

$$w^s = \begin{bmatrix} 0 & \frac{w_{12}}{w_{12} + w_{13}} & \frac{w_{13}}{w_{12} + w_{13}} \\ \frac{w_{21}}{w_{21} + w_{23}} & 0 & \frac{w_{23}}{w_{21} + w_{23}} \\ \frac{w_{31}}{w_{31} + w_{32}} & \frac{w_{32}}{w_{31} + w_{32}} & 0 \end{bmatrix}.$$

Each element in a given row of the resulting normalized weight matrix ($\mathbf{W}^s$) indicates the weight that some other firm's dividend decision has for the dividend behavior of the firm that is identified by the given row. If each company in the industry group receives an equal weight, each weight is simply equal to $1/(n\text{-}1)$.

The spatial lag variable, which serves as a proxy for the peer variable, is constructed as the matrix product of $\mathbf{W}^s$ and the dependent variable vector $\mathbf{y}$, which contains the positive values of the dividend outcome variable for all firms in the selected sample. For the remainder of this study the dependent variable vector $\mathbf{y}$ contains the dividend per share (*DPS*). In constructing the spatial lag variable for the current period t, the following matrix product is employed

$$\mathbf{W}_{t-1}^{s}\mathbf{y}_{t-1}.$$

The weight matrix is based on values that relate to the previous year and so does the vector of dividends.

Chapter 4: Data and Empirical Methodology

4.1 Basic Data Considerations

The hypothesis addressed in this study is that the dividend decisions of other companies in the firm's peer group influence the firm's dividend behavior. This is tested empirically using the data from the Compustat database[27] for samples drawn from the S&P 1500 Super Composite Index and the S&P 500 index. Compustat databases contain fundamental financial and market data for U.S. corporations, banks, and industries, such as dividends and earnings, capital expenditures, stock prices, market capitalizations, firm value, book value of assets, and more. The reason I choose the firms in the S&P indices is that these firms are representatives for their respective industry groups while they still form a diverse set of firms.

Due to the change in dividend tax laws in 2003,[28] there is much change in dividend behavior around/after this point of time. This study is limited to explaining dividend behavior after 2003 using data from the years 2003 to 2006 through a series of independent cross-sections study. The panel estimations are based on the seven years of S&P 1500 data for the period from 2000 to 2006. All firms from the utilities and financial sectors [29] are excluded due to regulation issue. The regulated industries excluded encompass electric utilities, commercial banking, investment and brokerage services, as well as life, property, and casualty insurance.

---

[27] Compustat database is accessed through Research Insight.

[28] Jobs and Growth Tax Relief Reconciliation Act (JGTRRA) of May 2003.

[29] Utilities and financial firms are defined as firms with Standard Industrial Classification (SIC) codes between 4900-4949 and 6000-6999) respectively.

4.1.1 S&P 1500 Super Composite Index

The S&P 1500 Index is a capitalization-weighted index, the stock price is multiplied by the shares outstanding. This means that a firm with a higher market value has more influence on the index's performance than one with a lower market value. The index includes 1,500 companies that are in the S&P 500, S&P 400 and S&P 600 indices. The firms are chosen based on market capitalization, liquidity and industry representation. A number of firms are excluded from the sample. First, following previous studies,[30] I exclude firms incorporated in foreign countries[31] and select only U.S. publicly traded firms listed on the NYSE (New York Stock Exchange), AMEX (American Stock Exchange) and NASDAQ (National Association of Securities Dealers Automated Quotations) Stock Exchange[32] according to Compustat. This results in a sample of 1,486 companies. Second, all firms from the utilities and financial sectors are eliminated. This limits the sample to 1,138 companies.

Third, I restrict estimation to only those firms for which the decision to pay or not to pay a dividend has not changed in the current period compared to the past period. This is done to account for the fact that, compared with a decision of increasing or decreasing dividends, a firm's decision to initiate or terminate dividends is materially different. Baker and Wurgler (2004) find that that the average market reaction to dividend

---

[30] e.g., DeAngelo et al. (2004), Fama and Babiak (1968) and Fama and French (1999, 2001).

[31] S&P 1500 has 14 companies that are incorporated in the foreign countries. Specifically, there are eight companies incorporated in the Bermuda, four companies incorporated in the Cayman Islands, one incorporated in the Netherlands Antilles and one incorporated in the Panama.

[32] Full list of exchanges in the U.S can be found on http,//finance.yahoo.com/exchanges.

initiations is three times as large as that to dividend increases. Lie (2005) reports that the average market reaction to dividend omissions is twice as large as that to dividend decreases. Because the probability of a firm switching from being a payer to being a non-payer or vice versa is low, as these are "out-of-the-ordinary" decisions, such exclusion criteria will not lose many observations.

Finally, I include only firms that have valid values[33] of all variables required by the empirical analysis for a given year needed. Because data availability may vary from year to year, the total number of companies in the resulting sample may be different for each year. Table 1 presents the number of firms and dividend payers for the various S&P samples. There are 1,022 firms in the S&P 1500 sample for 2004, while there are 1,048 in the 2005 sample. Table 1 also reports the number of dividend payers for the years from 2004 to 2007.[34] The number of firms in the S&P 1500 paying dividends rose from 2004 to 2006, but started to fall in 2007. Despite the decreasing number of payers in 2007, the relative percentage of dividend payers has kept increasing over the four year period.

4.1.2 S&P 500 Index

Like the S&P 1500 Index, the S&P 500 is also a market-weighted index and comprises 500 large publicly traded companies in the United States. Companies included in the S&P 500 are traded on the NYSE or NASDAQ, the two largest American stock markets. The exclusion and screening criteria for the S&P 500 index are the same as

[33] Other than excluding observations because of missing values, additional Compustat data availability conditions are imposed when the empirical tests are conducted (see Appendix STATA codes for details).
[34] The dividend payer dummy is a dependent variable, as discussed in section 4.4.1. Data from the previous year are used to explain the likelihood of paying dividends of the current year. Because the data used for the independent variables are from 2003 to 2006, the data for the dependent variable is from 2004 to 2007. For example, data from 2003 is used to explain the dividend behavior for the year 2004.

those for the S&P 1500 index. This means, among other things, that companies incorporated outside the U.S.[35] are excluded as are utilities and financial firms.[36] Summary statistics on the annual samples of S&P 500 firms are given in Table 1 below those for the S&P 1500 samples.

For the purpose of comparison, the final sample sizes and the number of dividend payers for both the S&P 400 and the S&P 600 indices are also reported in Table 1. The S&P MidCap 400 Index consists of 400 medium-sized U.S.firms.[37] The size of the firms included in the S&P 400 is between that typical of the S&P 500 Index and that of the S&P SmallCap 600 Index. There are 96 firms excluded from the S&P 400 sample because of the restrictions regarding utilities, nonfinancial firms, and foreign firms. The S&P 600 index contains a diverse sample of 600 small-cap companies, with a market capitalization between $300 million and $2 billion. 126 regulated firms are in the S&P 600.[38] Similar to the other samples, those drawn from the S&P 600 also reveal an increasing trend in the percentage of dividend payers for the years 2004 to 2007. As it is expected,

[35] The S&P 500 index contains 13 companies that are incorporated in foreign countries. Specifically, there are seven companies incorporated in Bermuda, four companies incorporated in the Cayman Islands, one incorporated in the Netherlands Antilles and one incorporated in Panama.
[36] 360 firms are dropped out of the sample.
[37] A company is defined as a mid-cap stock when the market capitalization ranges from about $2 billion to $10 billion.
[38] No foreign firms are included in the S&P 600 index.

Table 1: Number of Dividend Payers by S&P Final Samples for the Years 2004-2007

Sample	2004	2005	2006	2007
S&P 1500 Super Composite Index				
Total Number of firms	1022	1048	1076	1027
Dividend payers	469	515	542	524
Percentage of dividend payers (%)	45.89	49.14	50.37	51.02
S&P 500				
Total Number of firms	318	335	338	323
Dividend payers	214	234	240	230
Percentage of dividend payers (%)	67.3	69.85	71.01	71.21
S&P MidCap 400				
Total Number of firms	278	283	286	269
Dividend payers	121	133	140	137
Percentage of dividend payers (%)	43.53	47	48.95	50.93
S&P SmallCap 600				
Total Number of firms	426	430	452	435
Dividend payers	134	148	162	157
Percentage of dividend payers (%)	31.46	34.42	35.84	36.09

Notes: A firm in the Compustat sample is defined as a dividend payer in year *t* if it has positive dividend per share by the exdate (Compustat item 26) in the fiscal year that ends in *t*. S&P Final Samples are obtained after three restrictions are imposed. First, the sample is restricted to nonfinancial, nonutility and domestic firms, which are traded on the NYSE, NASDAQ and AMEX. Second, a firm switching from being a payer to a non-payer or vice versa for a selected year is excluded. Third, a firm must have no missing values for Compustat items that are used to generate the dependent and independent variables (see Appendix D STATA codes for details).

Summary statistics of dividend per share by the four S&P final samples are shown in Table 2.[39] Over the four-year period, the S&P 500 has the highest average dividend payments and the S&P 600 has the lowest, which confirm the finding from previous studies that large firms are more likely to pay more dividends.

[39] The dividend per share is a dependent variable, as discussed in section 4.4.1. Data from the previous year are used to explain the likelihood of paying dividends of the current year. Because the data used for the independent variables are from 2003 to 2006, the data for the dependent variable is from 2004 to 2007. For example, data from 2003 is used to explain the dividend behavior for the year 2004.

Table 2: Summary Statistics of Dividend per Share by S&P Final Samples for the Years 2004 to 2007

Sample	N	Mean	Std.Dev.	Min.	Max.
Panel A: Descriptive Statistics of Dividend per Share By S&P Final Samples for Year 2004					
S&P 1500 Super Composite	1022	0.20	0.38	0.00	5.80
S&P 500	318	0.37	0.54	0.00	5.80
S&P MidCap 400	278	0.16	0.30	0.00	1.72
S&P SmallCap 600	426	0.10	0.22	0.00	2.18
Panel B: Descriptive Statistics of Dividend per Share By S&P Final Samples for Year 2005					
S&P 1500 Super Composite	1048	0.26	0.56	0.00	8.20
S&P 500	335	0.42	0.60	0.00	7.00
S&P MidCap 400	283	0.21	0.40	0.00	3.20
S&P SmallCap 600	430	0.17	0.59	0.00	8.20
Panel C: Descriptive Statistics of Dividend per Share By S&P Final Samples for Year 2006					
S&P 1500 Super Composite	1076	0.28	0.52	0.00	7.40
S&P 500	338	0.48	0.65	0.00	7.40
S&P MidCap 400	286	0.23	0.40	0.00	4.00
S&P SmallCap 600	452	0.15	0.43	0.00	7.26
Panel D: Descriptive Statistics of Dividend per Share By S&P Final Samples for Year 2007					
S&P 1500 Super Composite	1027	0.31	0.70	0.00	17.27
S&P 500	323	0.52	0.61	0.00	4.75
S&P MidCap 400	269	0.31	1.11	0.00	17.27
S&P SmallCap 600	435	0.15	0.28	0.00	2.12

Notes: Numbers represent the amount of annual cash dividend per share by exdate (Compustat item 26), adjusted for all stock splits and stock dividends that occurred during the period (in dollars and cents). S&P Final Samples are obtained after three restrictions are imposed. First, the sample is restricted to non-financial, nonutility and domestic firms that have been traded on the NYSE, NASDAQ and AMEX. Second, a firm switching from being a payer to a non-payer or vice versa for a selected year is excluded. Third, a firm must have non- missing values for Compustat items that are used to generate the dependent and independent variables (see Appendix D STATA codes for details).

4.2 Construction of the Spatial Lag Variable

4.2.1 Weight Matrix Calculations

A spatial lag variable is constructed only on those firms that pay a positive dividend. If it were constructed on all observations, including those firms that pay no dividends, then the weighted industry average would be lower than if the spatial lag variable were constructed only for those firms paying dividends.

The steps of constructing the spatial lag variable (*SPLAG*) are as follows. First, a weight matrix is created on the basis of the same SIC codes. There are three alternative weight matrices considered. One is based on the criterion that all firms have the same 2-digit SICs code. Another uses the same 3-digit SICs as a criterion and the last one is constructed for the same 4-digit SICs. In other words, an industry peer group is defined by the SICs.

For each of the three weight matrices, the entries are set equal to one if the first two, three, or four digits of the SICs are the same across firms and zero otherwise. The structure of the weight matrix implies that a firm is theoretically affected by all the other firms within the same industry peer group.

In order to create the spatial lag variable, the weight matrix has to be standardized. For that purpose, each row element is divided by the sum of the row after removing ones from main diagonal. For illustration purposes, Table 3 lists 10 firms that are randomly selected from the S&P 1500 sample using data from year 2006.

Table 3: 10 Randomly Selected Firms from the S&P 1500 Sample

Firm ID	Company Name	SIC Code	DPS_06
1	3M CO	2670	1.84
2	AT&T INC	4813	1.33
3	BIG 5 SPORTING GOODS CORP	5940	0.34
4	CVS CAREMARK CORP	5912	0.16
5	DISNEY (WALT) CO	4833	0.27
6	EASTMAN KODAK CO	3861	0.50
7	EXXON MOBIL CORP	2911	1.28
8	FACTSET RESEARCH SYSTEMS INC	7370	0.22
9	FAIR ISAAC CORP	7373	0.08
10	FAMILY DOLLAR STORES	5331	0.40

Notes: *DPS_06* is defined as dividend per share (in dollars and cents) for 2006. It is retrieved from the Compustat (item 26).

Table 3 tells that firms 2 and 5 have the same 2-digit SIC number (48); thus, they are peers if the peer group is defined by the first two digits of the SIC. Firms 3 and 4 also share the same 2-digit SIC (59), and so do firms 8 and 9 (73). Firms 8 and 9 also belong to the 3-digit SIC group of 737. None of the selected ten firms have the identical 4-digit SIC. According to the hypothesis, peers' dividend decisions would affect a firm's dividend decision making. For instance, firm 2 would consider dividends payments from peer firm 5 when making dividend decision.

If the peer group is defined by the first two digits of the SIC, the weight matrix should contain a one whenever the first two digits of the SIC are the same between any two firms; zero otherwise. This is illustrated by the weight matrix below. For example, the element in the second row and fifth column and the element in the fifth row and second column are equal to one because the first two digits of the SIC code of firms 2 and 5 are identical. The main diagonal contains all ones because a firm is considered to be its own peer.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Weight matrix based on 2-digit SIC

The corresponding weight matrix for 3-digit SIC industries is given as below. In addition to the ones on the main diagonal, there are ones in the eighth row and ninth column and an element in the ninth row and eighth column because firms 8 and 9 share the same first three digits of their SICs.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Weight matrix based on the same 3-digit SICs

Because none of the ten firms have the same 4-digit code, all elements off the main diagonal are zero in the weight matrix defined for 4-digit SIC industries.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Weight matrix based on the same 4-digit SICs

In order to create a spatial lag variable, the weight matrices have to be standardized. After the ones are removed from the main diagonal, the standardized weight matrix is derived by dividing each row element by the row sum.[40] Each row should sum to one except the rows with all zeros. The standardized weight matrices that correspond to the matrices shown above are presented below. The matrix for the 4-digit SIC category is left out as it contains all zeros.

[40] The row sum is calculated after the elements of the main diagonal are set to zero.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Standardized weight matrix based on the same 2-digit SICs

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Standardized weight matrix based on the same 3-digit SICs

4.2.2 Constructing the Spatial Lag Variables

The spatial lag variable is constructed as the matrix product of the standardized weight matrix and the dependent variable vector. For instance, if the standardized weight matrix $\mathbf{W}^s$ is of size nxn and the dependent variable vector y is of size $nx1$, then the spatial lag vector/variable $\mathbf{W}^s\mathbf{y}$ is of size $nx1$. In the dividend case, I use dividend per share (*DPS*) as the dependent variable. Therefore, the spatial lag variable is a weighted average of peers' dividend per share.

In response to the three alternatives in defining weight matrices, three alternative spatial lag variables are created based on equal 2-digit, 3-digit, or 4-digit SIC codes, respectively. To illustrate the methodology, I continue to use the same ten firms from the S&P 1500 index.

From Table 3, one can see that the 3M Company pays $1.84 dividends per share in 2006, which is higher than the other nine selected firms. FAIR ISAAC Corporation has the lowest dividend per share ($0.08) among the ten selected firms. Calculated spatial lag variables are reported in Table 4.

Table 4: Spatial Lag Variables for the Randomly Selected 10 Firms from the S&P 1500 Sample

Firm ID	Company Name	SIC Code	DPS_06	SPLAG2_06	SPLAG3_06	SPLAG4_06
1	3M CO	2670	1.84	0	0	0
2	AT&T INC	4813	1.33	0.27	0	0
3	BIG 5 SPORTING GOODS CORP	5940	0.34	0.16	0	0
4	CVS CAREMARK CORP	5912	0.16	0.34	0	0
5	DISNEY (WALT) CO	4833	0.27	1.33	0	0
6	EASTMAN KODAK CO	3861	0.50	0	0	0
7	EXXON MOBIL CORP	2911	1.28	0	0	0
8	FACTSET RESEARCH SYSTEMS INC	7370	0.22	0.08	0.08	0

Table 4: Spatial Lag Variables for the Randomly Selected 10 Firms from the S&P 1500 Sample

9	FAIR ISAAC CORP	7373	0.08	0.22	0.22	0
10	FAMILY DOLLAR STORES	5331	0.40	0	0	0

Notes: DPS_06 is defined as dividend per share (in dollars and cents) for 2006. It is retrieved from the Compustat (item # 26). *SPLAG2_06* is constructed as the product of the normalized weight matrix based on the 2-digit SICs and *DPS_06*. *SPLAG3_06* is constructed as the product of the standardized weight matrix based on the 3-digit SIC and *DPS_06*. *SPLAG4_06* is constructed as the product of the standardized weight matrix based on the 4-digit SIC and DPS_06.

Because firm 1 does not have any peers in the selected ten firms, thus the weighted average of peers' dividends per share is zero as reported in Table 4. One can that the weighted peers' dividends per share for firm 2 is \$0.27, this is the same amount as firm 5 pays due to the fact that firm 5 is the only peer in these ten firms. The similar relationship exists between the firms 3 and 4, firms 8 and 9.

4.3 Empirical Methodology

4.3.1 Cross-Sections, Heckman Selection Models

This study models the effects of peers' dividend decisions on a particular firm. The dividend decision can be thought of as a two-stage process where a firm first decides whether to pay a dividend or not, and second how much to pay. Heckman's (1979) selection model approach accounts for both decisions and the fact that they are interdependent; thus any bias resulting from both decisions being considered separately is avoided. Two equations are estimated,

Stage 1, $d_{it}^{*} = z_{it-1}\alpha + u_{it}$ (dividend-paying decision / selection equation)

Stage 2, $y_{it}^{*} = x_{it-1}\beta + \varepsilon_{it}$ (amount of dividends paid / outcome equation)

where

$$d_{it}^{*} = 1 \text{ if dividends are paid by firm } i \text{ in period } t$$

$$d_{it}^{*} = 0 \text{ if no dividends are paid by firm } i \text{ in period } t$$

and

$$y_{it}^{*} = \text{ dividends paid by firm } i \text{ in period } t \text{ if } d_{it}^{*} = 1.$$

The first regression describes the firm's choice of paying dividends or not. It is estimated as a logit regression. Variable d_{it}^{*} on the left side of the selection equation is a 0/1 indicator variable that identifies whether a dividend is paid or not *(DIVPY).* It equals one for firm i if the annual amount of dividend per share is positive in year t, and zero otherwise.

The second equation considers only firms paying positive amounts of dividends. This equation explains the size of the dividend that a firm pays. It can be estimated by ordinary least squares (OLS). The continuous dividend variable in the outcome equation $\left(y_{it}^{*}\right)$ equals dividend per share when firm i pays a dividend in period t $\left(d_{it}^{*} = 1\right)$. For the purpose of econometric identification, the selection equation should have at least one independent variable that is not included in the second outcome equation.

If one ignores the selection equation, and OLS is used on the subsample of dividend payers without any correction factor included, it would induce a sample selection bias in the estimation of β. Heckman (1979) characterizes the selection problem as a special case of the omitted variables problem because the correction factor, also known as the Inverse Mills Ratio (*IMR*), is being omitted in the second-stage equation. One corrects for this sample selection bias by including the *IMR* as an additional regressor in the second equation. The *IMR* can be derived from the predicted value of the logit selection equation.[41] Thus, the two-stage Heckman (1979) methodology involves estimating the first logit model, the selection equation, computing the *IMR*, and then adding it as an additional regressor in the second stage regression model, the outcome equation.[42]

4.3.2 Basic Model for a Panel Data

Because the Heckman two-stage estimation method only applies to cross-sectional models, one needs to find a way to correct for the sample selection bias in a panel data setting. There are several approaches discussed in the literature. However, there is no agreement on the best way to estimate panel data models while one considers sample selection issues.

