2016

Three Essays on Asset Pricing in Security and Housing Markets

Minrong Zheng
University of Central Florida

THREE ESSAYS ON ASSET PRICING IN SECURITY AND HOUSING MARKETS

by

MINRONG ZHENG
B.E. Daqing Petroleum Institute, 1993
M.E. Dalian University of Technology, 1996
M.S. University of Nebraska-Lincoln, 2006, 2007

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Finance
in the College of Business Administration
at the University of Central Florida
Orlando, Florida

Fall Term
2016

Major Professors: Honghui Chen, Geoffrey K. Turnbull

ABSTRACT

In my first essay, I investigate the relationship between IPO long-run underperformance (Ritter, 1991) and the idiosyncratic risk puzzle (Ang, Hodrick, Xing and Zhang, 2006), the phenomenon of abnormally low returns for stocks with high idiosyncratic risk. I show that IPO long-run underperformance is in fact a manifestation of the surprisingly low returns for high idiosyncratic risk stocks. IPO underperformance disappears after I control for the idiosyncratic risk. Specifically, the underperformance of IPO firms only presents following the months in which they are classified into the highest idiosyncratic risk quintile. On the other hand, I find that the idiosyncratic risk puzzle is magnified by the IPO underperformance for two reasons. First, IPOs are over-represented in the highest volatility quintile. Second, while stocks in the highest volatility quintile underperform in general, the intra-quintile underperformance is substantially more severe for the IPO firms. My results are robust to different sample requirements.

My second essay examines school quality and quality risk capitalization when school quality is uncertain, taking into account uncertainty induced by low signal content in quality measures available to parents or stochastic quality outcomes. Extending the residential bid rent theory to the uncertainty environment, the theory shows that greater school quality increases housing prices steepens the price gradient, whereas the quality risk decreases the housing prices and flattens the price gradient. The empirical models incorporate two sources of quality risk, the variance in measured school quality and school attendance zone instability. Coupling an output based measure using the over-period average of school normalized math test scores based on the Orange County public elementary school average scores with an input based measure using student/teacher ratios provides quality measures that appear to correlate sufficiently with parents'

perceptions of elementary school quality, but school peer effects play important role as well. Estimates reveal capitalization of quality and uncertainty that are consistent with theory as well as systematic patterns across housing market phases and neighborhood in income level.

My third essay is a meta-analysis of the body of empirical results for school quality capitalization in house prices. One puzzling aspect of the housing markets literature is that, while public school quality is a major concern of many households, empirical studies of school quality capitalization into house prices yield mixed and sometimes inconsistent results not only across studies, but also within studies when using different school quality measures and models. These differences are reflected in the capitalization coefficient value, level of significance, and even direction of capitalization effects. This paper conducts meta-analysis of the school quality capitalization estimates to identify the factors contributing to this variation. It reveals that the way the school quality is measured matters. Peer effects measures yield less significant capitalization estimates than input and output based measures and value added measures exhibit lower significance than other output based measures. Moreover, both boundary fixed effects and neighborhood fixed effect approaches can effectively and significantly control for the influence of neighborhood amenities. Adding more school quality variables reduces the capitalization significance of individual school quality variables. The most unexpected finding is that school quality capitalization significance is much less in the South than in other regions. Also surprising is that econometric methods do not appear to be driving results.

To honor the memory of my late father

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my committee members, Professor Honghui Chen, Professor Geoffrey K. Turnbull, Professor Melissa Frye, and Professor Velma Zahirovic-Herbert at the University of Georgia for their support, encouragements, guidance, patience, and persistent help.

In addition, I would like to thank seminar participants in 2013 SFA Annual Conference, 2014 FMA Annual Conference, 2015 ARES Annual Meeting, 2016 AREUEA-ASSA Annual Conference, 2016 SWFA Annual Conference and 2016 ARES Annual Meeting. Their thoughtful questions and comments were valued greatly.

Finally, I would like also to thank my husband, Dr. Jianbin Zhu, my daughter, Ms. Mingyin Zhu, and my son, Mr. Samuel Zhu, and all my family members in China. Without their unconditional love and support during my whole Ph.D. period, I would not have been able to complete this dissertation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EAASY1: IPO UNDERPERFORMANCE AND THE IDIOSYNCRATIC RISK PUZZLE

1. Introduction

In this paper, we investigate the relation between two well-known puzzles, the IPO long-run underperformance and the idiosyncratic risk puzzle, or the abnormally low return for stocks with high idiosyncratic risk, first documented by Ang, Hodrick, Xing and Zhang (AHXZ 2006). While both phenomena have been extensively investigated on an individual level, we are not aware of any systematic attempts to link the two. We aim to develop a better understanding of these two puzzles by identifying similar underlying causes.