In this study, I utilize the estimation method suggested by Wooldridge (1995) and Jackle and Himmler (2007) to correct for sample selection. Because the firms that are paying dividends belong to a self-selected sample, the selection equation is specified as,

$$DIVPY^{*}_{i,t} = \alpha_0 + z_{i,t}\alpha + k_i + e_{i,t}, \qquad (1)$$

where $DIVPY^{*}_{i,t}$ is the binary dependent variable, which equals one if a firm pays dividends, zero otherwise. $z_{i,t}$ represents the vector of explanatory variables, and

[41] The Inverse Mills Ratio (*IMR*) can be derived as the ratio of the standard normal density of the predicted value to the cumulative standard normal function of the predicted value.
[42] It is more efficient to use the one-step ML estimator rather than the two-step approach.

k_i contains unobserved characteristics that are assumed not to vary over time. $e_{i,t}$ is an error term and assumed to be uncorrelated with $z_{i,t}$ and k_i.

If a firm pays dividends ($DIVPY^*_{i,t}$ = 1), positive dividend per share (*DPS*) are observed. The outcome dividend equation is given as,

$$DPS^*_{i,t} = \beta_0 + x_{i,t}\beta + c_i + u_{i,t}, \tag{2}$$

where $DPS^*_{i,t}$ is an annual dividend per share of firm i at time t, $x_{i,t}$ is a vector of explanatory variables, c_i represents unobserved characteristics of firm i, and $u_{i,t}$ is the error term.

Following Wooldridge (1995), I write k_i as a linear projection onto the time averages of z_i, a constant θ_0 and an error term ξ_i, such as,

$$k_i = \theta_0 + \bar{z}_i\theta + \varepsilon_i. \tag{3}$$

Then, by substituting k_i into equation (1), one can rewrite the selection equation (1) as,

$$DIVPY^*_{i,t} = \gamma_0 + \bar{z}_i\theta + z_{i,t}\alpha + v_{i,t}, \tag{4}$$

where $\gamma_0 = \alpha_0 + \theta_0$ and $v_{i,t} = \varepsilon_i + e_{i,t}$.

Similar to the selection equation, I assume that the unobserved effect c_i can be written as a linear projection of the averages of x_i over time (denoted $\bar{x}_i$), a constant ϕ_0, and an error term b_i,

$$c_i = \phi_0 + \bar{x}_i\phi + b_i. \tag{5}$$

Thus, equation (2) can be rewritten as,

$$DPS^*_{i,t} = \omega_0 + \bar{x}_i\phi + x_{i,t}\beta + r_{i,t}, \tag{6}$$

where $\omega_0 = \beta_0 + \phi_0$ and $r_{i,t} = b_i + u_{i,t}$. In order to correct for the sample selection bias, one needs to add the inverse Mills ratios (*IMRs*) as an additional regressor into the dividend outcome equation. The *IMR*s are obtained by estimating (4) with standard probit estimation methods for each time period in the sample. Finally, the outcome equation is

$$DPS^{*}_{i,t} = \omega_0 + \bar{x}_i\phi + x_{i,t}\beta + \xi_t\lambda_{i,t} + r_{i,t}, \tag{7}$$

where $\lambda_{i,t}$ is the inverse Mills ratio (*IMR*) for firm *i* at time *t*. It is noted that IMR is differs by time period because the coefficient of *IMR* (ξ_t) varies over time.

4.4 Dependent and Independent Variables

4.4.1 Variables for Cross Sections

In order to deal with the sample selection problem of estimating the outcome equation only on dividend paying firms, Heckman's two-stage selection model is employed. The statistical analysis includes in the first step all the firms from a given sample. This includes dividend paying and non-paying firms. For the second stage, I exclude non-payers from the dataset and, thus, the sample is confined to those firms that pay dividends.

4.4.1.1 First Stage Variables

At the first stage, the dependent variable indicates whether or not a dividend is paid (*DIVPY*). It is defined as one if a firm *i* has a positive dividend per share (*DPS*) in year *t* and zero otherwise. This is the most commonly used dependent variable in the dividend literature (e.g., DeAngelo et al., 2006; Denis, 2008; Eije et al., 2008; Li and

Zhao, 2008). It has been confirmed in the literature that the dividend-paying decision is related in a statistically significant manner to firm size, profitability, a firm's growth (Fama and French, 2001), and a firm's retained earnings (DeAngelo et al., 2006).

I include these main determinants of the decision to pay dividends as proposed by DeAngelo et al. (2006), Fama and French (2001) and Skinner (2008) in the first stage selection model. Firm size (*SIZE*) is defined as the natural log of total assets (Skinner, 2008). As Fama and French (2001) conclude that a small size firm is less likely to pay dividends. Hence, a positive relationship between size and probability to pay a dividend is predicted. I follow DeAngelo et al. (2006) and Skinner (2008) in measuring profitability as the return on assets (*ROA*). Because of the traditionally strong relation between earnings and dividends, as documented by Fama and Babiak (1968), earnings adjusted for the effect of special items (*E*) (Skinner, 2008) is included as an alternative measure of profitability. More profitable companies might be better able to afford paying dividends (DeAngelo et al., 2006; Fama and French, 2001; Skinner, 2008). This view implies that a positive relationship should be observed between profitability and the likelihood to pay a dividend.

Growth is measured by asset growth (*ASG*) and the market-to-book ratio (*MTB*). Similar to Skinner (2008), I define asset growth as the change in total assets from the prior year and the market-to-book ratio as the market value of equity divided by the book value of common equity. Fama and French (2001) state that firms that never pay dividends have better growth opportunities. Therefore, one would expect a negative association between growth potential and the probability to pay a dividend.

DeAngelo et al. (2006) test the life-cycle theory of dividends and provide strong evidence that dividend decisions are tied to the stage of the life cycle that a company happens to be in. Specifically, they find that a firm is more likely to pay dividends if the ratio of retained earnings to total equity (*RETE*) is high. This is their proxy for the firm's life-cycle stage. I use the same life-cycle variable (*RETE*) suggested by the model of DeAngelo et al. (2006).

Based on the Jenson's (1986) agency theory, a firm is better off distributing free cash to shareholders as dividend payout in order to reduce agency costs. Consistent with the life-cycle theory, the likelihood to pay dividends is higher when the ratio of cash to total assets (*CTA*) is larger. A positive relationship is predicted between the probability to pay a dividend and *CTA*.

Having paid dividends in the previous year has been demonstrated to be a highly reliable indicator for a firm to pay dividends also in the current period. However, because our sample does not include dividend initiations and omissions, which means that a lagged indicator variable would always be identical to the dependent variable, the past dividend indicator variable should not be included in the first-stage regression model.

4.4.1.2 Second Stage Variables

In the second stage, I exclude non-payers and only study the amounts paid by dividend payers. The focus of the second-stage outcome equation is to find whether peer effects have a significant and positive impact on dividend payments. The dependent variable for the second-stage outcome equation is dividend per share-exdate (*DPS*) adjusted for stock splits and dividends, which is similar to Eije et al. (2008). Evidence by

Baker and Powell (2000) suggests that about three-quarters of firms make dividend decisions annually, thus all the empirical tests are conducted using annual dividends per share.

The key independent variable in the second-stage model is the peer effects variable. The spatial lag variable is the proxy for the peer effects. If there are peer effects, this variable should be significantly positively associated with a firm's dividend decision. As discussed in the section 4.2.3, three alternative spatial lag variables are created. Specifically, spatial lag variables *SPLAG2*, *SPLAG3* and *SPLAG4* are constructed based on the company sharing the same 2-digit, 3-digit, or 4-digit SIC code, respectively.

I use most of the control variables from the first stage to explain the dividends paid by payers in the second stage. These controls include Firm size (*SIZE*), return on assets (*ROA*), asset growth (*ASG*) and market-to-book ratio (*MTB*). Li and Lie (2006) report that firms tend to increase their dividends if they are large and profitable and the market-to-book ratio is low. Therefore, the coefficients for firm size and return on assets are expected to be positive, while the coefficient for the market-to-book ratio should be negative. Earnings per share (*ESP*), adjusted for stock splits and dividends, is added as an additional control variable. Variable definitions are given in Table 5. In order to reduce possible endogeneity problems, I lag all independent variables once.

The lagged dividend payments has been identified as one of the most important factors affecting a firm's yearly dividend decision (Lintner, 1956; Fama and Babiak, 1968; Benartzi, et al., 1997; DeAngelo, et al., 1992). However, some studies exclude the lagged dividend variable from the set of regressors (Fama and French, 2001; Fama and French, 2002; Li and Zhao, 2008; Denis et al., 2008). Fama and French (2001) suggest that profitability, investment opportunities, and firm size [43] are three characteristics that affect a firm's decision to pay dividends. They conclude that larger and more profitable firms are more likely to pay dividends. They argue that using lagged dividend status as an explanatory variable may be problematic, "We are interested in long-term dividend patterns. Under reasonable assumptions the regression approach that ignores lagged dividend status (Table 6) does a better job capturing the long-term effects of changing characteristics and propensity to pay...In this situation, regression function that ignores lagged dividend status captures the pre-1978 long term propensity to pay, given characteristics. And applying the base period regression function to the samples of firm characteristics of subsequent years produces estimates of the long-term effects of changing characteristics and propensity to pay".

Fama and French (2002) directly attempt to explain the dividend payout ratio[44] as a function of investment opportunities, profitability, target leverage, and other driving forces. The driving variables included in the study are profitability, investment opportunities and firm size.[45] Their results from the cross-section regression indicate that the dividend payout ratio is positively related to profitability and negatively related to investment opportunities and volatility.

Li and Zhao (2008) examine how informational asymmetries affect firms' dividend policies. They follow Fama and French (2001) to include four firm

---

[43] Fama and French (2001) use the ratio of a firm's earnings before interest to its total assets as a proxy for the profitability. The proxies for investment opportunities are a firm's rate of growth of assets and its market-to-book ratio.

[44] Dividend payout ratio is defined as the ratio of the dividends to net income for the same year.

[45] Fama and French (2002) use the ratio of earnings to assets as a proxy for profitability; the market-to-book ratio, the ratio of R&D to assets, and the growth in assets as proxies for investment opportunities; firm size (natural logarithm of total book assets) as a proxy for volatility. They find out that larger firms are likely to have less volatile earnings, higher dividend payouts.

characteristics variables, profitability, market to book ratio, asset growth and firm size. In addition to these control variables that may affect a firm's dividend policy, they add the analysts' earnings forecast errors and the dispersion in forecasts as proxies for the degree of information asymmetry. They find that firms with a higher degree of information asymmetry are less likely to pay or increase dividends. Their results do not support the signaling theory of dividends.

Denis et al. (2008) extend the Fama and French (2001) work by examining the evidence on the firms' likelihood to pay dividends in several developed countries. Their findings are consistent with those of Fama and French (2001), indicating that the likelihood of paying dividends is associated with firm size, growth opportunities, and profitability. As commonly observed, including past dividend payments improves the forecasting performance of the models. However, because forecasting is not the purpose of the current paper but structural long-run analysis is, this study follows the example of Fama and French (2001) and the follow-up papers and does not include lagged dividend per share as an explanatory variable.

Table 5: Variable Definitions

Variable	Definition
First-stage dependent variables	
DIVPY	1 if the annual amount of dividend per share is positive, 0 otherwise.
Second-stage dependent variables	
DPS	Dividend per share –exdate (Compustat item 26), adjusted for stock splits and dividends
Independent variables	
SPLAG2#	Spatial lag variable; constructed as the matrix product of the weight matrix for 2-digit SIC codes and dividend per share (Compustat item 26)
SPLAG3#	Spatial lag variable; constructed as the matrix product of the weight matrix for 3-digit SIC codes and dividend per share (Compustat item 26)
SPLAG4#	Spatial lag variable; constructed as the matrix product of the weight matrix for 4-digit SIC codes and dividend per share (Compustat item 26)
*SIZE**	Natural log of total assets (Compustat item 6)
*ROA**	Return on assets; defined as operating income before depreciation (Compustat item 13) divided by total assets (Compustat item 6)
E_t@	Earnings adjusted for the effect of special items (Compustat item 18 - 0.6 x Compustat item 17)
ESP#	Earnings per share (Compustat item 58) before extraordinary items, adjusted for stock splits and dividends
*ASG**	Asset growth; change in total assets from the prior year (Compustat item 6)
*MTB**	Market -to-book ratio; market value of equity (Compustat item 25 x Compustat item 199) divided by book value of common equity (Compustat item 60)
RETE@	The ratio of retained earnings (Compustat item 36) to total equity (Compustat item 60)
CTA@	The ratio of cash (Compustat item 1) to total assets (Compustat item 6)

Notes: For cross sections, all independent variables are lagged one year. Values equal those at the end of the fiscal year. The independent variables with superscript * indicate that they are included in both stages of the Heckman selection models; the independent variables with superscript # indicate that they are included only in the second stage of the Heckman selection models; the independent variables with superscript @ indicate that they are included only in the first stage of the Heckman selection models.

4.4.2 Variables for Panel Estimation

In the selection equation (4), as derived in section 4.3.2, the dependent variable is the 0/1 indicator variable indicating dividend payment (*DIVPY*). The explanatory variables $z_{i,t}$ include earnings (*E*), firm size (*SIZE*), return on assets (*ROA*), asset growth

(*ASG*), market-to-book ratio (*MTB*), the ratio of retained earnings to total equity (*RETE*), and the ratio of cash to total assets (*CTA*), as defined in Table 5.

The dependent variable in the outcome equation (7) is defined as dividend per share (*DPS*). Only observations with positive values of *DPS* are included. The vector of explanatory variables $x_{i,t}$ contains firm size (*SIZE*), return on assets (*ROA*), asset growth (*ASG*), market-to-book ratio (*MTB*), earnings per share (*ESP*) and one of the spatial lag variables (*SPLAG2, SPLAG3 or SPLAG4*). The variables are defined as given in Table 5.

4.5 Summary Statistics for Variables

4.5.1 Summary Statistics for Cross Sections

Table 6 gives summary statistics for the explanatory variables and the dependent variables for the sample of firms listed in the S&P 1500 index. Because only firms with positive values of dividend per share are used in the second-stage equation, the number of observations for the second-stage regression is always far less than the number of observations for the first-stage equation. For instance, in 2004, the number of observations used for the second-stage equation is 469 while 1,021 firms are used for the first-stage equation. This shows that about 46 % of the 1,021 firms in the 2004 sample are dividend payers. Over the four year period, the percentage of dividend payers has slightly increased from 46 % in 2004 to 51% in 2007.

In Table 6, dividend per share (*DPS*) is the dependent variable in the second – stage OLS regression. Hence, the mean value of *DPS* is computed only on those firms with a positive amount of dividends per share. It should be much higher than if *DPS* is calculated for all observations, including those firms that pay no dividends. For example, in 2004, the average of *DPS* is \$0.53 for S&P dividend payers in Table 6, while, the average dividend payment per share in Table 2 is only \$0.20. The latter figure is much lower because it includes many firms with zero values for variable *DPS*.

Consistent with Fama and French (2001), Table 6 also shows that dividend payers are more profitable and larger than all other firms including nonpayers.[46] Dividend payers typically exhibit less asset growth[47] and also tend to have lower market-to-book ratios,[48] although the differences in means are small. Table 6 indicates that cash ratios (*CTA*) have moved little across the four year period.

It is expected that firms have more peers if the peer group is defined broadly, that is if it is based on the 2-digit SIC codes rather than the 3-digit or 4-digit SIC codes. Therefore, the weight matrices for the 3-digit or 4-digit industry classifications have more zero entries than the one based on the equivalent 2-digit SIC code. This has the result that the average value of *SPLAG2* is higher than that of *SPLAG3* and of *SPLAG4*. For example, *SPLAG2*, *SPLAG3*, and *SPLAG4* are \$0.65, \$0.58 and \$0.55, respectively, in 2007. The means of the three spatial lag variables are all smaller than the mean of the dividends per share (*DPS*) for a given year. For instance, *DPS* is \$0.63 in 2006, while *SPLAG2*, *SPLAG3* and *SPLAG4* are \$0.54, \$0.47, and \$0.44, respectively.

---

[46] Compare mean values of *ROA* and *SIZE* for both stages, where the second-stage uses only dividend payers.
[47] See the mean values of variable *ASG* for both stages.
[48] Compare the means values of *MTB* for both stages.

Table 6: Summary Statistics of Dependent and Independent Variables for S&P 1500 Samples

Year	Stage	Variables	N	Mean	Std.Dev.	Min	Max
2004		*Dependent*					
	First	*DIVPY*	1021	0.46	0.50	0.00	1.00
	Second	*DPS*	469	0.53	0.65	0.02	8.20
2003		*Independent*					
	First	*E*	1021	293.77	1058.56	-1929.00	15589.00
	First	*SIZE*	1021	7.14	1.59	1.89	13.38
	Second		469	7.83	1.55	4.20	13.38
	First	*ROA*	1021	0.14	0.09	-0.38	0.74
	Second		469	0.15	0.08	-0.05	0.74
	First	*ASG*	1021	0.15	0.33	-0.48	5.00
	Second		469	0.10	0.15	-0.38	1.52
	First	*MTB*	1009	3.56	4.24	0.63	75.70
	Second		466	3.35	3.67	0.63	42.77
	First	*RETE*	1009	0.29	2.41	-40.03	18.58
	First	*CTA*	1021	0.18	0.19	0.00	0.91
	Second	*SPLAG2*	469	0.44	0.25	0.00	2.32
	Second	*SPLAG3*	469	0.39	0.33	0.00	2.64
	Second	*SPLAG4*	469	0.35	0.36	0.00	2.64
	Second	*EPS*	469	1.38	1.92	-22.04	25.19
2005		*Dependent*					
	First	*DIVPY*	1046	0.49	0.50	0.00	1.00
	Second	*DPS*	514	0.57	0.65	0.01	7.40
2004		*Independent*					
	First	*E*	1046	368.07	1233.69	-4230.76	16819.00
	First	*SIZE*	1046	7.28	1.55	2.55	13.53
	Second		514	7.89	1.55	4.07	13.53
	First	*ROA*	1046	0.15	0.09	-0.26	0.69
	Second		514	0.16	0.07	-0.03	0.59
	First	*ASG*	1046	0.18	0.33	-0.62	3.76
	Second		514	0.13	0.21	-0.62	1.68
	First	*MTB*	1039	3.84	6.25	0.68	107.51
	Second		513	3.61	5.44	0.68	107.51
	First	*RETE*	1039	0.34	2.27	-42.65	12.85
	First	*CTA*	1046	0.18	0.18	0.00	0.88
	Second	*SPLAG2*	514	0.54	0.31	0.00	3.20
	Second	*SPLAG3*	514	0.47	0.45	0.00	3.53
	Second	*SPLAG4*	514	0.44	0.63	0.00	8.20
	Second	*EPS*	514	1.76	2.35	-17.56	34.69
2006		*Dependent*					
	First	*DIVPY*	1076	0.50	0.50	0.00	1.00
	Second	*DPS*	542	0.63	0.94	0.01	17.27
2005		*Independent*					
	First	*E*	1076	419.70	1294.45	-2586.94	18633.00
	First	*SIZE*	1076	7.36	1.53	3.75	13.42
	Second		542	7.93	1.52	4.71	13.42
	First	*ROA*	1076	0.16	0.09	-0.37	0.87
	Second		542	0.17	0.08	0.03	0.68
	First	*ASG*	1076	0.16	0.36	-0.51	5.02
	Second		542	0.10	0.23	-0.51	1.69
	First	*MTB*	1066	3.95	6.88	0.77	127.62

Table 6: Summary Statistics of Dependent and Independent Variables for S&P 1500 Samples

Year	Stage	Variables	N	Mean	Std.Dev.	Min	Max
	Second		540	3.54	4.89	0.88	88.39
	First	*RETE*	1066	0.33	2.72	-50.59	7.78
	First	*CTA*	1076	0.18	0.18	0.00	0.90
	Second	*SPLAG2*	542	0.54	0.31	0.00	4.00
	Second	*SPLAG3*	542	0.47	0.39	0.00	4.00
	Second	*SPLAG4*	542	0.44	0.44	0.00	4.00
	Second	*EPS*	542	2.02	2.80	-20.57	32.66
2007		*Dependent*					
	First	*DIVPY*	1027	0.51	0.50	0.00	1.00
	Second	*DPS*	524	0.79	1.57	0.02	25.17
2006		*Independent*					
	First	*E*	1027	534.10	1968.25	-723.40	39500.00
	First	*SIZE*	1027	7.51	1.50	3.88	13.45
	Second		524	8.04	1.52	4.50	13.45
	First	*ROA*	1027	0.16	0.09	-0.48	0.76
	Second		524	0.17	0.08	-0.01	0.76
	First	*ASG*	1027	0.17	0.36	-0.40	4.00
	Second		524	0.10	0.22	-0.40	2.59
	First	*MTB*	1019	3.93	8.41	0.65	162.63
	Second		521	4.13	10.79	0.86	162.63
	First	*RETE*	1019	0.59	3.06	-42.57	76.02
	First	*CTA*	1027	0.16	0.17	0.00	0.91
	Second	*SPLAG2*	524	0.65	0.42	0.00	3.04
	Second	*SPLAG3*	524	0.58	0.64	0.00	6.40
	Second	*SPLAG4*	524	0.55	0.67	0.00	6.40
	Second	*EPS*	524	2.23	1.85	-10.00	11.88

Notes: The table reports summary statistics for the dependent and independent variables that are used in Heckman's two-stage selection models. Data from the previous year are used to explain the dividend behavior of the current year. Because the data used for the independent variables are from 2003 to 2006, the data for the dependent variable is from 2004 to 2007. For example, data from 2003 is used to explain the dividend behavior for the year 2004. The dependent variable in the first-stage is *DIVPY*, which is coded one for firms/years with positive values of dividends per share (as defined in Table 5) and zero otherwise. The dependent variable in the second-stage is *DPS* (Compustat item 26, dividends per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables for the first-stage equation are: (1) *SIZE*; (2) *ROA*; (3) *ASG*; (4) *MTB*; (5) *RETE*; (6) *CTA*; (7) *E*. The independent variables for the second-stage equation are: (1) *SIZE*; (2) *ROA*; (3) *ASG*; (4) *MTB*; (5) *SPLAG2/ SPLAG3/ SPLAG4* (varies with the model); (6) *EPS*. Independent variable definitions are given in Table 5.