Ritter (1991) and Loughran and Ritter (1995) find that IPO firms underperform their peers and indexes for a period of up to five years after IPO. Ritter has extended this analysis to US IPOs in recent years and summarized his findings on his website.[1] While his evidence is confined to the underperformance of IPOs in the US market, the IPO long-term underperformance is not unique to the US market. Other studies confirm the existence of IPO long-run underperformance in many different countries outside of the US.[2]

Since the documentation of the phenomena of abnormally low return for stocks with high idiosyncratic risk by AHXZ (2006), many explanations have been offered for the idiosyncratic risk

[1] http://bear.warrington.ufl.edu/ritter/IPOs2013-5years.pdf

[2] For example, Aggarwal, Leal, and Hermandez (1993) report three-year market-adjusted negative returns of IPOs in Brazil, Mexico, and Chile. In the UK, Levis (1995) shows that IPOs experience price declines of 23% after adjusting for all-share-index returns for three years after the first trade day. Similarly, Dimovski and Brook (2004) show that IPOs underperform in the market using data from 1994-1998 in Australia. Agarwal, Liu, and Rhee (2008) document IPO long-term underperformance in Hong Kong.

puzzle[3]. One prominent explanation is related to Miller's (1977) theory, which argues that, when divergence of opinion is coupled with short sale constraints, the pessimistic investors are forced to sit out of the market and the price will be set by the most optimistic investors, which leads to overvaluation. This overvaluation effect is expected to be more pronounced for securities with greater divergence of opinion. Overvaluation is followed by low performance. If high idiosyncratic volatility indicates high divergence of opinion, stocks with high idiosyncratic volatility should demonstrate the postulated overvaluation effect and be followed by poor performance. AHXZ (2006) shows that Miller's theory can partially explain the idiosyncratic risk puzzle. Guo and Qiu (2014) also find a stronger negative relation between the options-implied variance and future stock returns for stocks with short-selling constraints or when it is more difficult to short sell stocks.

Another explanation developed by Stambaugh, Yu, and Yuan (2014) also connects high idiosyncratic risk with high levels of mispricing. This perspective argues that high idiosyncratic risk has an asymmetric effect on arbitrage activity.[4] More specifically, the asymmetric effect results from the fact that high idiosyncratic volatility has a greater impact on a short position than on a long position. Therefore, it is more difficult to eliminate the overvaluation among the more volatile stocks than among the less volatile stocks. Consequently, we expect a much greater effect

[3] Recent papers confirming and intending to explain the idiosyncratic risk puzzle include: Bali and Cakici (2008), Ang, Hodrick, Xing and Zhang (2009), Boehme, Danielsen, Kumar, and Sorescu (2009), Boyer, Mitton, and Vorkink (2009), Fu (2009), Chen, Huang, and Jha (2010), Chen, Chollete, and Ray (2010), Huang, Liu, Rhee, ang Zhang (2010), Bali, Cakici, and Whitelaw (2011), Han and Lesmond (2011), George and Hwang (2013), Stambaugh, Yu, and Yuan (2015), Hou and Loh (2016), and so on.

[4] Arbitrage can eliminate mispricing and enhance market efficiency. However, existence of arbitrage risk and arbitrage costs limits this activity and impedes the stock's price reverting to its true value (Pontiff, 1996, Shleifer and Vishny, 1997, Cheng, Hong and Stein, 2002, Mitchell, Pulvino, and Stafford, 2002, Wurgler, and Zhuravskaya, 2002, Ofek, Richardson,and Whitelaw, 2004, Mashruwala, Rajgopal, and Shevlin, 2006, Pontiff, 2006, Cao and Han, 2010, Duan, Hu, McLean, 2010, Li and Sullivan, 2011).

of idiosyncratic risk on subsequent returns for the overvalued stocks than for the undervalued stocks. This stronger effect on overvalued stocks generates the overall poorer performance for the more volatile stocks. This explanation is similar to Miller's (1977) explanation, since both propose that higher volatility makes short sale more difficult.

While both explanations that relate high volatility to long-run underperformance discussed above are applicable to all firms, there are reasons to expect that these explanations may be especially valid for the IPO firms. Schultz (2003) finds that the variance of excess returns and the sensitivity of the number of offerings to the level of an index of recent IPOs jointly determine the expected level of the abnormal returns. This supports the idea that long-run IPO underperformance may share the same underlying causes as the idiosyncratic risk puzzle by suggesting that the higher volatility of the excess returns worsens IPO firm performance.

With respect to Miller (1977), if IPO firms are more likely to have greater divergence of opinion than non-IPO firms, they are expected to be overvalued and will be followed by poor performance as divergence of opinion narrows. This not only suggests that IPO firms may be overrepresented among highly volatile stocks and contribute to the idiosyncratic risk puzzle, but also suggests that high idiosyncratic risk resulting from greater divergence of opinion following initial offering worsens underperformance, even among IPO firms. Many empirical studies provide evidence consistent with such a prediction connecting the divergence of opinion to long-term performance. Houge, Loughran, Suchanek, and Yan (2001) use the percentage opening bid-ask spread and the time of the first trade to capture the uncertainty about an IPO and use the flipping ratio as a proxy for the divergence of opinion between institutional and individual investors. They find that a wide opening spread and late opening trade, indicating a high uncertainty, in conjunction with a high flipping ratio, indicating high divergence of opinion, are associated with poor long-run returns. Gao, Mao, and Zhong (2006) also investigate the relationship between divergence of opinion and long-term returns for IPO firms using early-market return volatility as a proxy for divergence of opinion and find that divergence of opinion is negatively related to subsequent IPO long-term abnormal returns.

Moreover, higher idiosyncratic volatility may also be related to poorer IPO firm performance if idiosyncratic volatility reflects not only divergence of opinion but also inferior information quality and greater uncertainty. Teoh, Welch, and Wong (1998) further support the idea that high idiosyncratic volatility is relevant in understanding the IPO underperformance puzzle and show that issuers with unusually high accruals in the IPO year experience poor stock return performance during the three years thereafter. This connection is important, because several studies relate high abnormal accruals to high idiosyncratic volatility. For example, Rajgopal and Venkatachalam (2011) show that deteriorating earnings quality, as reflected by high abnormal accruals, is associated with higher idiosyncratic return volatility. Chen, Huang, and Stein (2012) show that, in addition to pre-managed earnings volatility, both discretionary accruals volatility and the correlation between the above two aspects contribute to high return idiosyncratic volatility. Moreover, managerial discretion in accruals affects the return idiosyncratic volatility via information quality. Fan and Yu (2013) document a positive relation between the abnormal accrual and idiosyncratic volatility with international data. To the extent that firms with unusually high accruals present more uncertainty and hence greater idiosyncratic risk, the evidence from these studies also suggests a relation between idiosyncratic risk puzzle and IPO performance. Therefore, we expect the IPO underperformance puzzle to be a manifestation of the idiosyncratic risk puzzle, since explanations for the idiosyncratic risk puzzle also hold true in an IPO-only sample.

Although there is compelling evidence connecting the IPO's poor long-run performance to high idiosyncratic volatility, we are uncertain about the extent to which the two are connected, especially given the evidence from Edwards and Hanley (2010), who discount the validity of the explanation that short selling constraints are an underlying cause for the IPO underperformance. They argue that short selling is not as constrained as suggested by the prior literature and cannot be the reason for the short-term overpricing of IPO firms. Therefore, the empirical question of whether IPO long-run underperformance persists after controlling for idiosyncratic volatility arises.

If IPOs are more volatile than non-IPOs during a period of 5 years after issuance, we also expect to see greater representation of IPO firms among stocks with greater idiosyncratic risk and expect that the idiosyncratic risk puzzle should be more pronounced among IPOs than among non-IPOs, based on the explanation by Stambaugh, Yu, and Yuan (2014) that overpricing is more difficult to eliminate among higher volatility stocks. In addition, the average performance of IPO firms is known to be lower than that of non-IPO firms. Consequently, IPO firms' underperformance may contribute to the idiosyncratic risk puzzle. However, it is not clear whether the idiosyncratic risk puzzle still holds in the universe of non-IPO stocks, since proposed explanations for the idiosyncratic risk puzzle in theory are not exclusively related to IPO firms. For example, according to Stambaugh, Yu, and Yuan (2014) regarding the idiosyncratic risk puzzle, as long as there is a significant variation in the firm-specific volatility among non-IPO stocks, the asymmetry effect of arbitrage risk can be at work for non-IPO firms as well. Therefore, we investigate whether the idiosyncratic risk puzzle remains significant after controlling for the underperformance of IPOs.