One can observe a number of similarities for the first-stage and second-stage variables of the S&P 1500 and S&P 500 samples (Table 7). Firm sizes and profits of dividend payers are larger for companies from the S&P 500 index than for those from the S&O 1500 index. However, dividend payers from the S&P 500 index tend to have less growth potential and smaller market-to-book ratios. Since the S&P 500 index consists of

500 large publicly traded companies in the United States, it is expected that the percentage of dividend payers should be higher in the S&P 500 than in the S&P 1500 index. As Table 7 shows, about 67 % of the firms in the S&P 500 pay dividends in 2004 and about 71 % do in 2007. As revealed in Table 6, only 46 % of the S&P 1500 companies pay dividends in 2004. The percentage rises to 51 % in 2007.

Average dividend payments per share (*DPS*) for S&P 500 firms, as shown in Table 7, are generally higher than those for S& P 1500 firms, as reported in Table 6. For example, in 2004, average *DPS* is $0.62 for the S&P 500 sample, while the average *DPS* for the S&P 1500 sample is $0.53. As explained in the previous chapter, the spatial lag variable is the proxy for the peer variable. For each firm, it is the weighted average of the dividend per share (*DPS*) paid by all other firms in the peer group, which is defined in terms of SICs. Table 7 reveals that the mean of the spatial lag variable is typically larger for the peer group defined in terms of 2-digit SIC codes than for the other two peer groups. The average values of the spatial lag variable for the S&P 500 index are higher than the mean values for the S&P 1500 index.

Table 7: Summary Statistics of Dependent and Independent Variables for S&P 500

Year	Stage	Variables	N	Mean	Std.Dev.	Min	Max
2004		*Dependent*					
	First	*DIVPY*	318	0.67	0.47	0.00	1.00
	Second	*DPS*	214	0.62	0.65	0.02	7.00
2003		*Independent*					
	First	*E*	318	847.18	1773.32	-1929.00	15589.00
	First	*SIZE*	318	8.82	1.21	5.63	13.38
	Second		214	9.09	1.18	6.81	13.38
	First	*ROA*	318	0.15	0.08	-0.12	0.74
	Second		214	0.16	0.08	-0.05	0.74
	First	*ASG*	318	0.14	0.40	-0.46	5.00
	Second		214	0.10	0.16	-0.38	1.52
	First	*MTB*	312	4.36	4.31	0.63	42.77
	Second		212	4.28	4.56	0.63	42.77
	First	*RETE*	312	0.62	2.94	-40.03	18.58

Table 7: Summary Statistics of Dependent and Independent Variables for S&P 500

Year	Stage	Variables	N	Mean	Std.Dev.	Min	Max
	First	*CTA*	318	0.15	0.17	0.00	0.84
	Second	*SPLAG2*	214	0.53	0.38	0.00	2.32
	Second	*SPLAG3*	214	0.40	0.46	0.00	3.39
	Second	*SPLAG4*	214	0.36	0.47	0.00	3.39
	Second	*EPS*	214	1.58	2.60	-22.04	25.19
2005		*Dependent*					
	First	*DIVPY*	334	0.70	0.46	0.00	1.00
	Second	*DPS*	233	0.68	0.68	0.01	7.40
2004		*Independent*					
	First	*E*	334	1024.67	2029.92	-4230.76	16819.00
	First	*SIZE*	334	8.91	1.17	5.21	13.53
	Second		233	9.14	1.15	6.99	13.53
	First	*ROA*	334	0.16	0.08	-0.09	0.59
	Second		233	0.17	0.07	0.02	0.59
	First	*ASG*	334	0.13	0.25	-0.34	2.80
	Second		233	0.12	0.18	-0.30	1.68
	First	*MTB*	331	4.73	7.15	0.83	107.51
	Second		233	4.62	7.73	0.83	107.51
	First	*RETE*	331	0.55	3.05	-42.65	12.85
	First	*CTA*	334	0.16	0.17	0.00	0.79
	Second	*SPLAG2*	233	0.58	0.39	0.00	2.45
	Second	*SPLAG3*	233	0.46	0.54	0.00	4.02
	Second	*SPLAG4*	233	0.41	0.55	0.00	4.02
	Second	*EPS*	233	2.13	2.86	-17.56	34.69
2006		*Dependent*					
	First	*DIVPY*	338	0.71	0.45	0.00	1.00
	Second	*DPS*	240	0.78	0.76	0.02	7.80
2005		*Independent*					
	First	*E*	338	1185.41	2112.59	-2586.94	18633.00
	First	*SIZE*	338	9.00	1.11	6.77	13.42
	Second		240	9.21	1.09	7.18	13.42
	First	*ROA*	338	0.17	0.09	-0.07	0.68
	Second		240	0.18	0.08	0.03	0.68
	First	*ASG*	338	0.14	0.40	-0.38	3.88
	Second		240	0.10	0.23	-0.37	1.58
	First	*MTB*	334	4.81	7.72	0.88	88.39
	Second		239	4.43	6.94	0.88	88.39
	First	*RETE*	334	0.49	3.32	-50.59	6.39
	First	*CTA*	338	0.16	0.16	0.00	0.78
	Second	*SPLAG2*	240	0.65	0.43	0.00	2.63
	Second	*SPLAG3*	240	0.51	0.55	0.00	4.26
	Second	*SPLAG4*	240	0.45	0.57	0.00	4.26
	Second	*EPS*	240	2.71	3.19	-10.54	32.66
2007		*Dependent*					
	First	*DIVPY*	323	0.71	0.45	0.00	1.00
	Second	*DPS*	230	0.86	0.91	0.02	11.30
2006		*Independent*					
	First	*E*	323	1519.37	3301.94	-723.40	39500.00
	First	*SIZE*	323	9.14	1.12	6.88	13.45

Table 7: Summary Statistics of Dependent and Independent Variables for S&P 500

Year	Stage	Variables	N	Mean	Std.Dev.	Min	Max
	Second		230	9.34	1.12	7.27	13.45
	First	*ROA*	323	0.17	0.09	-0.01	0.76
	Second		230	0.18	0.09	0.02	0.76
	First	*ASG*	323	0.15	0.38	-0.40	4.00
	Second		230	0.11	0.27	-0.40	2.59
	First	*MTB*	318	5.50	13.74	0.98	162.63
	Second		228	5.75	15.94	0.98	162.63
	First	*RETE*	318	0.87	5.11	-42.57	76.02
	First	*CTA*	323	0.14	0.14	0.00	0.76
	Second	*SPLAG2*	230	0.75	0.51	0.00	4.75
	Second	*SPLAG3*	230	0.58	0.55	0.00	4.50
	Second	*SPLAG4*	230	0.52	0.57	0.00	4.50
	Second	*EPS*	230	2.81	2.00	-3.93	11.88

Notes: The Table reports summary statistics for dependent and independent variables of Heckman two-stage selection models. Data from the previous year are used to explain the dividend behavior of the current year. Because the data used for the independent variables are from 2003 to 2006, the data for the dependent variable is from 2004 to 2007. For example, data from 2003 is used to explain the dividend behavior for the year 2004.The dependent variable in the first-stage is *DIVPY*, which is coded one for firms/years with positive values of dividends per share (as defined in Table 5) and zero otherwise. The dependent variable in the second-stage is *DPS* (Compustat item 26, dividends per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables for the first-stage equation are: (1) *SIZE*; (2) *ROA*; (3) *ASG*; (4) *MTB*; (5) *RETE*; (6) *CTA*; (7) *E*. The independent variables for the second-stage equation are, (1) *SIZE*; (2) *ROA*; (3) *ASG*; (4) *MTB*; (5) *SPLAG2/ SPLAG3/ SPLAG4* (varies with the model); (6) *EPS*. Independent variable definitions are given in Table 5.

4.5.2 Summary Statistics for Panel Data

The S&P 1500 data for the seven years from 2000 to 2006 are employed to examine the impact of peer effects on dividend decisions in a panel data framework. As for the cross section analysis, the dependent variable of the outcome equation consists of the positive values of the variable dividend per share (*DPS*). Among the independent variables are the four traditional controls (*SIZE*, *ROA*, *ASG* and *MTB*). In addition, I include the variable *ESP*, seven *IMRs*, six year-specific dummy variables and one of the three spatial lag variables (*SPLAG2, SPLAG3, or SPLAG4)*. Table 8 reports summary statistics for the outcome equation that is related to the S&P 1500 panel.

The original panel contains 10,500 observations, which are made up of seven years and 1,500 firms per year. After dropping foreign firms, utilities and financial firms, 7,966 observations remain. The sample drops down to 3,554 observations the second-stage equation after removing observations with missing data or no dividend payments. The sample declines to 2,294 observations if cases of dividend initiation and termination are excluded.

Table 8: Summary for the Second-Stage Outcome Dividend Equation, Panel S&P 1500, Years 2000-2006

Variables	N	Mean	Std.Dev.	Min	Max
Dependent					
DPS	2294	0.49	0.57	0.01	17.27
Independent					
SIZE	2294	7.86	1.54	4.18	13.53
SIZEbar	2294	7.86	1.52	4.39	13.31
ROA	2294	0.16	0.07	-0.08	0.53
ROAbar	2294	0.16	0.06	0.02	0.50
ASG	2294	2.12	21.21	-1.00	506.99
ASGbar	2294	2.11	7.78	-0.16	72.43
MTB	2294	3.39	4.95	0.41	162.63
MTBbar	2294	3.39	3.71	0.72	50.12
SPLAG2	2294	0.51	0.36	0.00	5.18
SPLAG2bar	2294	0.52	0.27	0.00	2.33
SPLAG3	2294	0.45	0.48	0.00	7.76
SPLAG3bar	2294	0.45	0.36	0.00	2.65
SPLAG4	2294	0.42	0.54	0.00	8.20
SPLAG4bar	2294	0.43	0.40	0.00	2.65
EPS	2294	1.55	1.57	-22.04	11.78
EPSbar	2294	1.55	1.00	-1.22	5.90
IMR2000	2294	0.09	0.26	0.00	2.08
IMR2001	2294	0.09	0.26	0.00	2.04
IMR2002	2294	0.08	0.25	0.00	2.85
IMR2003	2294	0.07	0.23	0.00	2.39
IMR2004	2294	0.07	0.21	0.00	2.16
IMR2005	2294	0.07	0.20	0.00	1.92
IMR2006	2294	0.06	0.19	0.00	1.72
yd1	2294	0.14	0.35	0.00	1.00
yd2	2294	0.14	0.35	0.00	1.00
yd3	2294	0.14	0.35	0.00	1.00
yd4	2294	0.14	0.35	0.00	1.00
yd5	2294	0.14	0.35	0.00	1.00
yd6	2294	0.14	0.35	0.00	1.00

Notes: The dependent variable is *DPS* (Compustat item 26, dividend per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables are, (1) *SIZE*; (2) *ROA*; (3) *ASG*; (4) *MTB*; (5) *SPLAG2/ SPLAG3/ SPLAG4* (varies with the model); (6) *EPS*; (7) averages for each explanatory variable. For instance, *MTBbar* is the average for the *MTB* over the seven years, thus *MTBbar= (MTB2000+MTB2001+MTB2002+MTB2003+MTB2004+MTB2005+MTB2006)/7* and this value is used for every year for a firm *i* in the panel; (8) seven *IMRs*. For example, *IMR 2000* is computed from the probit selection model using 2000 data and (9) six year dummies *yd1-yd6* since there are seven years. For instance, *yd1* equals one if the year is 2000 and zero otherwise. Definitions of the independent variables are given in Table 5.

Chapter 5: Empirical Models and Estimation Results

5.1 Cross-Sectional Empirical Models

Three cross-sectional empirical models are estimated using Heckman's two-stage estimation method. Each model contains two equations, i.e., selection equation in the first-stage and outcome interest equation in the second-stage. The first-stage selection equation for Model 1(*M1*) is

$$DIVPY_{it} = \alpha + \beta_1 E_{it-1} + \beta_2 SIZE_{it-1} + \beta_3 ROA_{it-1} + \beta_4 ASG_{it-1} + \beta_5 MTB_{it-1} \qquad (8)$$

$$+ \beta_6 RETE_{it-1} + \beta_7 CTA_{it-1} + \varepsilon_{it},$$

where α is a constant term, β_j is the regression coefficient for the j^{th} independent variable and ε is the residual error term. The first-stage equation is estimated as a logit regression. As discussed in Chapter 4, one can derive the *IMR* from the selection model and include it as an additional independent variable in the second-stage equation, which corrects the sample selection problem.

In Model 1, the dependent variable of the selection equation is an indicator variable $DIVPY_{it}$. It equals one for firm i if the annual dividend per share paid is positive in year t, and zero otherwise. The explanatory variables included in the first-stage equation are earnings (*E*), firm size (*SIZE*), return on assets (*ROA*), asset growth (*ASG*), market-to-book ratio (*MTB*), ratio of retained earnings to total equity (*RETE*), and ratio of cash to total assets (*CTA*). All variables are defined in Table 5 of Chapter 4.

The second-stage outcome equation of Model 1 can be written as

$$DPS_{it} = \alpha + \beta_1 EPS_{it-1} + \beta_2 SIZE_{it-1} + \beta_3 ROA_{it-1} + \beta_4 ASG_{it-1} + \beta_5 MTB_{it-1} + \beta_6 SPLAG2_{it-1} + \varepsilon_{it}, \tag{9}$$

where the coefficient of the peer proxy (β_6) should be positive and significant if firms try not to deviate too much from what their peers are doing in terms of dividend policy.

In Model 1, the dependent variable of the second equation is a continuous dividend variable (*DPS_{it}*). It is defined in Table 5 of Chapter 4. It should be noted that only observations for which dividend per share are positive are used in the OLS regression. In addition to the peer variable (*SPLAG2*), four of the traditional variables used to explain the probability of paying dividends are also included in the second-stage equation, firm size (*SIZE*), return on assets (*ROA*), asset growth (*ASG*) and the market-to-book ratio (*MTB*). I also include earnings per share (*ESP*) as a control variable. The variable definitions are given in Table 5.

For Models 2 and 3, the first-stage selection equations are identical as one in Model 1. But the second-stage outcome equations vary for three models because of the use of the different spatial lag variable. The main focus in the second-stage regression is to identify the existence of peer effects. Thus three models are estimated, each with a different peer variable. In order to reduce possible endogenity problems, I lag all independent variables once.

The second-stage equation of Model 2 (*M2*) is given as

$$DPS_{it} = \alpha + \beta_1 EPS_{it-1} + \beta_2 SIZE_{it-1} + \beta_3 ROA_{it-1} + \beta_4 ASG_{it-1} + \beta_5 MTB_{it-1} + \beta_6 SPLAG3_{it-1} + \varepsilon_{it}, \tag{80}$$

while the second-stage equation of Model 3 (*M3*) is formed as

$$DPS_{it} = \alpha + \beta_1 EPS_{it-1} + \beta_2 SIZE_{it-1} + \beta_3 ROA_{it-1} + \beta_4 ASG_{it-1} + \beta_5 MTB_{it-1} + \beta_6 SPLAG4_{it-1} + \varepsilon_{it}, \quad (91)$$

where the parameters and variables are defined as above.

5.2 Cross-Sectional Estimation Results

Cross-sectional regressions for three alternative models are estimated on a sample of S&P 1500 and S&P 500 firms using data from 2003 to 2006. Heckman's two-stage estimation method is employed in order to correct for sample selection bias. The estimation results for each regression model are provided in Tables 9 and 10. Each table reports the regression results for a particular S&P sample. By construction, peer effects are measured only to the extent that they exist within the same SIC group.

Consistent with the findings of DeAngelo et al. (2006) and Fama and French (2001), the four traditional variables (*SIZE, ROA, ASG* and *MTB*) used in explaining the probability of a firm paying dividends also prove to be significant and with the expected sign in the first-stage selection regressions. This applies to all three models and is shown in Table 9. The size variable (*SIZE*) is significant at the one percent level for four years. The return on asset variable (*ROA*) always has a significant positive impact. The growth of assets variable (*ASG*) always has a negative and significant sign, and the market-to-book ratio (*MTB*) is significant with the expected negative sign for each year from 2003 to 2006.[49] The ratio of retained earnings to total equity (*RETE*) proves to be highly significant and positively related to the probability of paying dividend in all four periods.

[49] Years refer to the year that data are used. For example, year 2003 means that the data from 2003 is used.

This implies that firms with a relatively high ratio of earned to total equity are more likely to pay dividends. In Table 9, the ratio of cash to total assets (*CTA*) is significantly negatively related to the probability of paying dividends for all four years. This suggests that larger cash holdings are more likely to be retained for funding new projects.

Table 9 also provides evidence on the relation between dividends per share (*DPS*) and the peer variable. The spatial lag variable *SPLAG2* is highly significant with the expected positive sign for four years. However, the magnitude of the estimated coefficient associated with the variable *SPLAG2* varies considerably from year to year. For example, the coefficient estimates are 0.82 and 0.22 for the years 2003 and 2004, respectively. This can be interpreted to mean that a firm would adjust its dividend per share by 0.82 % in year 2004 when the weighted average of peers' dividends per share is increased by one percent in 2003. This result suggests that the weighted average of peers' dividend per share last year has a strong impact on the amount of dividend per share that a firm pays this year.

The alternative spatial lag variables (*SPLAG3* and *SPLAG4*) also have the predicted positive signs. But *SPLAG4* is statistically significant only in three years, 2003, 2005 and 2006. It is also clear from the Heckman second-stage regressions that the signs of the other control variables, such as *SIZE*, *ASG* and *EPS*, conform to expectations.

There is also strong evidence in support of the peer effects hypothesis in the sample drawn from the S&P 500 index (Table 10). I consistently observe a positive and highly significant relation between the firm's dividends per share and peers' dividend payouts. The coefficients on *SPLAG2, SPLAG3* and *SPLAG4* are of the predicted sign and highly significant in every model for every year. The second-stage OLS regressions also consistently reveal statistically significant relations between the firms' dividend per share and its size, growth and profitability, which is consistent with Fama and French (2001).

The results for Models 1 to 3 of Table 10 are fully consistent with the findings of Fama and French (2001). They confirm that the probability that a firm pays dividends is significantly and positively related to size, profitability, and negatively related to growth. As predicted, the life-cycle variable (*RETE*) is significantly positively related to the probability of paying dividends. The expected sign of the alternative life-cycle variable *CTA* is also positive. However, opposite to what is predicted, Table 10 reveals that cash holdings (*CTA*) are significantly negatively related to the probability of paying dividends in all models. This suggests that the *CTA* life-cycle variable is empirically distinct from the life-cycle variable *RETE*.

Table 9: Cross-Sectional Results on Heckman Two-Stage Selection Models for Samples of S&P 1500 Companies

Year	*Model*	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*SPLAG2*	*SPLAG3*	*SPLAG4*	*EPS*
2003	M1	First	-2.00***	0.00**	0.26***	1.94***	-0.84***	-0.04***	0.30***	-1.25***				
	M2		(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)				
	M3													
	M1	Second	-0.98***		0.10***	0.01	-0.77***	0.00			0.82***			0.15***
			(0.01)		(0.00)	(0.99)	(0.00)	(0.94)			(0.00)			(0.00)
	M2	Second	-0.76*		0.10**	0.01	-0.87***	0.00				0.49***		0.14***
			(0.06)		(0.01)	(0.98)	(0.00)	(0.73)				(0.00)		(0.00)
	M3	Second	-0.68*		0.10**	-0.07	-0.91***	0.00					0.45***	0.14***
			(0.09)		(0.01)	(0.90)	(0.00)	(0.68)					(0.00)	(0.00)
2004	M1	First	-1.99***	0.00**	0.25***	1.47***	-0.57***	-0.02***	0.41***	-1.12***				
	M2		(0.00)	(0.05)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)				
	M3													
	M1	Second	-0.20		0.07***	-0.90***	-0.36***	0.00***			0.22***			0.13***
			(0.54)		(0.01)	(0.01)	(0.01)	(0.00)			(0.01)			(0.00)
	M2	Second	-0.12		0.07**	-0.91***	-0.37***	0.00***				0.10*		0.13***
			(0.71)		(0.02)	(0.01)	(0.01)	(0.00)				(0.06)		(0.00)
	M3	Second	-0.07		0.07**	-0.97***	-0.37***	0.00***					0.05	0.13***
			(0.83)		(0.02)	(0.01)	(0.01)	(0.00)					(0.19)	(0.00)
2005	M1	First	-2.15***	0.00	0.27***	1.83***	-0.61***	-0.06***	0.62***	-1.19***				
	M2		(0.00)	(0.18)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)				
	M3													
	M1	Second	-0.12		0.05	-0.43	-0.41**	0.01			0.57***			0.09***
			(0.78)		(0.24)	(0.50)	(0.03)	(0.24)			(0.00)			(0.00)
	M2	Second	-0.06		0.04	-0.70	-0.46***	0.02**				0.33***		0.09***
			(0.89)		(0.26)	(0.27)	(0.01)	(0.02)				(0.00)		(0.00)
	M3	Second	-0.12		0.04	-0.76	-0.46***	0.02**					0.27***	0.09***
			(0.78)		(0.28)	(0.23)	(0.01)	(0.02)					(0.00)	(0.00)
2006	M1	First	-1.59***	0.00**	0.21***	1.36***	-0.85***	-0.05***	0.55***	-1.66***				
	M2		(0.00)	(0.03)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)				
	M3													
	M1	Second	-0.56		0.10	-0.53	-0.68*	0.00			0.38**			0.10**
			(0.47)		(0.15)	(0.59)	(0.06)	(0.77)			(0.02)			(0.02)
	M2	Second	-0.56		0.11	-0.47	-0.69**	0.00				0.31***		0.10**
			(0.47)		(0.14)	(0.64)	(0.05)	(0.53)				(0.00)		(0.02)
	M3	Second	-0.50		0.11	-0.53	-0.69**	0.00					0.26***	0.10**
			(0.52)		(0.14)	(0.60)	(0.05)	(0.53)					(0.01)	(0.02)

Notes: The table reports Heckman two-stage selection regressions. The year refers to the year that data are used. Dependent variables are leaded one year. For example, year 2003 data is used to explain the dividend behavior for year 2004.The dependent variable in the first-stage is *DIVPY*, which is coded one for firms/years with positive values of dividends per share (as defined in Table 5) and zero otherwise. The dependent variable in the second-stage is *DPS* (Compustat item 26, dividends per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables for the first-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *RETE* (6) *CTA* (7) *E*; The independent variables for the second-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *SPLAG2/ SPLAG3/ SPLAG4* (varies with the model) (6) *EPS*. Independent variable definitions are given in Table 5. The number of observations for the first-stage logit regression is 1010, 1041, 1066 and 1019 for the years 2003, 2004, 2005, and 2006, respectively. The number of observations for the second-stage OLS regressions is 466, 514, 540 and 521 for the years 2003, 2004, 2005, and 2006, respectively. P-values are in parentheses.* indicates significance at the 10% level; ** indicates significance at the 5% level;*** indicates significance at the 1% level.

Table 10 : Cross-Sectional Results on Heckman Two-Stage Selection Models for Samples of S&P 500 Companies

Year	*Model*	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*SPLAG2*	*SPLAG3*	*SPLAG4*	*EPS*
2003	M1 M2	First	-2.02*	0.00	0.24**	6.23***	-0.64	-0.06**	0.14**	-2.80***				
	M3		(0.08)	(0.17)	(0.04)	(0.00)	(0.12)	(0.02)	(0.02)	(0.00)				
	M1	Second	-1.23**		0.13***	0.19	-0.74***	0.00			0.70***			0.14***
			(0.02)		(0.00)	(0.81)	(0.00)	(0.94)			(0.00)			(0.00)
	M2	Second	-0.70		0.10**	-0.05	-0.78***	0.00				0.51***		0.14***
			(0.17)		(0.02)	(0.95)	(0.00)	(0.80)				(0.00)		(0.00)
	M3	Second	-0.66		0.10**	-0.25	-0.79***	0.00					0.49***	0.13***
			(0.19)		(0.02)	(0.75)	(0.00)	(0.60)					(0.00)	(0.00)
2004	M1	First	-3.17***	0.00	0.37***	4.70***	-0.26	-0.01**	0.21***	-2.20***				
	M2		(0.01)	(0.39)	(0.00)	(0.00)	(0.44)	(0.02)	(0.00)	(0.00)				
	M3													
	M1	Second	-1.36***		0.16***	0.13	-0.46***	0.00*			0.46***			0.15***
			(0.01)		(0.00)	(0.83)	(0.01)	(0.06)			(0.00)			(0.00)
	M2	Second	-0.93*		0.14***	-0.18	-0.36**	0.00***				0.24***		0.14***
			(0.08)		(0.00)	(0.76)	(0.05)	(0.00)				(0.00)		(0.00)
	M3	Second	-0.90*		0.14***	-0.25	-0.33*	0.00***					0.22***	0.14***
			(0.09)		(0.00)	(0.67)	(0.07)	(0.00)					(0.00)	(0.00)
2005	M1	First	-3.64	0.00	0.44***	3.05**	-0.43*	-0.06***	0.74***	-2.00***				
	M2			(0.81)	(0.00)	(0.04)	(0.08)	(0.00)	(0.00)	(0.00)				
	M3													
	M1	Second	-1.06**		0.11**	1.00	-0.38**	0.00			0.45***			0.12***
			(0.05)		(0.02)	(0.16)	(0.03)	(0.60)			(0.00)			(0.00)
	M2	Second	-0.71		0.10**	0.61	-0.46***	0.01*				0.27***		0.12***
			(0.18)		(0.04)	(0.40)	(0.01)	(0.06)				(0.00)		(0.00)
	M3	Second	-0.74		0.10**	0.51	-0.44***	0.02**					0.25***	0.12***
			(0.16)		(0.03)	(0.47)	(0.01)	(0.05)					(0.00)	(0.00)
2006	M1	First	-0.78	0.00*	0.14	1.13	-0.41*	-0.03**	0.61***	-3.44***				
	M2		(0.63)	(0.09)	(0.41)	(0.42)	(0.08)	(0.02)	(0.00)	(0.00)				
	M3													
	M1	Second	-0.67		0.11	-0.47	-0.49**	0.00			0.27**			0.12**
			(0.35)		(0.11)	(0.59)	(0.03)	(0.44)			(0.02)			(0.00)
	M2	Second	-0.51		0.10	-0.57	-0.48**	0.01				0.21**		0.12***
			(0.48)		(0.14)	(0.51)	(0.03)	(0.20)				(0.05)		(0.00)
	M3	Second	-0.54		0.10	-0.59	-0.48**	0.01					0.23**	0.11***
			(0.45)		(0.12)	(0.50)	(0.03)	(0.19)					(0.02)	(0.00)

Notes: The table reports Heckman two-stage selection regressions. The year refers to the year that data are used. Dependent variables are leaded one year. For example, year 2003 data is used to explain the dividend behavior for year 2004.The dependent variable in the first-stage is *DIVPY*, which is coded one for firms/years with positive values of dividends per share (as defined in Table 5) and zero otherwise. The dependent variable in the second-stage is *DPS* (Compustat item 26, dividends per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables for the first-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *RETE* (6) *CTA* (7) *E*; The independent variables for the second-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *SPLAG2/ SPLAG3/ SPLAG4* (varies with the model) (6) *EPS*. Independent variable definitions are given in Table 5. The number of observations for the first-stage logit regression is 312, 332, 334 and 318 for the years2003, 2004, 2005, and 2006, respectively. The number of observations for the second-stage OLS regressions is 212, 234, 239 and 228 for the years 2003, 2004, 2005, and 2006, respectively. P-values are in parentheses.* indicates significance at the 10% level; ** indicates significance at the 5% level;*** indicates significance at the 1% level.

5.3 Sensitivity Analysis

5.3.1 Do Peer Effects Exist when the Lagged Dependent Variable is Added in the Models?

My central prediction is that the amount of dividends a firm pays increases with the weighted average of peers' dividends as measured by the spatial lag variables *SPLAG2* or *SPLAG3* or *SPLAG4*. As indicated in Tables 9 and 10, the coefficients of the spatial lag variables are consistently positive, for both the S&P 1500 samples and the S&P 500 samples.

Lagged dividend per share is not included in the second-stage regressions in Tables 9 and 10. It is well known that Lintner's (1956) model explains dividend policy fairly well (Fama and Babiak, 1968). Lintner (1956) finds that his model explains 85% of the dividend changes from year to year. The empirical work, including Benartzi et al. (1997), DeAngelo et al. (1992), confirm Lintner's result that past dividend payouts and earnings are the most important factors influencing year-to-year dividend decisions. Benartzi et al. (1997) conclude that "Lintner's model of dividends remains the best description of the dividend setting process available."

Survey research by Baker et al. (1985), Baker and Powell (2000) and Farrelly et al. (1986) also support that earnings and past dividends are the key determinants affecting the dividend decision. The key question is whether the peer effects variables are still significantly and positive if a past dividends variable is added in the models.

I run the Heckman selection models for the S&P 1500 samples with lagged dividend per share added as an additional explanatory variable. The cross-sectional estimation results are presented in Table 11. In general, significant peer effects can be identified only for the two most recent years (2005 and 2006).[50] None of the estimated coefficients for the peer variables is significant when data of 2003 is used. Because of dividend tax changes in 2003, it would take firms a while to adjust to this policy change instead of following peers.

Table 11 shows that the lagged dividend variable is highly significant for all models across the four year periods. This is fully consistent with the previously literature. However, when the data from years 2005 and 2006 are used, the lagged dividend per share and peer effects variables are both entered into the second-stage OLS models, the lagged dividend variable seems to capture the full effect of the other explanatory variables. The four traditional explanatory variables size, profitability, growth and market-to-book ratio all become insignificant. This suggests that using lagged dividends as an explanatory variable might be problematic for attempts to identify the impact of the driving forces of dividends other than persistence.

Previously identified peer effects no longer appear to exist when lagged dividend per share is added to the models using data from the years 2003 and 2004. The results suggest that firms stick toward their past dividends for years 2004 and 2005.[51] One potential reason for the fact that the very strong persistence of dividends over time dominates the peer effects at some of the times might be related to changes in the

---

[50] Years refer to the year that data are used. For example, year 2005 means that the data from 2005 is used.
[51] The data from previous years is used to explain the dividend behavior in the current year.

dividend tax law. As there is uncertainty of how to react to the new laws, firms have followed their previous dividend polices. However, peer effects have become significant again for the data years 2005 and 2006 after firms have adjusted to the new laws and have had time to see what others are doing.

Table 11: Sensitivity Analysis, Cross-Sectional Results for a Sample of S&P 1500 with Lagged Dividends per Share added as an Explanatory Variable

Year	*Model*	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*SPLAG2*	*SPLAG3*	*SPLAG4*	*EPS*	*DPS*
2003	M1	First	-2.00***	0.00**	0.26***	1.94***	-0.84***	-0.04***	0.30***	-1.25***					
	M2		(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)					
	M3														
	M1	Second	-0.62**		0.04*	0.73*	-0.28*	-0.01			0.06			0.04***	1.04***
			(0.04)		(0.10)	(0.07)	(0.08)	(0.27)			(0.53)			(0.00)	(0.00)
	M2	Second	-0.59**		0.04*	0.73*	-0.29*	-0.01				0.01		0.03***	1.05***
			(0.05)		(0.10)	(0.07)	(0.07)	(0.27)				(0.85)		(0.00)	(0.00)
	M3	Second	-0.61**		0.04*	0.73*	-0.29*	-0.01					0.04	0.04***	1.04***
			(0.04)		(0.10)	(0.07)	(0.07)	(0.28)					(0.44)	(0.00)	(0.00)
2004	M1	First	-1.99***	0.00**	0.25***	1.47***	-0.57***	-0.02***	0.41***	-1.12***					
	M2		(0.00)	(0.05)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)					
	M3														
	M1	Second	-0.31		0.06**	-0.45	-0.14	0.00**			0.12*			0.09***	0.36***
			(0.30)		(0.02)	(0.18)	(0.27)	(0.03)			(0.09)			(0.00)	(0.00)
	M2	Second	-0.26		0.06**	-0.46	-0.14	0.00***				0.05		0.09***	0.36***
			(0.38)		(0.02)	(0.17)	(0.27)	(0.01)				(0.35)		(0.00)	(0.00)
	M3	Second	-0.23		0.06**	-0.49	-0.14	0.00***					0.01	0.09***	0.36***
			(0.43)		(0.02)	(0.14)	(0.26)	(0.01)					(0.70)	(0.00)	(0.00)
2005	M1	First	-2.15***	0.00	0.27***	1.83***	-0.61***	-0.06***	0.62***	-1.19***					
	M2		(0.00)	(0.18)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)					
	M3														
	M1	Second	-0.05		0.01	-0.12	-0.07	0.00			0.38***			0.04***	0.62***
			(0.91)		(0.77)	(0.84)	(0.70)	(0.68)			(0.00)			(0.01)	(0.00)
	M2	Second	-0.08		0.01	-0.29	-0.09	0.01				0.21***		0.04***	0.64***
			(0.84)		(0.82)	(0.62)	(0.60)	(0.23)				(0.02)		(0.01)	(0.00)
	M3	Second	-0.12		0.01	-0.33	-0.09	0.01					0.17**	0.04***	0.64***
			(0.76)		(0.84)	(0.58)	(0.60)	(0.23)					(0.04)	(0.01)	(0.00)
2006	M1	First	-1.59***	0.00**	0.21***	1.36***	-0.85***	-0.05***	0.55***	-1.66***					
	M2		(0.00)	(0.03)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)					
	M3														
	M1	Second	-0.42		0.09	-0.57	-0.51	0.00			0.29*			0.07	0.24***
			(0.58)		(0.23)	(0.56)	(0.16)	(0.80)			(0.08)			(0.11)	(0.00)
	M2	Second	-0.44		0.09	-0.49	-0.52	0.00				0.28***		0.07*	0.24***
			(0.57)		(0.22)	(0.62)	(0.15)	(0.60)				(0.01)		(0.10)	(0.00)
	M3	Second	-0.38		0.09	-0.55	-0.52	0.00					0.23**	0.07	0.25***
			(0.62)		(0.22)	(0.58)	(0.15)	(0.61)					(0.02)	(0.11)	(0.00)

Notes: The Table reports Heckman two-stage selection regressions. The year refers to the year that data are used. Dependent variables are leaded one year. For example, year 2003 data is used to explain the dividend behavior for year 2004.The dependent variable in the first-stage is *DIVPY*, which is coded one for firms/years with positive values of dividends per share (as defined in Table 5) and zero otherwise. The dependent variable in the second-stage is *DPS* (Compustat item 26, dividends per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables for the first-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *RETE* (6) *CTA* (7) *E*; The independent variables for the second-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *SPLAG2/ SPLAG3/ SPLAG4* (varies with the model) (6) *EPS*. Independent variable definitions are given in Table 5. The number of observations for the first-stage logit regression is 1010, 1041, 1066 and 1019 for year 2003, 2004, 2005 and 2006 respectively. The number of observations for the second-stage OLS regression is 466, 514, 540 and 521 for year 2003, 2004, 2005 and 2006 respectively. P-values are in parentheses.* indicates significance at the 10% level; ** indicates significance at the 5% level;*** indicates significance at the 1% level.

5.3.2 Peer Effects and the Probability of a Firm Paying Dividends

The events of firms initiating or terminating dividends have been viewed as having more serious consequences than just changing the amount of the dividend (Baker and Wurgler, 2004; Lie, 2005). In this study, the analysis of peer effects is limited to dividend paying firms that also paid dividends in the previous year. However, a natural question is whether peer effects also play a role for the first-stage regression, which explains the probability of a firm paying dividends.

Since the first stage comprises far more companies than the second stage, the weight matrix for the second stage cannot be used for the first stage. As an alternative, one can construct a peer variable that contains the percentage of companies in the same 2-digit, 3-digit or 4-digit SIC group that pay dividends. This peer variable is added as an additional regressor to the first-stage regression.

For example, suppose there are a total of 5 companies, including dividend payers and non-payers. Companies 1, 2, and 4 are in the same 3-digit SIC group. If companies 1 and 2 are paying dividends for the year 2003 but company 4 not, then the first, second and fourth entries in the peer variable vector are 2/3, 2/3, 2/3 for that 3-digit SIC group for the year 2003.[52] If there is just one firm in the SIC group, then the percentage of others paying dividends is set to zero.[53] To avoid the problem of simultaneity, the

[52] Because the total number of companies in the same 3-digit SIC group is 3 and the total number of companies paying dividends in this 3-digit SIC group is 2, the percentage of companies that are in the same 3-digit SIC group paying dividends is 2/3.

[53] Though there is alternative to substitute the percentage of companies in the same 2-digit group that pay dividends. It seems that one can view it as the strength of a signal from nearby firms on whether to pay a

previous year's value is used for the percentage of companies paying dividends and only non-missing values are counted for each SIC group.

The results from cross-section regressions using the S&P 1500 sample are reported in Table 12. The peer variable entered in the first stage (*PV2, PV 3 or PV4*)[54] is positive and statistically significant for four years and for all three models. This indicates that peer effects matter for the decision whether to pay or not to pay dividends. The magnitudes of the peer variables reveal that the 2-digit peer results are stronger than those for the 3-digit SICs or for the 4-digit SICs. Previously identified strong peer effects at the second stage can also be observed.

---

dividend. When there are no nearby firms the strength is zero, just as it would be when there are nearby firms, but none pay dividends.

[54] The peer variable *PV2* is constructed on the percentage of firms in the same 2-digit SIC group that pay dividends. Similarly, the peer variable *PV3* is constructed on the percentage of firms in the same 3-digit SIC group that pay dividends. The same definition applies to the peer variable *PV4*.

Table 12: Sensitivity Analysis, Cross-Sectional Results for a Sample of S&P 1500 with Peer Variable added in the First-Stage Regression

Year	*Model*	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*PV2*	*SPLAG2*	*PV3*	*SPLAG3*	*PV4*	*SPLAG4*	*EPS*
2003	M1	First	-2.56***	0.00**	0.26***	2.03***	-0.80***	-0.04***	0.28***	-1.01***	1.06***						
			(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)						
		Second	-0.59*		0.07**	-0.28	-0.68***	0.00				0.85***					0.14***
			(0.08)		(0.02)	(0.57)	(0.00)	(0.64)				(0.00)					(0.00)
	M2	First	-2.41***	0.00**	0.27***	2.01***	-0.82***	-0.04***	0.28***	-0.97***			0.77***				
			(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)			(0.00)				
		Second	-0.27		0.06**	-0.31	-0.75***	0.01						0.49***			0.14***
			(0.40)		(0.05)	(0.54)	(0.00)	(0.39)						(0.00)			(0.00)
	M3	First	-2.22***	0.00**	0.26***	1.96***	-0.84***	-0.04***	0.29***	-1.06***					0.50***		
			(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)					(0.00)		
		Second	-0.44		0.08**	-0.25	-0.84***	0.01								0.47***	0.14***
			(0.21)		(0.02)	(0.62)	(0.00)	(0.49)								(0.00)	(0.00)
2004	M1	First	-2.59***	0.00*	0.25***	1.64***	-0.54***	-0.02***	0.40***	-0.81***	1.08***						
			(0.00)	(0.06)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)						
		Second	-0.42		0.09***	-0.83**	-0.41***	0.00***				0.24***					0.14***
			(0.14)		(0.00)	(0.02)	(0.00)	(0.00)				(0.00)					(0.00)
	M2	First	-2.36***	0.00**	0.26***	1.54***	-0.54***	-0.02***	0.39***	-0.84***			0.66***				
			(0.00)	(0.05)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)			(0.00)				
		Second	-0.28		0.08***	-0.84**	-0.40***	0.00***						0.12**			0.13***
			(0.33)		(0.00)	(0.02)	(0.00)	(0.00)						(0.04)			(0.00)
	M3	First	-2.23***	0.00**	0.25***	1.53***	-0.57***	-0.02***	0.40***	-0.90***					0.48***		
			(0.00)	(0.05)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)					(0.00)		
		Second	-0.16		0.08***	-0.93***	-0.39***	0.00***								0.06	0.13***
			(0.59)		(0.00)	(0.01)	(0.00)	(0.00)								(0.16)	(0.00)
2005	M1	First	-2.71***	0.00	0.27***	1.91***	-0.60***	-0.06***	0.60***	-0.93***	1.01***						
			(0.00)	(0.19)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)						
		Second	0.08		0.03	-0.52	-0.37**	0.01				0.54***					0.08***
			(0.85)		(0.41)	(0.41)	(0.04)	(0.19)				(0.00)					(0.00)
	M2	First	-2.54***	0.00	0.27***	1.81***	-0.63***	-0.06***	0.57***	-0.94***			0.71***				
			(0.00)	(0.16)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)			(0.00)				
		Second	0.18		0.03	-0.76	-0.43**	0.02**						0.32***			0.08***
			(0.66)		(0.35)	(0.23)	(0.02)	(0.02)						(0.00)			(0.00)
	M3	First	-2.41***	0.00	0.27***	1.78***	-0.64***	-0.06***	0.59***	-0.98***					0.55***		
			(0.00)	(0.16)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)					(0.00)		
		Second	0.17		0.04	-0.79	-0.45**	0.02**								0.26***	
			(0.68)		(0.30)	(0.21)	(0.01)	(0.02)								(0.00)	
2006	M1	First	-2.12***	0.00**	0.19***	1.43**	-0.74***	-0.05***	0.56***	-1.34***	1.03***						
			(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)						
		Second	-0.37		0.09	-0.63	-0.61*	0.00				0.41**					0.10**
			(0.61)		(0.19)	(0.52)	(0.07)	(0.76)				(0.02)					(0.03)

Table 12: Sensitivity Analysis, Cross-Sectional Results for a Sample of S&P 1500 with Peer Variable added in the First-Stage Regression

Year	*Model*	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*PV2*	*SPLAG2*	*PV3*	*SPLAG3*	*PV4*	*SPLAG4*	*EPS*
	M2	First	-1.97***	0.00**	0.20***	1.28**	-0.74***	-0.04***	0.55***	-1.32***			0.79***				
			(0.00)	(0.02)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)			(0.00)				
		Second	-0.32		0.09	-0.56	-0.61*	0.00						0.33***			0.10**
			(0.64)		(0.19)	(0.56)	(0.07)	(0.50)						(0.00)			(0.02)
	M3	First	-1.82***		0.19***	1.26**	-0.75***	-0.04***	0.56***	-1.39***					0.58***		
			(0.00)		(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)					(0.00)		
		Second	-0.24		0.08	-0.62	-0.61*	0.00								0.28***	0.10**
			(0.73)		(0.20)	(0.53)	(0.07)	(0.50)								(0.01)	(0.03)

Notes: The Table reports Heckman two-stage selection regressions. The year refers to the year that the regressors are defined for. The dependent variables lead by one year. For example, year 2003 regressors are used to explain the dividend behavior for year 2004.The dependent variable in the first-stage is *DIVPY*, which is coded one for firms/years with positive values of dividends per share (as defined in Table 5) and zero otherwise. The dependent variable in the second-stage is *DPS* (Compustat item 26, dividends per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables for the first-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *RETE* (6) *CTA* (7) *E (8) PV2/PV3/PV4;* The independent variables for the second-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *SPLAG2/ SPLAG3/ SPLAG4 (*varies with the model) (6) *EPS*. Independent variable definitions are given in Table 5. The number of observations for the first-stage logit regression is 1012, 1044, 1074 and 1028 for year 2003, 2004, 2005 and 2006 respectively. The number of observations for the second-stage OLS regression is 466, 514, 540 and 521 for year 2003, 2004, 2005 and 2006 respectively. P-values are in parentheses.* indicates significance at the 10% level; ** indicates significance at the 5% level;*** indicates significance at the 1% level.