The prediction of the importance of IPO firms in the idiosyncratic risk puzzle is consistent with the observation that the idiosyncratic volatility puzzle depends on the interval over which the idiosyncratic risk is estimated. The negative relation between the expected return and idiosyncratic volatility is obtained when volatility is estimated from the in-month daily return data. Malkiel and Xu (2002), however, document a positive relation between the expected return and the idiosyncratic risk estimated from monthly returns. We notice that the estimation of idiosyncratic volatility with monthly returns usually requires 24 to 60 monthly returns. The period of the first 60 months after IPO is essentially the period during which IPO underperforms. Thus, data restriction excludes IPOs from the analysis when their performance is poor. Therefore, we reason that IPO underperformance is likely to be responsible for inconsistency in the relation between expected returns and the idiosyncratic risk estimated using daily data or monthly data.

In our empirical analysis, we strictly follow the methodology of AHXZ (2006). We estimate the one-month-lag idiosyncratic volatility as the standard error of the regression of daily return on the Fama-French (1993) three-factor (FF3) model and group the stocks to idiosyncratic volatility (IVOL) quintiles. Our sample is based on the universe of CRSP common stocks. We find that IPO stocks are more volatile than non-IPO stocks on average, and as a result, there are more IPO stocks in the highest IVOL quintile. We also notice that, although there are more IPO stocks in the highest IVOL quintile than in the lowest IVOL quintile, there are substantial IPO observations in the lower IVOL quintiles as well. Next, we separate stocks into IPOs and non-IPOs in each IVOL quintile to test the difference in the idiosyncratic risk puzzle for IPOs and non-

IPOs. We examine the gross and four risk-adjusted returns of IPOs and the relative performance between IPOs and non-IPOs in each idiosyncratic risk quintile. We find that the performance of the IPOs in the highest IVOL quintile is much poorer than that in other IVOL quintiles. Furthermore, the performance of IPOs relative to non-IPOs varies across IVOL quintiles. In the lowest three IVOL quintiles, the IPOs do not significantly underperform non-IPOs, suggesting that the IPO underperformance puzzle is basically an idiosyncratic risk puzzle. However, the performance of the IPOs in the highest IVOL quintile is significantly lower than that for non-IPOs, confirming that the idiosyncratic risk puzzle is more significant for the IPOs. For non-IPOs firms, although the idiosyncratic risk puzzle is weaker, it remains significant and economically important, suggesting that IPO underperformance is only partially responsible for the idiosyncratic risk puzzle and that there might be other reasons for this phenomenon.

We conduct Fama-Macbeth regressions to investigate the relation between the two puzzles as well. We find that, in regressions to explain returns, coefficients for both the IPO variable and IVOL are negative and significant when they are included separately. However, the coefficient for the IPO variable becomes positive or insignificantly negative when it is included with IVOL, again suggesting that the idiosyncratic risk puzzle dominates the IPO puzzle, and the latter is only a manifestation of the former.

We also evaluate the robustness of our results. Since IPO firms are typically characterized as small, low-priced, and illiquid growth stocks, one might wonder if the results we observe are a consequence of such characteristics. To address this concern, we repeat our analysis on a subset of the sample after controlling for price, size, and illiquidity effects; we find that both the statistical significance and economic importance of our results remain unchanged. Moreover, we also consider the situations in which portfolio returns are equal-weighted and when NYSE stocks only are under consideration. At last, considering the large amount of studies on the hot IPO market, we divide the full sample into three different sub-periods according to the new issuance volumes for each month to investigate the two puzzles' relation under hot and cold IPO issuing periods.

The remainder of the paper is organized as follows. Section 2 describes data and the empirical methodology and presents some descriptive statistics. Section 3 reports the results from our empirical analysis, and Section 4 concludes the paper.

2. Data, Variables, Methodology and Descriptive Statistics

We obtain our data from four sources. Both the daily and monthly stock returns are collected from the Center for Research in Security Prices (CRSP). The daily and monthly Fama/French common factors are from the Kenneth R. French data library. The initial public offering information is obtained from SDC Platinum and Jay Ritter IPO data website. Following the methodology in Ang et al. (2006), we regress the daily stock returns on the Fama-French three factors (Fama and French, 1993) to estimate the in-month idiosyncratic volatility. We focus our analysis on the universe of CRSP stocks. For completeness, we later repeat all of our analysis on the subset of the stocks traded on the New York Stock Exchange (NYSE) as a robustness check. Given that not all IPO firms receive sufficient daily returns in the first month for reliable idiosyncratic volatility estimation, we calculate the idiosyncratic volatility for each month, starting with the second month after IPO, and analyze the relation between the returns and the idiosyncratic risk starting from the third month.