5.3.3 Strength of Peer Effects when Peers Exist at Multiple SIC Levels

According to Table 9, the 2-digit peer effect results (*SPLAG2*) are stronger than those for the 3-digit industries (*SPLAG3*) or those for the 4-digit industries (*SPLAG4*). It may be of interest to test whether the coefficients of the peer variables are larger or more significant when peers at the two digit level coexist with peers also at the three or four digit SIC levels than those where peers are only at the 2-digit levels.

Following the first step as before to construct the peer variable (*SPLAG*), I need to calculate the weight matrix. I add the three raw matrices for 2-digit, 3-digit, and 4-digit SICs, but with zeros on the main diagonal. Then I standardize the resulting matrix and use the standardized matrix as the weight matrix. The peer variable is defined as the product of this weight matrix and the dividend per share for a specified year.

The results from the S&P 1500 sample are shown in Table 13. Significant peer effects can be identified for every year from 2003 to 2006. The sizes of the coefficients of *SPLAG* for the years 2003 and 2006 are 0.83 and 0.42, respectively, which are slightly larger than for those peer variables that are based on the 2-digit SICs (*SPLAG2*). However, for the years 2004 and 2005, the peer variable (*SPLAG*) has smaller coefficients compared to the corresponding variable *SPLAG2*. Therefore, one cannot conclude that matching SIC codes at the 3-digit or 4-digit level in addition to matches at the 2-digit level produces larger peer effects than matching SIC codes only at the 2-digit level.

Table 13: Sensitivity Analysis, Cross-Sectional Results for a Sample of S&P 1500 with SPLAG Constructed on the Combined Two, Three or Four digit SICs

Year	*Model*	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*SPLAG2*	*SPLAG3*	*SPLAG4*	*SPLAG*	*EPS*
2003	M1	First	-2.00***	0.00**	0.26***	1.94***	-0.84***	-0.04***	0.30***	-1.25***					
	M2		(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)					
	M3														
	M4														
	M1	Second	-0.98***		0.10***	0.01	-0.77***	0.00			0.82***				0.15***
			(0.01)		(0.00)	(0.99)	(0.00)	(0.94)			(0.00)				(0.00)
	M2	Second	-0.76*		0.10**	0.01	-0.87***	0.00				0.49***			0.14***
			(0.06)		(0.01)	(0.98)	(0.00)	(0.73)				(0.00)			(0.00)
	M3	Second	-0.68*		0.10**	-0.07	-0.91***	0.00					0.45***		0.14***
			(0.09)		(0.01)	(0.90)	(0.00)	(0.68)					(0.00)		(0.00)
	M4	Second	-1.03***		0.11***	0.06	-0.76***	0.00						0.83***	0.15***
			(0.01)		(0.00)	(0.91)	(0.00)	(0.92)						(0.00)	(0.00)
2004	M1	First	-1.99***	0.00**	0.25***	1.47***	-0.57***	-0.02***	0.41***	-1.12***					
	M2		(0.00)	(0.05)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)					
	M3														
	M4														
	M1	Second	-0.20		0.07***	-0.90***	-0.36***	0.00***			0.22***				0.13***
			(0.54)		(0.01)	(0.01)	(0.01)	(0.00)			(0.01)				(0.00)
	M2	Second	-0.12		0.07**	-0.91***	-0.37***	0.00***				0.10*			0.13***
			(0.71)		(0.02)	(0.01)	(0.01)	(0.00)				(0.06)			(0.00)
	M3	Second	-0.07		0.07**	-0.97***	-0.37***	0.00***					0.05		0.13***
			(0.83)		(0.02)	(0.01)	(0.01)	(0.00)					(0.19)		(0.00)
	M4	Second	-0.19		0.07***	-0.88**	-0.36***	0.00***						0.19***	0.13***
			(0.56)		(0.01)	(0.02)	(0.01)	(0.00)						(0.01)	(0.00)
2005	M1	First	-2.15***	0.00	0.27***	1.83***	-0.61***	-0.06***	0.62***	-1.19***					
	M2		(0.00)	(0.18)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)					
	M3														
	M4														
	M1	Second	-0.12		0.05	-0.43	-0.41**	0.01			0.57***				0.09***
			(0.78)		(0.24)	(0.50)	(0.03)	(0.24)			(0.00)				(0.00)
	M2	Second	-0.06		0.04	-0.70	-0.46***	0.02**				0.33***			0.09***
			(0.89)		(0.26)	(0.27)	(0.01)	(0.02)				(0.00)			(0.00)
	M3	Second	-0.12		0.04	-0.76	-0.46***	0.02**					0.27***		0.09***
			(0.78)		(0.28)	(0.23)	(0.01)	(0.02)					(0.00)		(0.00)
	M4	Second	-0.10		0.05	-0.44	-0.42**	0.01						0.51***	0.08***
			(0.82)		(0.23)	(0.50)	(0.02)	(0.21)						(0.00)	(0.00)
2006	M1	First	-1.59***	0.00**	0.21***	1.36***	-0.85***	-0.05***	0.55***	-1.66***					
	M2		(0.00)	(0.03)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)					
	M3														
	M4														
	M1	Second	-0.56		0.10	-0.53	-0.68*	0.00			0.38**				0.10**
			(0.47)		(0.15)	(0.59)	(0.06)	(0.77)			(0.02)				(0.02)
	M2	Second	-0.56		0.11	-0.47	-0.69**	0.00				0.31***			0.10**
			(0.47)		(0.14)	(0.64)	(0.05)	(0.53)				(0.00)			(0.02)
	M3	Second	-0.50		0.11	-0.53	-0.69**	0.00					0.26***		0.10**
			(0.52)		(0.14)	(0.60)	(0.05)	(0.53)					(0.01)		(0.02)

Table 13 : Sensitivity Analysis, Cross-Sectional Results for a Sample of S&P 1500 with SPLAG Constructed on the Combined Two, Three or Four digit SICs

Year	*Model*	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*SPLAG2*	*SPLAG3*	*SPLAG4*	*SPLAG*	*EPS*
	M4	Second	-0.55		0.10	-0.46	-0.66*	0.00						0.42***	0.10**
			(0.48)		(0.17)	(0.64)	(0.07)	(0.80)						(0.01)	(0.03)

Notes: The Table reports Heckman two-stage selection regressions. The year refers to the year that the regressors are defined for. The dependent variables lead by one year. For example, year 2003 regressors are used to explain the dividend behavior for year 2004.The dependent variable in the first-stage is *DIVPY*, which is coded one for firms/years with positive values of dividends per share (as defined in Table 5) and zero otherwise. The dependent variable in the second-stage is *DPS* (Compustat item 26, dividends per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables for the first-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *RETE* (6) *CTA* (7) *E* ; The independent variables for the second-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *SPLAG2/ SPLAG3/ SPLAG4/ SPLAG* (varies with the model) (6) *EPS*. Independent variable definitions are given in Table 5. The number of observations for the first-stage logit regression is 1012, 1044, 1074 and 1028 for year 2003, 2004, 2005 and 2006 respectively. The number of observations for the second-stage OLS regression is 466, 514, 540 and 521 for year 2003, 2004, 2005 and 2006 respectively. P-values are in parentheses.* indicates significance at the 10% level; ** indicates significance at the 5% level;*** indicates significance at the 1% level.

As indicated in Tables 12 and 13, peers play a role at the first stage and the new combined peer variable (*SPLAG*) is significant. Table 14 presents the results from adding the peer variables at the first stage and the combined peer variable at the second stage. The size of the *SPLAG* at the second stage is not very sensitive in terms of different peer variables entered at the first stage (*PV2, PV3 or PV4*).

Table 14: Sensitivity Analysis, Cross-Sectional Results for a Sample of S&P 1500 with Peer Variable added in the First-Stage Regression and Combined SPLAG in the Second-Stage Regression

Year	*Model*	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*PV2*	*SPLAG*	*PV3*	*SPLAG*	*PV4*	*SPLAG*	*EPS*
2003	M1	First	-2.56***	0.00**	0.26***	2.03***	-0.80***	-0.04***	0.28***	-1.01***	1.06***						
			(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)						
		Second	-0.66**		0.07***	-0.22	-0.67***	0.00				0.87***					0.15***
			(0.05)		(0.01)	(0.66)	(0.00)	(0.63)				(0.00)					(0.00)
	M2	First	-2.41***	0.00**	0.27***	2.01***	-0.82***	-0.04***	0.28***	-0.97***			0.77***				
			(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)			(0.00)				
		Second	-0.43		0.06**	-0.33	-0.62***	0.01						0.82***			0.14***
			(0.18)		(0.05)	(0.49)	(0.00)	(0.48)						(0.00)			(0.00)
	M3	First	-2.22***	0.00**	0.26***	1.96***	-0.84***	-0.04***	0.29***	-1.06***					0.50***		
			(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)					(0.00)		
		Second	-0.60*		0.07**	-0.24	-0.66***	0.00								0.83***	0.15***
			(0.08)		(0.02)	(0.64)	(0.00)	(0.58)								(0.00)	(0.00)
2004	M1	First	-2.59***	0.00*	0.25***	1.64***	-0.54***	-0.02***	0.40***	-0.81***	1.08***						
			(0.00)	(0.06)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)						

Table 14 : Sensitivity Analysis, Cross-Sectional Results for a Sample of S&P 1500 with Peer Variable added in the First-Stage Regression and Combined SPLAG in the Second-Stage Regression

Year	*Model*	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*PV2*	*SPLAG*	*PV3*	*SPLAG*	*PV4*	*SPLAG*	*EPS*
		Second	-0.40		0.09***	-0.81**	-0.40***	0.00***				0.22***					0.14***
			(0.17)		(0.00)	(0.02)	(0.00)	(0.00)				(0.00)					(0.00)
	M2	First	-2.36***	0.00**	0.26***	1.54***	-0.54***	-0.02***	0.39***	-0.84***			0.66***				
			(0.00)	(0.05)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)			(0.00)				
		Second	-0.32		0.08***	-0.83**	-0.38***	0.00***						0.21***			0.13***
			(0.27)		(0.00)	(0.02)	(0.00)	(0.00)						(0.00)			(0.00)
	M3	First	-2.23***	0.00**	0.25***	1.53***	-0.57***	-0.02***	0.40***	-0.90***					0.48***		
			(0.00)	(0.05)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)					(0.00)		
		Second	-0.28		0.08***	-0.85**	-0.37***	0.00***								0.20***	0.13***
			(0.36)		(0.00)	(0.02)	(0.01)	(0.00)								(0.01)	(0.00)
2005	M1	First	-2.71***	0.00	0.27***	1.91***	-0.60***	-0.06***	0.60***	-0.93***	1.01***						
			(0.00)	(0.19)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)						
		Second	0.12		0.03	-0.53	-0.38**	0.01				0.49***					0.08***
			(0.77)		(0.42)	(0.40)	(0.04)	(0.15)				(0.00)					(0.00)
	M2	First	-2.54***	0.00	0.27***	1.81***	-0.63***	-0.06***	0.57***	-0.94***			0.71***				
			(0.00)	(0.16)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)			(0.00)				
		Second	0.11		0.03	-0.54	-0.38**	0.01						0.50***			0.08***
			(0.79)		(0.41)	(0.39)	(0.04)	(0.15)						(0.00)			(0.00)
	M3	First	-2.41***	0.00	0.27***	1.78***	-0.64***	-0.06***	0.59***	-0.98***					0.55***		
			(0.00)	(0.16)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)					(0.00)		
		Second	0.02		0.04	-0.50	-0.40**	0.01								0.50***	0.08***
			(0.95)		(0.32)	(0.43)	(0.03)	(0.17)								(0.00)	(0.00)
2006	M1	First	-2.12***	0.00**	0.19***	1.43**	-0.74***	-0.05***	0.56***	-1.34***	1.03***						
			(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)						
		Second	-0.40		0.09	-0.54	-0.60*	0.00				0.45***					0.09**
			(0.57)		(0.20)	(0.58)	(0.08)	(0.79)				(0.00)					(0.03)
	M2	First	-1.97***	0.00**	0.20***	1.28**	-0.74***	-0.04***	0.55***	-1.32***			0.79***				
			(0.00)	(0.02)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)			(0.00)				
		Second	-0.21		0.07	-0.59	-0.55	0.00						0.44***			0.09**
			(0.76)		(0.28)	(0.55)	(0.10)	(0.78)						(0.00)			(0.03)
	M3	First	-1.82***		0.19***	1.26**	-0.75***	-0.04***	0.56***	-1.39***					0.58***		
			(0.00)		(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)					(0.00)		
		Second	-0.22		0.07	-0.58	-0.56	0.00								0.43***	0.09**
			(0.75)		(0.28)	(0.56)	(0.10)	(0.78)								(0.00)	(0.03)

Notes: The Table reports Heckman two-stage selection regressions. The year refers to the year that the regressors are defined for. The dependent variables lead by one year. For example, year 2003 regressors are used to explain the dividend behavior for year 2004.The dependent variable in the first-stage is *DIVPY*, which is coded one for firms/years with positive values of dividends per share (as defined in Table 5) and zero otherwise. The dependent variable in the second-stage is *DPS* (Compustat item 26, dividends per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables for the first-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *RETE* (6) *CTA* (7) *E (8) PV2/PV3/PV4;* The independent variables for the second-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *SPLAG* (6) *EPS*. Independent variable definitions are given in Table 5. The number of observations for the first-stage logit regression is 1012, 1044, 1074 and 1028 for year 2003, 2004, 2005 and 2006 respectively. The number of observations for the second-stage OLS regression is 466, 514, 540 and 521 for year 2003, 2004, 2005 and 2006 respectively. P-values are in parentheses.* indicates significance at the 10% level; ** indicates significance at the 5% level;*** indicates significance at the 1% level.

Table 15 reports the estimation results from a model that adds the combined peer variables for both stages. The newly constructed combined peer variables (*PV* and *SPLAG*) are posivitive and statistically significant at the 1 % level for four years. However, the magnitude of the peer variable at the first stage is much larger than the one at the second stage.

Table 15: Cross-Sectional Results for a Sample of S&P 1500 with the Combined Peer Variables PV and SPLAG Added in the Both Stages

Year	*Stage*	*Int.*	*E*	*SIZE*	*ROA*	*ASG*	*MTB*	*RETE*	*CTA*	*PV*	*SPLAG*	*EPS*
2003	First	-2.84***	0.00**	0.26***	2.00***	-0.79***	-0.03**	0.27***	-0.77***	1.73***		
		(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.03)	(0.00)	(0.01)	(0.00)		
	Second	-0.36		0.05**	-0.38	-0.61***	0.01				0.83***	0.14***
		(0.21)		(0.05)	(0.43)	(0.00)	(0.43)				(0.00)	(0.00)
2004	First	-2.94***	0.00*	0.25***	1.74***	-0.51***	-0.02***	0.39***	-0.54*	1.81***		
		(0.00)	(0.06)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.06)	(0.00)		
	Second	-0.09		0.06***	-0.93***	-0.34***	0.00***				0.20***	0.13***
		(0.70)		(0.00)	(0.01)	(0.01)	(0.00)				(0.01)	(0.00)
2005	First	-3.05***	0.00	0.26***	1.90***	-0.59***	-0.06***	0.58***	-0.66**	1.76***		
		(0.00)	(0.15)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.00)		
	Second	0.13		0.03	-0.52	-0.38**	0.01				0.48***	0.08***
		(0.72)		(0.39)	(0.40)	(0.04)	(0.15)				(0.00)	(0.00)
2006	First	-2.39***	0.00**	0.18***	1.41**	-0.69***	-0.05***	0.57***	-1.03***	1.71***		
		(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
	Second	-0.22		0.07	-0.61	-0.56*	0.00				0.45***	0.09**
		(0.73)		(0.25)	(0.53)	(0.09)	(0.78)				(0.00)	(0.03)

Notes: The Table reports Heckman two-stage selection regressions. The year refers to the year that the regressors are defined for. The dependent variables lead by one year. For example, year 2003 regressors are used to explain the dividend behavior for year 2004.The dependent variable in the first-stage is *DIVPY*, which is coded one for firms/years with positive values of dividends per share (as defined in Table 5) and zero otherwise. The dependent variable in the second-stage is *DPS* (Compustat item 26, dividends per share), adjusted for stock splits and dividends (as defined in Table 5). The independent variables for the first-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *RETE* (6) *CTA* (7) *E (8) PV;* The independent variables for the second-stage equation are: (1) *SIZE* (2) *ROA* (3) *ASG* (4) *MTB* (5) *SPLAG* (6) *EPS*. Independent variable definitions are given in Table 5. The number of observations for the first-stage logit regression is 1012, 1044, 1074 and 1028 for year 2003, 2004, 2005 and 2006 respectively. The number of observations for the second-stage OLS regression is 466, 514, 540 and 521 for year 2003, 2004, 2005 and 2006 respectively. P-values are in parentheses.* indicates significance at the 10% level; ** indicates significance at the 5% level;*** indicates significance at the 1% level.

5.4 Panel Data Estimates of Peer Effects

In order to estimate the selection model (4) and the outcome model (7), one needs to calculate the time averages $\bar{z}_i$ and $\bar{x}_i$. To do this, I compute the sample average of each explanatory variable over the time period T for every firm i. For example, the average of the variable *ROA*, denoted *ROAbar*, is calculated as the sum of the *ROAs* over the sample years divided by the number of the years.[55] Next, instead of estimating the selection equation (4) with a panel probit estimation method, I estimate the probit models of the selection equation (4) for each time period in the sample. These are also called cross-sectional probit models. The purpose of separately estimating the probit selection model for each time period is to compute an inverse Mills ratio (*IMR*) $\lambda_{i,t}$ for each t for every firm i.[56] This mean that each firm i should have T IMRs.

Finally, I estimate the outcome dividend equation (7) using pooled OLS.[57] That means, no fixed effects or random effects estimator is used. What the fixed or random effects normally capture is instead absorbed through the inclusion of the sample averages. In addition, I add the year dummies to equation (7). It is important to note that only the observations with positive (non-missing) dividend per share are used in the second stage,

[55] I use the seven years data from 2000 to 2006 to form a panel, thus *ROAbar*= *(ROA2000+ROA2001+ROA2002+ROA2003+ROA2004+ROA2005+ROA2006)/7* and this value is used for every year for a firm i in the panel.

[56] For example, after probit selection model (8) is estimated using year 2005 data, the IMR for 2005 is derived as *IMR2005 = normalden(xb)/normal(xb)*, where *xb* is the linear prediction from the fitted model.

[57] Because the firm-specific average terms I include are basically proxies for firm-specific effects, the estimation strategy that has those firm-specific average terms is basically a fixed effects approach. Due to other issues created by allowing selection to change over time, the way I handle the fixed effects is slight different as the common ways.

while the entire sample is used in the first-stage selection models. In order to allow the error terms to vary for different firms, I adjust the standard errors using the cluster option.

5.5 Panel Estimation Results

Table 16 shows the panel estimation results for the outcome equation using the S&P 1500 sample from 2000 to 2006. Three models are presented. They differ in terms of the spatial lag variable that enters the estimation equation.

Statistically significant coefficients for the peer variable can only be identified for Model 1. The estimate of *SPLAG2* is 0.04 and statistically significant at the 1% level. It means that a firm would adjust its dividends per share by 0.04% if the weighted average of the peer firms' dividends per share is changed by 1%. Somewhat unexpectedly, the traditional explanatory variables of dividend payments that prove to be significant in the literature turn insignificant for all three models when a spatial lag variable is included in the model. However, the *IMRs*, time dummies and the average terms are all highly significant for three models. A possible explanation for the fact that these variables capture most of the explanatory power of the models is that there is not substantial variation in the data values for the same firm over time.

It is apparent that the panel estimation results from the pooled OLS models provide coefficient estimates for the peer variable that are much less significant than those from the cross-sections. One may conclude that the differences across firms are much more important than those within firms for the sample S&P data. Because the panel models are identified based on the variation within firms, it is possible that the results they produce differ from those of the cross-sectional models.

Table 16: Pooled OLS Estimation Results for the Outcome Equation on a Panel of Firms from the S&P1500 Index

	Model 1	Model 2	Model 3
SPLAG2	0.04*		
	(0.09)		
SPLAG3		0.01	
		(0.27)	
SPLAG4			0.01
			(0.32)
EPS	0.05	0.05	0.05
	(0.12)	(0.12)	(0.11)
SIZE	-0.08	-0.08	-0.08
	(0.55)	(0.54)	(0.52)
ROA	-0.30	-0.31	-0.33
	(0.54)	(0.53)	(0.51)
ASG	0.00*	0.00	0.00
	(0.09)	(0.11)	(0.12)
MTB	0.00	0.00	0.00
	(0.66)	(0.69)	(0.70)
Constant	0.02	0.21	0.29
	(0.93)	(0.25)	(0.12)
N	2294	2294	2294
R^2	0.20	0.18	0.17
Wald tests on the joint significance of			
7 IMRs	2.40**	2.66***	2.78***
	(0.02)	(0.01)	(0.01)
6 year dummies	4.72***	4.87***	5.26***
	(0.00)	(0.00)	(0.00)
Average terms	9.63***	5.61***	4.91***
	(0.00)	(0.00)	(0.00)

Notes: P-values are in parentheses.* indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level. The standard errors are adjusted by the cluster option.

5.6 Sensitivity Analysis for Panel Estimates

For the purpose of sensitivity analysis, I run the three models by adding interaction terms between the peer variable and year dummies. The results show that the interaction terms are not significant while all the year dummies are still highly significant. If the six year dummies are removed, the interaction terms become significant.

Using the combined peer variable (*SPLAG*), I run the panel estimates and the results are given in Table 17. The peer variable *SPLAG* is statistically significant in Model 4, when *SPLAG* is entered in the second-stage regression model. The model results are consistent with the previous panel results in that *IMRs*, time dummies and the average terms are all highly significant in Model 4, as in the other three models.

Table 17: Sensitivity Analysis, Pooled OLS Estimation Results for the Outcome Equation on a Panel of Firms from the S&P1500 Index

	Model 1	Model 2	Model 3	Model 4
SPLAG2	0.04*			
	(0.09)			
SPLAG3		0.01		
		(0.27)		
SPLAG4			0.01	
			(0.32)	
SPLAG				0.03*
				(0.09)
EPS	0.05	0.05	0.05	0.05
	(0.12)	(0.12)	(0.11)	(0.12)
SIZE	-0.08	-0.08	-0.08	-0.08
	(0.55)	(0.54)	(0.52)	(0.55)
ROA	-0.30	-0.31	-0.33	-0.30
	(0.54)	(0.53)	(0.51)	(0.55)
ASG	0.00*	0.00	0.00	0.00*
	(0.09)	(0.11)	(0.12)	(0.09)

Table 17 : Sensitivity Analysis, Pooled OLS Estimation Results for the Outcome Equation on a Panel of Firms from the S&P1500 Index

	Model 1	Model 2	Model 3	Model 4
MTB	0.00	0.00	0.00	0.00
	(0.66)	(0.69)	(0.70)	(0.67)
Constant	0.02	0.21	0.29	0.02
	(0.93)	(0.25)	(0.12)	(0.90)
N	2294	2294	2294	2294
R^2	0.20	0.18	0.17	0.21
Wald tests on the joint significance of				
7 IMRs	2.40**	2.66***	2.78***	2.43**
	(0.02)	(0.01)	(0.01)	(0.02)
6 year dummies	4.72***	4.87***	5.26***	4.73***
	(0.00)	(0.00)	(0.00)	(0.00)
Average terms	9.63***	5.61***	4.91***	10.17***
	(0.00)	(0.00)	(0.00)	(0.00)

Notes: P-values are in parentheses.* indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level. The standard errors are adjusted by the cluster option.

Chapter 6: Summary and Conclusions

The objective of this study is to add a peer variable to the traditional type of dividend model to identify whether peer effects play any role in explaining a firm's dividend behavior. Several survey studies (Baker and Powell, 2000; Baker, Veit, and Powell, 2001 and Brav et al., 2005) find that managers consider peer behavior an important influential factor in setting their own dividends. This study examines whether the survey results can be supported by coefficient estimates from a dividend regression model similar to the type commonly employed in the dividend literature. One would conclude that peer effects are present if the dividend behavior of firms depends on the dividend behavior of other firms after common variables that drive dividends, such as firm size, profitability etc., have been accounted for.

Dividend policy has been a puzzle for researchers and firm managers alike. Most studies on dividend policy focus on the traditional dividend theories[58] and analyze commonly used variables in explaining dividend behavior. This study extends the literature on dividend policy by adding to the traditional dividend model explaining a firm's dividend behavior a variable capturing peer effects. I consider dividend initiation and termination to be special cases and limit the analysis of the impact of peer effects only to those firms that have not initiated or terminated dividend payments in the current year compared to the past year.

[58] For example, Ang (1987), Frankfurter (1999), and Lease (2000) for a review of various dividend theories.

The estimation of peer effects is a difficult task because of the reflection problem, possible omitted variable bias, and data availability problems. I have attempted to overcome the omitted variables bias by including numerous common factors that may be driving dividends. The common factors are intended to capture unmeasured environmental effects and, thereby, reduce the omitted variable bias problem. The reflection problem is addressed by using lagged as opposed to current measures of peer influence throughout the study.

The data used in this study are retrieved from the Compustat database. The study focuses on companies that are listed in S&P 1500 Super Composite Index and in the S&P 500 index. The Compustat database contains fundamental financial and market data for U.S. corporations, banks, and industries, such as dividends and earnings information, capital expenditures, stock prices, market capitalizations, firm value, book value of assets, and more. The firms in the S&P 1500 and 500 indices are drawn from different industry groups so that a diverse set of firms is included in the study.

Peer groups can be specified based on the available standard industry classification (SIC) codes, which is selected on a priori grounds. There is no data instigated procedure for the selection of the groups and subgroups. For example, a firm from the mineral industries can never be in the same peer group as a firm from a manufacturing group.

The study suggests a methodology taken from spatial econometrics to identify peer effects. This involves constructing a peer measure of dividend behavior by weighting the dividend behavior of other firms within a pre-selected SIC group by some similarity weight. A spatial lag variable is used as a proxy for the unobserved peer effect.

The spatial lag variable is defined as the matrix product $\mathbf{W}^s\mathbf{y}$, where $\mathbf{W}^s$ is the standardized weight matrix and $\mathbf{y}$ is the dependent variable, which is dividend payments.

Each weight matrix is created based on the equality of SIC codes of the firms involved. Specifically, alternative weight matrices are defined in terms of equal 2-digit, 3-digit, and 4-digit SIC codes. The structure of the weight matrices implies that a firm is only affected by the firms within the same industry group, which is defined in terms of SIC code. Corresponding to the three alternative definitions of weight matrices, three alternative spatial lag variables are created.

In order to investigate the effects of peers on a firm's dividend decision making, the two-step Heckman (also known as heckit) estimation method is employed and applied to annual cross-sections using data from the years 2003, 2004, 2005, and 2006. The Heckman approach is the typical method to correct for sample selection bias. Due to the fact that the dividend paying firms are a self-selected sample, employing simple least squares on those payers would produce biased estimates and invalid inferences. The first stage, the selection equation, of the Heckman two-stage model is a probit model with the dependent variable defined as zero if no dividend is paid and one if a dividend is paid. The second stage regression, the outcome equation, is confined to those firms paying dividends. This is also the equation that contains the peer variable.

The results from the annual cross sections on the sample of firms drawn from the S&P 1500 index reveal a positive and statistically significant relationship between the peer variable and dividend per share for all three definitions of the peer variable. Strong evidence in support of peer effects can also be identified for the sample of firms drawn from the S&P 500 index. The coefficients of the peer variable are of the predicted sign and are statistically significant regardless of the particular definition of the peer variable.

The results are subjected to a sensitivity analysis for the sample of S&P 1500 firms. In particular, it is checked to what extent the addition of lagged dividend per share changes the results. The cross-sectional estimation results indicate that, for the amended equation, a significant peer effect can only be identified when the data from the two most recent years 2005 and 2006 are used. By contrast, the lagged dividend variable is highly significant for all models across all four years, which is consistent with the previous literature. It is interesting to note in this context that the explanatory variables typically employed in dividend regressions become insignificant for the data years 2005 and 2006, when lagged dividend per share and the peer effects variable are both entered into the second-stage OLS models.

The sensitivity analysis reveals that previously identified peer effects no longer appear to exist when lagged dividend per share is added to the models using data from the years 2003 and 2004. The results suggest that firms stick toward their past dividends for the years 2004 and 2005.[59] One potential reason for the fact that the very strong persistence of dividends over time dominates the peer effects at some of the times might be related to changes in the dividend tax law. As there is uncertainty of how to react to the new laws, firms have followed their previous dividend polices. However, peer effects have become significant again for the data years 2005 and 2006 after firms have adjusted to the new laws and have had time to see what others are doing.

---

[59] The data from previous years is used to explain the dividend behavior in the current year.

Panel estimates for the sample of S&P 1500 firms are run using seven years of data from 2000 to 2006. The results show that the peer variable is statistically significant only in the model where peers are identified in terms of two-digit SIC codes. Because there are relatively few observations for the peer variable that are not equal to zero for the three-digit and four-digit variables. Hence, there is not enough variation over time for a firm. The result suggests that there are fewer differences in dividend policy within S&P 1500 firms over time. It is consistent with the fact that the dividend policy is less volatile over time.

References

Aaronson, D. (1998). Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes. *Journal of Human Resources*, 33 (4), 915-946.

Aharony, J., & I. Swany. (1980). Quarterly Dividend and Earnings Announcement and Shareholders' Returns, An Empirical Analysis, *Journal of Finance,* 35, 1-12.

Akerlof, G. (1997) Social Distance and Economic Decisions. *Econometrica*, 65(5), 1005-1027.

Akerlof, G., & R.E. Kranton. (2000). Economics and Identity. *Quarterly Journal of Economics*, 115 (3), 715–753.

Akerlof, G. (2007). The Missing Motivation in Macroeconomics. *American Economic Review*, 97 (1), 5-36.

Alli, K. L., A. Q. Khan, & G. G. Ramirez. (1993). Determinants of Corporate Dividend Policy, A Factorial Analysis. *Financial Review*, 28(4), 523-547.

Ang, J. S. (1987). *Do Dividends Matter? A Review of Corporate Dividend Theories and Evidence.* Monograph Series in Finance and Economics, NY, Salomon Brothers Center for the Study of Financial Institutions, New York University.

Anselin, L. (1988). *Spatial Econometrics, Methods and Models.* Series, Studies in Operational Regional Science, Vol 4, Springer.

Asquith, P., & D. Mullins. (1983).The Impact of Initiating Dividend Payments on Shareholders' Wealth. *Journal of Business,* 56, 77-96.

Baker, H. K., G. E. Farrelly, & R. B. Edelman. (1985). A Survey of Management Views of Dividend Policy. *Financial Management,* 14 (3), 78-84.

Baker, H. K. and G. E. Powell. (2000). Determinants of Corporate Dividend Policy, A Survey of NYSE Firms. *Financial Practice and Education,* 10 (1), 29-40.

Baker, H. K., E. T. Veit, & G. E. Powell. (2001). Factors Influencing Dividend Policy Decisions of NASDAQ Firms. *The Financial Review,* 36 (3), 19-39.

Baker, M., & J. Wurgler.(2004). A Catering Theory of Dividends. *Journal of Finance*, 59, 1125–1165.

Becker, G.S., & K. M. Murphy. (2001). *Social Economics, Market Behavior in a Social Environment.* Harvard University Press, Cambridge.

Benartzi, S., R. Michaely, & R. Thaler. (1997). Do Changes in Dividends Signal the Future or the past? *Journal of Finance*, 52 (3), 1007-1034.

Bhattacharya, S. (1979). Imperfect Information, Dividend Policy, and "the Bird in the Hand" Fallacy. *Bell Journal of Economics*, 10, 259-270.

Black, F., & M. Scholes (1974). The Effect of Dividend Yield and Dividend Policy on Common Stock Prices and Returns. *Journal of Financial Economics,* 1 (1), 1-22.

Black, F. (1976). The Dividend Puzzle. *Journal of Portfolio Management,* 2, 5-8.

Boozer, M., & S. Cacciola. (2001). *Inside the Black Box of Project Star, Estimation of Peer Effects Using Experimental Data.* Yale University Economic Growth Center Discussion Paper no. 832.

Brav, A., J. R. Graham, C. R. Harvey, & R. Michaely. (2005). Payout Policy in the 21st Century. *Journal of Financial Economics*, 77, 483–527.

Brittain, J. (1964). The Tax Structure and Corporate Dividend Policy. *The American Economic Review*, 54 (3), 272-287.

Brock, W. (1993). Pathways to Randomness in the Economy, Emergent Nonlinearity and Chaos in Economics and Finance. *Estudios Economicos*, 8, 3-55.

Brock, W.A. & S.N. Durlauf. (2001a). *Interactions-Based Models.* In, Heckman, J.J., Leamer, E. (Eds.). Handbook of Econometrics. V. 5. North-Holland, Amsterdam.

Brock, W.A., & S. N. Durlauf. (2001b). Discrete Choice with Social Interactions. *Review of Economic Studies*, 68 (2), 235–260.

Charitou, A., & N. Vafeas. (1998). The Association between Operating Cash Flows and Dividend Changes, An Empirical Investigation. *Journal of Business Finance and Accounting,* 25 (1&2), 225-248.

Cont, R., & J. P. Bouchard. (2000). Herd Behavior and Aggregate Fluctuations in Financial Markets. *Macroeconomic Dynamics*, 4(2), 170-196.

Cooper, R., & A. John. (1988). Coordinating Coordination Failures in Keynesian Models. *Quarterly Journal of Economics*, 103(3), 441-463.

DeAngelo, H., L. DeAngelo, & D. J. Skinner. (1992). Dividends and Losses. *Journal of Finance*, 47 (5), 1837-1863.

DeAngelo, H., L. DeAngelo, & D. J. Skinner. (2004). Are Dividends Disappearing? Dividend Concentration and the Consolidation of Earnings. *Journal of Financial Economics,* 72(3), 425-456.

DeAngelo, H., L. DeAngelo, & René M. Stulz. (2006). Dividend Policy and the Earned/Contributed Capital Mix, a Test of the Life-Cycle Theory. *Journal of Financial Economics,* 81 (2), 227-254.

Denis, D. J., & Igor Osobov. (2008). Why Do Firms Pay Dividends? International Evidence on the Determinants of Dividend Policy. *Journal of Financial Economics,* 89 (1), 62-82.

Dhillon, U., & H. Johnson. (1994). Effect of Dividend Changes on Stock and Bond Prices. *The Journal of Finance*, 49 (1), 281-289.

Drewianka, S. (2003). Estimating Social Effects in Matching Markets, Externalities in Spousal Search. *Review of Economics and Statistics*, 85(2), 408-423.

Easterbrook, F. H. (1984). Two Agency Cost Explanations of Dividends. *American Economic Review,* 74 (4), 650-659.

Eije, H. V., & W. L. Megginson. (2008). Dividends and share repurchases in the European Union. *Journal of Financial Economics,* 89 (2), 347-374.

Evans, W. N., W. E. Oates, & R. M. Schwab. (1992). Measuring Peer Group Effects, A Study of Teenage Behavior. *Journal of Political Economy*, 100 (5), 966-991.

Fama, E., & H. Babiak (1968). Dividend Policy-An Empirical Analysis. *American Statistical Association Journal,* December, 1132-1161.

Fama, E., & K. French. (2001). Disappearing Dividends, Changing Firm Characteristics or Lower Propensity to Pay? *Journal of Financial Economics*, 60 (1), 3-43.

Fama, E. F., & K. R. French. (2002). Testing Tradeoff and Pecking Order Predictions about Dividends and Debt. *Review of Financial Studies,* 15, 1-33.

Farrelly, G. E., H. K. Baker, & R. B. Edelman. (1986).Corporate Dividends, Views of the Policymakers. *Akron Business and Economic Review,* 17 (4), 62-74.

Fertig, M. (2003). *Educational Production, Endogenous Peer Group Formation and Class Composition – Evidence From the PISA 2000 Study.* IZA Discussion Paper

No. 714; RWI Discussion Paper No. 2. Available at SSRN: http://ssrn.com/abstract=385163.

Fluck, Z. (1995). *The Optimality of Debt Versus Outside Equity*. NYU Mimeo.

Focardi, S., S. Cincotti, & M. Marchesi. (2002). Self-Organization and Market Crashes. *Journal of Economic Behavior and Organization*, 49, 241-267.

Frankfurter, G. M. (1999). What is the Puzzle in the 'the Dividend Puzzle'? *Journal of Investing*, 8(2), 76-85.

Gaviria, A., & S. Raphael. (2001). School-Based Peer Effects and Juvenile Behavior. *Review of Economics and Statistics*, 83 (2), 257–268.

Glaeser, E., B. Sacerdote, & J. Scheinkman. (1996). Crime and Social Interactions. *Quarterly Journal of Economics*, 111 (2), 507-548.

Gomes, A. (1996). *Dynamics of Stock Prices, Manager Ownership, and Private Benefits of Control*. Manuscript Harvard University.

Hanushek, E. A., J. F. Kain, J. M. Markman, & S. G. Rivkin. (2003). Does Peer Ability Affect Student Achievement? *Journal of Applied Econometric,* 18 (5), 527-544.

Haurin, D.R., R. D. Dietz, & B. A. Weinberg. (2003). The Impact of Neighborhood Homeownership rates, A Review of the Theoretical and Empirical Literature. *Journal of Housing research*, 13(2), 119-151.

Healy, P., & K. G. Palepu (1988). Earnings Information Conveyed by Dividend Initiations and Omissions. *Journal of Financial Economics,* 21 (2), 149-176.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Economometrica*, 47(1), 153-161.

Henderson, J. V., P. Mieszkowski, & Y. Sauvageau. (1978). Peer Group Effects and Educational Production Functions, *Journal of Public Economics*, 10 (1), 97–106.

Hoxby, C.M. (2000). *Peer Effects in the Classroom, Learning From Gender and Race Variation.* Working Paper 7867, NBER.

Hoxby, C. M., & B. Terry. (1999). *Explaining Rising Income and Wage Inequality Among the College Educated.* Working Paper 6873, NBER.

Ioannides, Y., & J. Zabel. (2002a). Neighborhood Effects and Housing Demand. *Journal of Applied Econometrics*, forthcoming.

Ioannides, Y., & J. Zabel. (2002b). *Interactions, Neighborhood Selection, and Housing Demand.* Working Paper, Tufts University.

Jensen, M.C. (1986). Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *American Economic Review*, May, 323-330.

Jackle, R., & O. Himmler. (2007). *Health and Wages, Panel data estimates considering selection and endogeneity.* Munich Personal RePEc Archive Working Paper No. 11578.

John, K., & J. Williams. (1985). Dividends, Dilution, and Taxes, A Signaling Equilibrium. *Journal of Finance*, 40, 1053-1070.

Katz, L.F., J. R. Kling, & J. B. Liebman. (2001). Moving to Opportunity in Boston, Early Results of a Randomized Mobility Experiment. *Quarterly Journal of Economics*, 116 (2), 607–654.

Krauth, B. V. (2003). A Dynamic Model of Job Networking and Social Influences in Employment. *Journal of Economic Dynamics and Control*, forthcoming.

Krauth, B. V. (2006). Simulation-Based Estimation of Peer Effects. *Journal of Econometrics*, 133, 243-271.

Kremer, M., & D.M. Levy. (2001). *Peer Effects from Alcohol Use Among College Students.* Working Paper 9876, NBER.

Lauenstein, M. C. (1987). A New Look at Dividend Strategy. *Journal of Business Strategies,* 8(1), 80-88.

Lease, R. C., K. John, A. Kalay, U. Loewenstein, & O.H. Sarig. (2000). *Dividend Policy, Its Impact on Firm Value.* Harvard Business School Press, Boston, MA.

Li, Kai, & Xinlei Zhao. (2008). Asymmetric Information and Dividend Policy. *Financial Management*, 37 (4), 673 – 694.

Li, Wei, & Erik Lie. (2006). Dividend Changes and Catering Incentives. *Journal of Financial Economics*, 80, 293-308.

Lie, Erik. (2005). Operating Performance following Dividend Decreases and Omissions *Journal of Corporate Finance*, 12, 27– 53.

Lintner, J. (1956). Distribution of Incomes of Corporations among Dividends, Retained Earnings and Taxes. *The American Economic Review,* 46 (2), 97-113.

Litzenberger, R. H., & K. Ramaswamy. (1979). The effect of personal taxes and dividends on capital asset prices. *Journal of Financial Economics*, 7, 163-95.

Lloyd, W. P., J. S. Jahera, & D. E. Page. (1985). Agency Costs and Dividend Payout Ratios. *Quarterly Journal of Business and Economics*, 24, 19-29.

Ludwig, J., G. Duncan, & P. Hirshfield. (2001). Urban Poverty and Juvenile Crime, Evidence from a Randomized Housing–Mobility Experiment. *Quarterly Journal of Economics*, 116 (2), 655-680.

Mancinelli, L., & A.Ozkan. (2006). Ownership Structure and Dividend Policy, Evidence from Italian Firms. *The European Journal of Finance*, 12 (3), 265-282.

Manski, C. F. (1993). Identification of Endogenous Social Effects, The Reflection Problem. *Review of Economic Studies*, 60(3), 531-542.

_______. (2000). Economic Analysis of Social Interactions. *Journal of Economic Perspectives*, 14 (3), 115-136.

Michael, R., R. H. Thaler, & K.L. Womack. (1995). Price Reactions to Dividend Initiations and Omissions, Overreaction or Drift? *The Journal of Finance*, L (2).

Miller, M. H., & F. Modigliani (1961). Dividend Policy, Growth, and the Valuation of Shares. *The Journal of Business,* 34 (4), 411-433.

Miller, M., & K. Rock. (1985). Dividend Policy under Asymmetric Information. *Journal of Finance*, 40, 1031-1052.

Minkin, A. (2002). *Heterogeneous Social Interactions Models*. Mimeo, Department of Economics, University of Wisconsin.

Moffitt, R. A. (2001). *Policy Interventions, Low-Level Equilibria, and Social Interactions*. In Durlauf, S.N., Young, H.P. (Eds.), Social Dynamics. MIT Press, Cambridge, 45–82.

Myers, S. (1996). *Outside Equity Financing*. MIT Working Paper.

Pettit, R. R. (1972). Dividend Announcements, Security Performance and Capital Market Efficiency. *The Journal of Finance,* 27 (5), 993-1007.

Plotnick, R., & S. Hoffman. (1999). The Effect of Neighborhood Characteristics on Young Adult Outcomes, Alternative Estimates. *Social Science Quarterly*. 80 (1), 1-18.

Perez-Gonzalez, F. (2003). *Large Shareholders and Dividends, Evidence from U.S. Tax Reforms.* Columbia University Working Paper.

Poterba, J. (2004). Corporate Payout Policy. *American Economic Review,* 94(2), 171-175.

Poterba, J. M., & L.H. Summers (1984). New Evidence that Taxes Affect the Valuation of Dividends. *Journal of Finance,* 39 (5), 1397-1415.

Poterba, J., & L. Summers. (1985). *The Economic Effects of Dividend Taxation.* In E. Altman and M. Subrahmanyam, eds. Recent Advances in Corporation Finance, (Homewood, IL, Dow Jones-Irwin), 227-284.

Pruitt, S. W., & L.W. Gitman (1991). The interactions between the investment, financing, and dividend decisions of major US firms. *Financial review*, 26 (33), 409-430.

Rivkin, S. G. (1997). The Estimation of Peer Group Effects. Amherst College mimeo.

Rivkin, S. G. (2001). Tiebout Sorting, Aggregation and the Estimation of Peer Group Effects. *Economics of Education Review,* 20 (3).

Rosenhaum, E., & L. Harris. (2001). Residential Mobility and Opportunities, Early Impacts of the Moving to Opportunity Demonstration Program in Chicago. *Housing Policy Debate*, 12(2), 321-346.

Ross, S. A. (1977). The Determination of Financial Structure, The Incentive Signaling Approach. *The Bell Journal of Economics*, 8, 23-40.

Rozeff, M. S. (1982). Growth, Beta and Agency Costs as Determinants of Dividend Payout Ratios. *Journal of Financial Research,* 5 (3), 249-258.

Sacerdote, B. (2001). Peer Effects with Random Assignment, Results for Dartmouth Roommates. *Quarterly Journal of Economics*, 116 (2), 681-704.

Shefrin, H. M., & M. Statman. (1984). Explaining Investor Preference for Cash Dividends. *Journal of Financial Economics*, 13 (2), 253-82.

Shiller, R. J. (1984). Stock Prices and Social Dynamics. *Brookings Papers on Economic Activity*, 457-510.

_________, (1990). Market Volatility and Investor Behavior. *The American Economic Review*, 80(2), 58-62.

Sirakaya, S. (2003). *Recidivism and Social Interactions*. Mimeo, Department of Economics, University of Wisconsin.

Skinner, D. (2008). The Evolving Relation between Earnings, Dividends, and Stock Repurchases. *Journal of Financial Economics*, 87, 582–609.

Weinberg, B., P.Reagan, & J.Yankow. (2002). *Do Neighborhoods Affect Hours Worked? Evidence from Longitudinal Data.* Working paper. Ohio State University.

Wilson, W. J. (1987). The Truly Disadvantaged, The Inner City, the Underclass, and Public Policy. University of Chicago Press, Chicago.

Woolridge, J. R., & C. Ghosh. (1988). An Analysis of Shareholder Reaction to Dividend Cuts and Omissions. *Journal of Financial Research*, 11(4), 281-294.

Woolridge, J. R., & C. Ghosh. (1991). Dividend Omissions and Stock Market Rationality. *The Journal of Business Finance and Accounting*, 18 (3), 315-330.

Wooldridge, J. M. (1995). Selection correction for panel data models under conditional mean independence assumption. *Journal of Econometrics*, 68, 115-132.

Zimmerman, D. J. (2003). Peer Effects in Academic Outcomes, Evidence from a Natural Experiment. *The Review of Economics and Statistics*, 85 (1), 9-23.

Zhou, P., & W. Ruland. (2006).Dividend Payout and Future Earnings Growth. *Financial Analysts Journal,* 62 (3), 58-69.

Appendices

Appendix A: Related Firm-Level Variables Listed by Compustat Data Items

Table 18: Variables Listed by Compustat Data Items

Variables	Compustat data items
Cash & Short Term Investment (millions of dollars)	1
Assets-Total (millions of dollars)	6
Operating Income Before Depreciation (millions of dollars)	13
Special Items (millions of dollars)	17
Income Before Extra Items (millions of dollars)	18
Common Shares Outstanding (millions of dollars)	25
Dividends per Share by Ex-Date (dollars and cents)	26
Retained Earnings (millions of dollars)	36
EPS Basic Exc Extra Items (dollars and cents)	58
Common Equity-Total (millions of dollars)	60
Price-Close Fiscal Year	199

Appendix B: Matlab Codes for Calculating the Standardized Weight Matrix ($\mathbf{W}^s$)

In this study, a spatial lag variable is constructed as a proxy for the peer effects. In order to create the spatial lag variable, a standardized weight matrix $\mathbf{W}^s$ must be calculated. $\mathbf{W}^s$ is a symmetric matrix with zeros on the main diagonal, and its size is *n*x*n*, where *n* is the number of firms in an industry group. The sum of each row has to be one. For instance, if there are 100 firms, the size of the matrix will be 100x100. Each element in the standardized weight matrix $\mathbf{W}^s$ indicates the weight that another firm's dividend decision in the same industry group has on a particular firm's dividend behavior. The weight matrix among firms represents how close these firms are related. Closeness or similarity is measured in this context on the basis of Standard Industry Classification (SIC) Codes.

There are three alternative weight matrices considered, one is based on the same 2-digit SIC codes, another on the same 3-digit SICs, and one is for the same 4-digit SICs. In order to obtain $\mathbf{W}^s$, one needs the information on the number of firms having positive values of dividend per share and the SIC codes for each firm. All firms in the same SIC group receive the same weight.

For the purpose of illustration, the Matlab code is given for calculating the standardized weight matrix $\mathbf{W}^s$ using 2004 data for the S&P 1500 sample,

```
clc; format short g; clear all; close all;

% import data from Excel
X = importdata('W04_SP.xls')

% define variables
id = X.data.Sheet1(,,1);
SIC = X.data.Sheet1(,,3);
n = 533

% sic code is needed
% n identifies the number of firms
% sic contains the SIC codes for each firm
%create matrix with values equal to 1 if two digit ind. is the same
 SIC2=floor(SIC/100);
  for i = 1,n
  mSIC2(,,i) = (abs(SIC2 - SIC2(i)) == 0);
  end

%create matrix with values equal to 1 if three digit ind. is the same
 SIC3=floor(SIC/10);
  for i = 1,n
  mSIC3(,,i) = (abs(SIC3 - SIC3(i)) == 0);
  end

%create matrix with values equal to 1 if four digit ind. is the same
 SIC4=floor(SIC);
  for i = 1,n
  mSIC4(,,i) = (abs(SIC4 - SIC4(i)) == 0);
  end

% remove ones from main diagonal
mSIC2=mSIC2-eye(n,n);
mSIC3=mSIC3-eye(n,n);
mSIC4=mSIC4-eye(n,n);

% the matrices msic2 to msic4 can be used to create a spatial lag
% variable for use in Stata but only after they are standardized

% standardize the matrix
% get the sum for each row
sumr2=sum(mSIC2');
sumr3=sum(mSIC3');
sumr4=sum(mSIC4');

% divide each row element by the above row sum
for i = 1,n
  mSIC2s(i,,) = mSIC2(i,,)./sumr2(i);
  mSIC3s(i,,) = mSIC3(i,,)./sumr3(i);
  mSIC4s(i,,) = mSIC4(i,,)./sumr4(i);
end

mSIC2s(isnan(mSIC2s))=0;
mSIC3s(isnan(mSIC3s))=0;
mSIC4s (isnan (mSIC4s)) =0;
```

5 by 5 submatices of the three alternative weight matrices and of the standardized weight matrices are reported below for illustration.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Weight matrix based on the same 2-digit SICs

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Weight matrix based on the same 3-digit SICs

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Weight matrix based on the same 4-digit SICs

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.033 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0.033 & 0 & 0 & 0 \end{bmatrix}$$

Standardized weight matrix based on the same 2-digit SICs

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Standardized weight matrix based on the same 3-digit SICs

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Standardized weight matrix based on the same 4-digit SICs

Appendix C: Matlab Codes for Constructing the Peer Variable

A spatial lag variable, which is a proxy for a peer effects variable, is defined as the standardized weight matrix multiplied by the dependent variable vector. The dependent variable is dividend per share. Since the peer variable is constructed based on only those firms with positive values of dividends per share, one first needs to identify the number of observations having positive dividends per share for each individual year. For example, there are 533 S&P firms paying positive amounts of dividends per share in year 2004. [60]

For the three alternatives in defining weight matrices, three alternative spatial lag variables are created based on the equality at the 2-digit SIC codes, 3-digit SIC codes, and 4-digit SIC codes, respectively. For consistence, I continue using 2004 data for the S&P 1500 sample to illustrate the calculations.

```
clc; format short g; clear all; close all;

% import data from Excel
X = importdata('W04_SP.xls')

% define variables
id = X.data.Sheet1(,,1);
DPS_04 = X.data.Sheet1(,,2);
SIC = X.data.Sheet1(,,3);
n = 533

% sic code is needed and dependent variable (dividend)
% n identifies the number of firms
% sic contains the SIC codes for each firm
% y contains the dividend data
```

[60] The calculations on the number of firms with positive values of dividends per share are done in STATA. The number of firms is different from the number of dividend payers reported in Table 1. This is because its calculation considers only nonnegative dividends per share and does not take into consideration other restrictions, such as excluding initiations and terminations.

```
%create matrix with values equal to 1 if two digit ind. is the same
 SIC2=floor(SIC/100);
  for i = 1,n
  mSIC2(,,i) = (abs(SIC2 - SIC2(i)) == 0);
  end

%create matrix with values equal to 1 if three digit ind. is the same
 SIC3=floor(SIC/10);
  for i = 1,n
  mSIC3(,,i) = (abs(SIC3 - SIC3(i)) == 0);
  end

%create matrix with values equal to 1 if four digit ind. is the same
 SIC4=floor(SIC);
  for i = 1,n
  mSIC4(,,i) = (abs(SIC4 - SIC4(i)) == 0);
  end

% remove ones from main diagonal
mSIC2=mSIC2-eye(n,n);
mSIC3=mSIC3-eye(n,n);
mSIC4=mSIC4-eye(n,n);

% the matrices msic2 to msic4 can be used to create a spatial lag
% variable for use in Stata
% but only after they are normalized!!!

%normalize the matrix
%get the sum for each row
sumr2=sum(mSIC2');
sumr3=sum(mSIC3');
sumr4=sum(mSIC4');

% divide each row element by the above row sum
for i = 1,n
  mSIC2s(i,,) = mSIC2(i,,)./sumr2(i);
  mSIC3s(i,,) = mSIC3(i,,)./sumr3(i);
  mSIC4s(i,,) = mSIC4(i,,)./sumr4(i);
end

mSIC2s(isnan(mSIC2s))=0;
mSIC3s(isnan(mSIC3s))=0;
mSIC4s(isnan(mSIC4s))=0;

% if one constructs the spatial lag as mSIC4s*DPS, the lag consists
% of a weighted average of peer DPSs, one for each observation

splag2_04=mSIC2s*DPS_04;
splag3_04=mSIC3s*DPS_04;
splag4_04=mSIC4s*DPS_04;
```

The three weight matrices **W** and standardized weight matrices $\mathbf{W}^s$ calculated from the 2004 data are shown in Appendix A. For comparison purposes, the three alternative peer variables reported below are constructed from the corresponding standardized weight matrices $\mathbf{W}^s$ of 2004. Outputs of the spatial lag variables for the first five observations are shown.

$$\begin{bmatrix} 1.08 \\ 0.35 \\ 0.63 \\ 0.34 \\ 0.34 \end{bmatrix}$$

Spatial lag variable constructed based on the 2-digit SICs

$$\begin{bmatrix} 1.07 \\ 0 \\ 0.51 \\ 0.14 \\ 0.11 \end{bmatrix}$$

Spatial lag variable constructed based on the 3-digit SICs

$$\begin{bmatrix} 1.07 \\ 0 \\ 0.57 \\ 0.14 \\ 0.11 \end{bmatrix}$$

Spatial lag variable constructed based on the 4-digit SICs

Appendix D: STATA Codes for the Heckman Two-Stage Estimations

Once the spatial lag variables are created from Matlab, they are added to the initial data base of the S&P 1500 as additional variables. For illustration purposes, STATA code for the Heckman two-stage estimation for the S&P 1500 sample (explaining dividend behavior in 2005 using 2004 data) is provided. STATA code for explaining dividend behavior in 2004 and 2006 are obtained by changing corresponding year-specific variables. .

```
set memory 1000m
set matsize 800

odbc load, dsn("Excel Files;DBQ=C,\Fang
Yang\Dissertation\Oct_2008\SP.xls") table("data$")
*remove financial (SIC, 6000-6999) and utility firms (SIC, 4900-
4949),as well as firms not incorporated in the US
gen du=(SIC>=4900 & SIC<=4949)
gen df=(SIC>=6000 & SIC<=6999)

drop if (INC!= 0 | du==1 | df==1)

*use 2004 data to explain dividend behavior in 2005

*delete missing values on dividends
drop if missing(DPS_04)
drop if missing(DPS_05)

*excluded dividend initiations and terminations obs
drop if (DPS_04==0 & DPS_05>0 )
drop if (DPS_04>0 & DPS_05==0 )

*delete missing values
drop if missing(TA_03)
drop if missing(TA_04)
drop if missing(Income_04)
drop if missing(Spitem_04)
drop if missing(Price_04)
drop if missing(Shares_04)
drop if missing(TE_04)
drop if missing(RE_04)
drop if missing(Cash_04)
drop if missing(OPI_04)
drop if missing(DPS_05)
```

```
*generate dependent variables
gen DIVPY_05=(DPS_05>0)

*generate independent variables
gen SIZE_04=ln(TA_04)
gen E_04=Income_04-0.6*Spitem_04
gen ASG_04=(TA_04-TA_03)/TA_03
gen MTB_04=(Price_04*Shares_04)/TE_04 if (TE_04>0)
gen RETE_04=RE_04/TE_04 if (TE_04>0)
gen CTA_04=Cash_04/TA_04
gen ROA_04=OPI_04/TA_04


*Heckman two-stage method
*first stage regression is probit/logit with no peer effects
*With dividend size on the left in the second equation

heckman DPS_05 SPLAG2_04 EPS_04 SIZE_04 ROA_04 ASG_04 MTB_04 ,twostep
select(DIVPY_05 = E_04 SIZE_04 ROA_04 ASG_04 MTB_04 RETE_04 CTA_04 )
rhosigma

heckman DPS_05 SPLAG3_04 EPS_04 SIZE_04 ROA_04 ASG_04 MTB_04 ,twostep
select(DIVPY_05 = E_04 SIZE_04 ROA_04 ASG_04 MTB_04 RETE_04 CTA_04 )
rhosigma

heckman DPS_05 SPLAG4_04 EPS_04 SIZE_04 ROA_04 ASG_04 MTB_04 ,twostep
select(DIVPY_05 = E_04 SIZE_04 ROA_04 ASG_04 MTB_04 RETE_04 CTA_04 )
rhosigma
```

Appendix E: STATA Codes for the Panel Estimation

```
set memory 1000m
set matsize 800

odbc load, dsn("Excel Files;DBQ=C,\Fang
Yang\Dissertation\Oct_2008\wide_SP.xls") table("data$")

*remove financial (SIC, 6000-6999) and utility firms (SIC, 4900-
4949),as well as firms not incorporated in the US
gen du=(SIC>=4900 & SIC<=4949)
gen df=(SIC>=6000 & SIC<=6999)

drop if (INC!= 0 | du==1 | df==1)

*calculate averages of independent variables
gen E2000=Income2000-0.6*Spitem2000
gen E2001=Income2001-0.6*Spitem2001
gen E2002=Income2002-0.6*Spitem2002
gen E2003=Income2003-0.6*Spitem2003
gen E2004=Income2004-0.6*Spitem2004
gen E2005=Income2005-0.6*Spitem2005
gen E2006=Income2006-0.6*Spitem2006
gen Ebar = (E2000+E2001+E2002+E2003+E2004+E2005+E2006)/7

gen SIZE2000=ln(TA2000)
gen SIZE2001=ln(TA2001)
gen SIZE2002=ln(TA2002)
gen SIZE2003=ln(TA2003)
gen SIZE2004=ln(TA2004)
gen SIZE2005=ln(TA2005)
gen SIZE2006=ln(TA2006)
gen
SIZEbar=(SIZE2000+SIZE2001+SIZE2002+SIZE2003+SIZE2004+SIZE2005+SIZE200)
/7

gen ROA2000=OPI2000/TA2000
gen ROA2001=OPI2001/TA2001
gen ROA2002=OPI2002/TA2002
gen ROA2003=OPI2003/TA2003
gen ROA2004=OPI2004/TA2004
gen ROA2005=OPI2005/TA2005
gen ROA2006=OPI2006/TA2006
gen ROAbar=(ROA2000+ROA2001+ROA2002+ROA2003+ROA2004+ROA2005+ROA2006)/7

gen ASG2000=(TA2000-TA1999)/TA1999
gen ASG2001=(TA2001-TA2000)/TA2000
gen ASG2002=(TA2002-TA2001)/TA2001
gen ASG2003=(TA2003-TA2002)/TA2002
gen ASG2004=(TA2004-TA2003)/TA2003
gen ASG2005=(TA2005-TA2004)/TA2004
gen ASG2006=(TA2006-TA2005)/TA2005
gen ASGbar=(ASG2000+ASG2001+ASG2002+ASG2003+ASG2004+ASG2005+ASG2006)/7
```

```
gen MTB2000=(Price2000*Shares2000)/TE2000 if (TE2000>0)
gen MTB2001=(Price2001*Shares2001)/TE2001 if (TE2001>0)
gen MTB2002=(Price2002*Shares2002)/TE2002 if (TE2002>0)
gen MTB2003=(Price2003*Shares2003)/TE2003 if (TE2003>0)
gen MTB2004=(Price2004*Shares2004)/TE2004 if (TE2004>0)
gen MTB2005=(Price2005*Shares2005)/TE2005 if (TE2005>0)
gen MTB2006=(Price2006*Shares2006)/TE2006 if (TE2006>0)
gen MTBbar=(MTB2000+MTB2001+MTB2002+MTB2003+MTB2004+MTB2005+MTB2006)/7

gen RETE2000=RE2000/TE2000 if (TE2000>0)
gen RETE2001=RE2001/TE2001 if (TE2001>0)
gen RETE2002=RE2002/TE2002 if (TE2002>0)
gen RETE2003=RE2003/TE2003 if (TE2003>0)
gen RETE2004=RE2004/TE2004 if (TE2004>0)
gen RETE2005=RE2005/TE2005 if (TE2005>0)
gen RETE2006=RE2006/TE2006 if (TE2006>0)
gen
RETEbar=(RETE2000+RETE2001+RETE2002+RETE2003+RETE2004+RETE2005+RETE2006
)/7


gen CTA2000=Cash2000/TA2000
gen CTA2001=Cash2001/TA2001
gen CTA2002=Cash2002/TA2002
gen CTA2003=Cash2003/TA2003
gen CTA2004=Cash2004/TA2004
gen CTA2005=Cash2005/TA2005
gen CTA2006=Cash2006/TA2006
gen CTAbar=(CTA2000+CTA2001+CTA2002+CTA2003+CTA2004+CTA2005+CTA2006)/7

gen EPSbar=(EPS2000+EPS2001+EPS2002+EPS2003+EPS2004+EPS2005+EPS2006)/7
gen SPLAG2bar
=(SPLAGTW2000+SPLAGTW2001+SPLAGTW2002+SPLAGTW2003+SPLAGTW2004+SPLAGTW20
05+SPLAGTW2006)/7
gen SPLAG3bar
=(SPLAGTR2000+SPLAGTR2001+SPLAGTR2002+SPLAGTR2003+SPLAGTR2004+SPLAGTR20
05+SPLAGTR2006)/7
gen SPLAG4bar
=(SPLAGFO2000+SPLAGFO2001+SPLAGFO2002+SPLAGFO2003+SPLAGFO2004+SPLAGFO20
05+SPLAGFO2006)/7

*to calculate IMR2000
*excluded dividend initiations and terminations obs
drop if missing(DPS2000)
drop if missing(DPS2001)
drop if (DPS2000==0 & DPS2001>0)
drop if (DPS2000>0 & DPS2001==0)

*delete missing values
drop if missing(E2000)
drop if missing(SIZE2000)
drop if missing(ROA2000)
drop if missing(ASG2000)
drop if missing(MTB2000)
drop if missing(RETE2000)
```

```
drop if missing(CTA2000)
drop if missing(Ebar)
drop if missing(SIZEbar)
drop if missing(ROAbar)
drop if missing(ASGbar)
drop if missing(MTBbar)
drop if missing(RETEbar)
drop if missing(CTAbar)

gen DIVPY2001=(DPS2001>0)

*estmate T years probit models of the selection equation involving the
averages of the independent variables
probit DIVPY2001 E2000 SIZE2000 ROA2000 ASG2000 MTB2000 RETE2000
CTA2000 Ebar SIZEbar ROAbar ASGbar MTBbar RETEbar CTAbar
*xb calculates the linear prediction from the fitted model.
predict xb, xb
*generate inverse of Mills' ratio
gen IMR2000 = normalden(xb)/normal(xb)
keep id IMR2000
save IMR2000

*to calculate IMR2001 (need to read into the data again)
*excluded dividend initiations and terminations obs
drop if missing(DPS2001)
drop if missing(DPS2002)
drop if (DPS2001==0 & DPS2002>0)
drop if (DPS2001>0 & DPS2002==0)

*delete missing values
drop if missing(E2001)
drop if missing(SIZE2001)
drop if missing(ROA2001)
drop if missing(ASG2001)
drop if missing(MTB2001)
drop if missing(RETE2001)
drop if missing(CTA2001)
drop if missing(Ebar)
drop if missing(SIZEbar)
drop if missing(ROAbar)
drop if missing(ASGbar)
drop if missing(MTBbar)
drop if missing(RETEbar)
drop if missing(CTAbar)

gen DIVPY2002=(DPS2002>0)

*estmate T years probit models of the selection equation involving the
averages of the independent variables
probit DIVPY2002 E2001 SIZE2001 ROA2001 ASG2001 MTB2001 RETE2001
CTA2001 ///
       Ebar SIZEbar ROAbar ASGbar MTBbar RETEbar CTAbar
*xb calculates the linear prediction from the fitted model.
predict xb, xb
*generate inverse of Mills' ratio
```

```
gen IMR2001 = normalden(xb)/normal(xb)
keep id IMR2001
save IMR2001

*to calculate IMR2002 (need to read into the data again)
*excluded dividend initiations and terminations obs
drop if missing(DPS2002)
drop if missing(DPS2003)
drop if (DPS2002==0 & DPS2003>0)
drop if (DPS2002>0 & DPS2003==0)

*delete missing values
drop if missing(E2002)
drop if missing(SIZE2002)
drop if missing(ROA2002)
drop if missing(ASG2002)
drop if missing(MTB2002)
drop if missing(RETE2002)
drop if missing(CTA2002)
drop if missing(Ebar)
drop if missing(SIZEbar)
drop if missing(ROAbar)
drop if missing(ASGbar)
drop if missing(MTBbar)
drop if missing(RETEbar)
drop if missing(CTAbar)

gen DIVPY2003=(DPS2003>0)

*estmate T years probit models of the selection equation involving the
averages of the independent variables
probit DIVPY2003 E2002 SIZE2002 ROA2002 ASG2002 MTB2002 RETE2002
CTA2002 ///
       Ebar SIZEbar ROAbar ASGbar MTBbar RETEbar CTAbar
*xb calculates the linear prediction from the fitted model.
predict xb, xb
*generate inverse of Mills' ratio
gen IMR2002 = normalden(xb)/normal(xb)
keep id IMR2002
save IMR2002

*to calculate IMR2003 (need to read into the data again)
*excluded dividend initiations and terminations obs
drop if missing(DPS2003)
drop if missing(DPS2004)
drop if (DPS2003==0 & DPS2004>0)
drop if (DPS2003>0 & DPS2004==0)

*delete missing values
drop if missing(E2003)
drop if missing(SIZE2003)
drop if missing(ROA2003)
drop if missing(ASG2003)
drop if missing(MTB2003)
drop if missing(RETE2003)
```

```
drop if missing(CTA2003)
drop if missing(Ebar)
drop if missing(SIZEbar)
drop if missing(ROAbar)
drop if missing(ASGbar)
drop if missing(MTBbar)
drop if missing(RETEbar)
drop if missing(CTAbar)

gen DIVPY2004=(DPS2004>0)

*estmate T years probit models of the selection equation involving the
averages of the independent variables
probit DIVPY2004 E2003 SIZE2003 ROA2003 ASG2003 MTB2003 RETE2003
CTA2003 Ebar SIZEbar ROAbar ASGbar MTBbar RETEbar CTAbar
*xb calculates the linear prediction from the fitted model.
predict xb, xb
*generate inverse of Mills' ratio
gen IMR2003 = normalden(xb)/normal(xb)
keep id IMR2003
save IMR2003

*to calculate IMR2004 (need to read into the data again)
*excluded dividend initiations and terminations obs
drop if missing(DPS2004)
drop if missing(DPS2005)
drop if (DPS2004==0 & DPS2005>0)
drop if (DPS2004>0 & DPS2005==0)

*delete missing values
drop if missing(E2004)
drop if missing(SIZE2004)
drop if missing(ROA2004)
drop if missing(ASG2004)
drop if missing(MTB2004)
drop if missing(RETE2004)
drop if missing(CTA2004)
drop if missing(Ebar)
drop if missing(SIZEbar)
drop if missing(ROAbar)
drop if missing(ASGbar)
drop if missing(MTBbar)
drop if missing(RETEbar)
drop if missing(CTAbar)

gen DIVPY2005=(DPS2005>0)

*estmate T years probit models of the selection equation involving the
averages of the independent variables
probit DIVPY2005 E2004 SIZE2004 ROA2004 ASG2004 MTB2004 RETE2004
CTA2004 Ebar SIZEbar ROAbar ASGbar MTBbar RETEbar CTAbar
*xb calculates the linear prediction from the fitted model.
predict xb, xb
*generate inverse of Mills' ratio
gen IMR2004 = normalden(xb)/normal(xb)
```

```
keep id IMR2004
save IMR2004

*to calculate IMR2005 (need to read into the data again)
*excluded dividend initiations and terminations obs
drop if missing(DPS2005)
drop if missing(DPS2006)
drop if (DPS2005==0 & DPS2006>0)
drop if (DPS2005>0 & DPS2006==0)

*delete missing values
drop if missing(E2005)
drop if missing(SIZE2005)
drop if missing(ROA2005)
drop if missing(ASG2005)
drop if missing(MTB2005)
drop if missing(RETE2005)
drop if missing(CTA2005)
drop if missing(Ebar)
drop if missing(SIZEbar)
drop if missing(ROAbar)
drop if missing(ASGbar)
drop if missing(MTBbar)
drop if missing(RETEbar)
drop if missing(CTAbar)

gen DIVPY2006=(DPS2006>0)

*estmate T years probit models of the selection equation involving the
averages of the independent variables
probit DIVPY2006 E2005 SIZE2005 ROA2005 ASG2005 MTB2005 RETE2005
CTA2005 Ebar SIZEbar ROAbar ASGbar MTBbar RETEbar CTAbar
*xb calculates the linear prediction from the fitted model.
predict xb, xb
*generate inverse of Mills' ratio
gen IMR2005 = normalden(xb)/normal(xb)
keep id IMR2005
save IMR2005

*to calculate IMR2006 (need to read into the data again)
*excluded dividend initiations and terminations obs
drop if missing(DPS2006)
drop if missing(DPS2007)
drop if (DPS2006==0 & DPS2007>0)
drop if (DPS2006>0 & DPS2007==0)

*delete missing values
drop if missing(E2006)
drop if missing(SIZE2006)
drop if missing(ROA2006)
drop if missing(ASG2006)
drop if missing(MTB2006)
drop if missing(RETE2006)
drop if missing(CTA2006)
drop if missing(Ebar)
```

```
drop if missing(SIZEbar)
drop if missing(ROAbar)
drop if missing(ASGbar)
drop if missing(MTBbar)
drop if missing(RETEbar)
drop if missing(CTAbar)

gen DIVPY2007=(DPS2007>0)

*estmate T years probit models of the selection equation involving the
averages of the independent variables
probit DIVPY2007 E2006 SIZE2006 ROA2006 ASG2006 MTB2006 RETE2006
CTA2006 Ebar SIZEbar ROAbar ASGbar MTBbar RETEbar CTAbar
*xb calculates the linear prediction from the fitted model.
predict xb, xb
*generate inverse of Mills' ratio
gen IMR2006 = normalden(xb)/normal(xb)
keep id IMR2006
save IMR2006

clear

use IMR2000
sort id
save IMR1, replace

use IMR2001
sort id
save IMR2, replace

use IMR2002
sort id
save IMR3, replace

use IMR2003
sort id
save IMR4, replace

use IMR2004
sort id
save IMR5, replace

use IMR2005
sort id
save IMR6, replace

use IMR2006
sort id
save IMR7, replace


*need to read into the data again
odbc load, dsn("Excel Files;DBQ=C,\Fang
Yang\Dissertation\Oct_2008\wide_SP.xls") table("data$")
```

```
*remove financial (SIC, 6000-6999) and utility firms (SIC, 4900-
4949),as well as firms not incorporated in the US
gen du=(SIC>=4900 & SIC<=4949)
gen df=(SIC>=6000 & SIC<=6999)

drop if (INC!= 0 | du==1 | df==1)

*calculate averages of independent variables
gen E2000=Income2000-0.6*Spitem2000
gen E2001=Income2001-0.6*Spitem2001
gen E2002=Income2002-0.6*Spitem2002
gen E2003=Income2003-0.6*Spitem2003
gen E2004=Income2004-0.6*Spitem2004
gen E2005=Income2005-0.6*Spitem2005
gen E2006=Income2006-0.6*Spitem2006
gen Ebar = (E2000+E2001+E2002+E2003+E2004+E2005+E2006)/7

gen SIZE2000=ln(TA2000)
gen SIZE2001=ln(TA2001)
gen SIZE2002=ln(TA2002)
gen SIZE2003=ln(TA2003)
gen SIZE2004=ln(TA2004)
gen SIZE2005=ln(TA2005)
gen SIZE2006=ln(TA2006)
gen
SIZEbar=(SIZE2000+SIZE2001+SIZE2002+SIZE2003+SIZE2004+SIZE2005+SIZE2006
)/7

gen ROA2000=OPI2000/TA2000
gen ROA2001=OPI2001/TA2001
gen ROA2002=OPI2002/TA2002
gen ROA2003=OPI2003/TA2003
gen ROA2004=OPI2004/TA2004
gen ROA2005=OPI2005/TA2005
gen ROA2006=OPI2006/TA2006
gen ROAbar=(ROA2000+ROA2001+ROA2002+ROA2003+ROA2004+ROA2005+ROA2006)/7

gen ASG2000=(TA2000-TA1999)/TA1999
gen ASG2001=(TA2001-TA2000)/TA2000
gen ASG2002=(TA2002-TA2001)/TA2001
gen ASG2003=(TA2003-TA2002)/TA2002
gen ASG2004=(TA2004-TA2003)/TA2003
gen ASG2005=(TA2005-TA2004)/TA2004
gen ASG2006=(TA2006-TA2005)/TA2005
gen ASGbar=(ASG2000+ASG2001+ASG2002+ASG2003+ASG2004+ASG2005+ASG2006)/7

gen MTB2000=(Price2000*Shares2000)/TE2000 if (TE2000>0)
gen MTB2001=(Price2001*Shares2001)/TE2001 if (TE2001>0)
gen MTB2002=(Price2002*Shares2002)/TE2002 if (TE2002>0)
gen MTB2003=(Price2003*Shares2003)/TE2003 if (TE2003>0)
gen MTB2004=(Price2004*Shares2004)/TE2004 if (TE2004>0)
gen MTB2005=(Price2005*Shares2005)/TE2005 if (TE2005>0)
gen MTB2006=(Price2006*Shares2006)/TE2006 if (TE2006>0)
gen MTBbar=(MTB2000+MTB2001+MTB2002+MTB2003+MTB2004+MTB2005+MTB2006)/7
```

```
gen RETE2000=RE2000/TE2000 if (TE2000>0)
gen RETE2001=RE2001/TE2001 if (TE2001>0)
gen RETE2002=RE2002/TE2002 if (TE2002>0)
gen RETE2003=RE2003/TE2003 if (TE2003>0)
gen RETE2004=RE2004/TE2004 if (TE2004>0)
gen RETE2005=RE2005/TE2005 if (TE2005>0)
gen RETE2006=RE2006/TE2006 if (TE2006>0)
gen
RETEbar=(RETE2000+RETE2001+RETE2002+RETE2003+RETE2004+RETE2005+RETE2006
)/7

gen CTA2000=Cash2000/TA2000
gen CTA2001=Cash2001/TA2001
gen CTA2002=Cash2002/TA2002
gen CTA2003=Cash2003/TA2003
gen CTA2004=Cash2004/TA2004
gen CTA2005=Cash2005/TA2005
gen CTA2006=Cash2006/TA2006
gen CTAbar=(CTA2000+CTA2001+CTA2002+CTA2003+CTA2004+CTA2005+CTA2006)/7

gen EPSbar=(EPS2000+EPS2001+EPS2002+EPS2003+EPS2004+EPS2005+EPS2006)/7
gen SPLAG2bar
=(SPLAGTW2000+SPLAGTW2001+SPLAGTW2002+SPLAGTW2003+SPLAGTW2004+SPLAGTW20
05+SPLAGTW2006)/7
gen SPLAG3bar
=(SPLAGTR2000+SPLAGTR2001+SPLAGTR2002+SPLAGTR2003+SPLAGTR2004+SPLAGTR20
05+SPLAGTR2006)/7
gen SPLAG4bar
=(SPLAGFO2000+SPLAGFO2001+SPLAGFO2002+SPLAGFO2003+SPLAGFO2004+SPLAGFO20
05+SPLAGFO2006)/7


*Reshaping data into panel format
drop DPS2007 SPLAGTW2007 SPLAGTR2007 SPLAGFO2007 EPS2007
reshape long DPS SPLAGTW SPLAGTR SPLAGFO EPS SIZE ROA ASG MTB, i(id)
j(year)
tsset id year
summarize

*create year dummies
tabulate year, gen(yd)
list year yd1 yd2 yd3 yd4 yd5 yd6

drop if missing(IMR2000)
drop if missing(IMR2001)
drop if missing(IMR2002)
drop if missing(IMR2003)
drop if missing(IMR2004)
drop if missing(IMR2005)
drop if missing(IMR2006)

replace IMR2000 = 0 if (year !=2000)
replace IMR2001 = 0 if (year !=2001)
replace IMR2002 = 0 if (year !=2002)
replace IMR2003 = 0 if (year !=2003)
```

```
replace IMR2004 = 0 if (year !=2004)
replace IMR2005 = 0 if (year !=2005)
replace IMR2006 = 0 if (year !=2006)

save SP_ Paneldata, replace

odbc load, dsn("Excel Files;DBQ=C,\Fang
Yang\Dissertation\Oct_2008\SP_Paneldata.xls") table("SPdata$")

*delete missing values
drop if missing(DPS)
drop if missing(EPS)
drop if missing(SIZE)
drop if missing(ROA)
drop if missing(ASG)
drop if missing(MTB)
drop if missing(EPSbar)
drop if missing(SIZEbar)
drop if missing(ROAbar)
drop if missing(ASGbar)
drop if missing(MTBbar)
drop if missing(SPLAG2bar)
drop if missing(SPLAG3bar)
drop if missing(SPLAG4bar)
drop if missing(IMR00)
drop if missing(IMR01)
drop if missing(IMR02)
drop if missing(IMR03)
drop if missing(IMR04)
drop if missing(IMR05)
drop if missing(IMR06)


*With dividend size on the left in the second equation
*Inverse Mills Ratio is included in the second equation as a control
for selection bias
*including IMR allows us to get unbiased coefficient estimates

*with spatial lag variable SPLAGTW

*results with clustered standard errors
sum DPS SPLAGTW EPS SIZE ROA ASG MTB IMR00 IMR01 IMR02 IMR03 IMR04
IMR05 IMR06 EPSbar SIZEbar ROAbar ASGbar MTBbar SPLAG2bar yd1 yd2 yd3
yd4 yd5 yd6

reg DPS SPLAGTW EPS SIZE ROA ASG MTB IMR00 IMR01 IMR02 IMR03 IMR04
IMR05 IMR06 EPSbar SIZEbar ROAbar ASGbar MTBbar SPLAG2bar yd1 yd2 yd3
yd4 yd5 yd6, cluster (id)
test IMR00  IMR01 IMR02  IMR03  IMR04  IMR05  IMR06
test EPSbar  SIZEbar ROAbar ASGbar MTBbar SPLAG2bar
test yd1 yd2 yd3 yd4 yd5 yd6


*with spatial lag variable SPLAGTR
```

```
*results with clustered standard errors
sum DPS SPLAGTR EPS SIZE ROA ASG MTB IMR00 IMR01 IMR02 IMR03 IMR04
IMR05 IMR06 EPSbar SIZEbar ROAbar ASGbar MTBbar SPLAG3bar yd1 yd2 yd3
yd4 yd5 yd6

reg DPS SPLAGTR EPS SIZE ROA ASG MTB IMR00 IMR01 IMR02 IMR03 IMR04
IMR05 IMR06 EPSbar SIZEbar ROAbar ASGbar MTBbar SPLAG3bar yd1 yd2 yd3
yd4 yd5 yd6, cluster (id)
test IMR00 IMR01 IMR02 IMR03 IMR04 IMR05 IMR06
test EPSbar SIZEbar ROAbar ASGbar MTBbar SPLAG3bar
test yd1 yd2 yd3 yd4 yd5 yd6


*with spatial lag variable SPLAGFO
*results with clustered standard errors
drop if missing(SPLAGFO)
reg DPS SPLAGFO EPS SIZE ROA ASG MTB IMR00 IMR01 IMR02 IMR03 IMR04
IMR05 IMR06 EPSbar SIZEbar ROAbar ASGbar MTBbar SPLAG4bar yd1 yd2 yd3
yd4 yd5 yd6, cluster (id)
test IMR00 IMR01 IMR02 IMR03 IMR04 IMR05 IMR06
test EPSbar SIZEbar ROAbar ASGbar MTBbar SPLAG4bar
test yd1 yd2 yd3 yd4 yd5 yd6
```