Following Ang et al. (2006), the in-month daily idiosyncratic volatility is estimated as a standard deviation of the residuals from the regression of the one-month daily excess returns on the FF3 factors. We require a minimum of fifteen returns for the idiosyncratic risk estimation. We define the monthly idiosyncratic volatility as the product of the daily idiosyncratic volatility and the square root of the number of daily returns for the stock in that month (Fu, 2009). In each month, we classify the stocks into five quintiles based on the idiosyncratic volatility estimated from the previous month as in Ang et al. (2006). We calculate both value-weighted (VW) returns and equally-weighted (EW) returns in each month for all idiosyncratic risk quintiles.

We obtain new public offering date information from the SDC Platinum and Jay Ritter IPO data website. In Loughran and Ritter (1995)[5], only IPOs with offering prices greater than $5 are included; ADRs, REITs, acquisition funds, closed-end funds, unit offers, small best efforts deals, and oil and gas limited partnerships are not considered either. Similar to the above treatments, we exclude observations with stock prices smaller than $5. Moreover, we also only consider the stocks with share code 10 and 11 ordinary common shares which have not been (or need not be) further defined. We then merge the IPO offering date with the CRSP universe data set. We define the IPO's age based on the number of months after the stocks go to public. Then, based on the empirical evidence on IPO underperformance as documented in Ritter (1991) and Loughran and Ritter (1995), we classify a stock as an IPO firm if it has been public for no more than 60 months. After a stock has gone to the public for more than 60 months, we consider it to be seasoned. In our later analysis, we use a dummy variable to indicate IPO, which takes a value of one if a firm is no more than 60 months old and zero otherwise. Finally, we notice that, from 01/1975 to 10/1982, the number of firms going to market has remained small; from11/1982, this number increases suddenly and remains large[6] till 12/2000 and then remains relatively small in the following years. Therefore, we start our analysis from 01/1983[7]. In our analysis, there are 9196 unique IPO stocks. We also remove observations of all other new listings that appear in CRSP for less than 60 months.

[5] https://site.warrington.ufl.edu/ritter/files/2015/05/Returns-on-IPOs-during-the-five-years-after-issuing-for-IPOs-from-1970-2012-2014-05-30.pdf

[6] Although during the period between 11/1982 and 12/2000 there are a few months with fewer public offerings, these months are discontinuous.

[7] Since the IPO data is available in Ritter's data from 1975 and the IPO data from SDC seems problematic before 1975, according to the definition of variable IPO, our analysis can start from 01/1980 as early as possible. Another reason that we start our analysis from 1983 is that, before 11/1982, the trading volume values in more than half observations are missing, affecting the sub-sample analysis of controlling for illiquidity and stock size. Moreover, by doing so, the overall period consists of a hot issuing period and a relatively cold issuing period.

Because IPO firms are often considered to be illiquid, we also calculate Amihud's (2002) illiquidity measure as follows.

$$Ill_{i,t} = \frac{1}{Dt}\sum_{nt=1}^{Dt}\frac{|ret_{i,nt}|}{prc_{i,nt}\times vol_{i,nt}} \quad (1)$$

where $Ill_{i,t}$ is the monthly illiquidity of stock i in month t; Dt is the number of observations in month t; $ret_{i,nt}$ is the hold period return of stock i in day nt; $prc_{i,nt}$is the price of the stock i in day nt; and $vol_{i,nt}$is the trading volume of stock i in day nt.

We examine the performance of each IVOL quintile by measuring gross returns and the alphas from the Capital Asset Pricing Model (CAPM), the FF3 model, the Carhart (1997) enhanced Fama-French four-factor (FF4) model, and the Fama-French five-factor (FF5) model. The alphas are estimated using the following model:

$$R_{i,t} - r_{f,t} = \alpha 0 + \sum b_j * Factor_{j,t} + \varepsilon_{i,t} \quad (2)$$

where $R_{i,t}$ refers to the average return for quintile i(i=1~5) on month t, $r_{f,t}$ is the monthly risk-free rate; for the CAPM model, the factor is the market excess return; for the FF3 model, the factors are market excess return, SMB (Small-Minus-Big), and HML (High-Minus-Low); for the FF4 model, we also include the MOM (Momentum) factor in addition to the FF3 factors; and for the FF5 model[8], RMW (Robust Minus Weak) and CMA (Conservative Minus Aggressive) are added to the FF3 factors. We define the intercepts from these regressions, α_0, as the alphas.

[8] In this new model proposed by Fama and French in 2015, another two factors, profitability (RMW) and investment (CMA) are added to the three factor model. Fama and French argue that although this model is rejected on the GRS test at the average return level, in practical applications, it offers "an acceptable description of average return". In our study, we include this new model and basically receive similar results with those at other levels of performance.

Table 1.1 reports the time series averages of sample characteristics. We find that the idiosyncratic volatility increases monotonically from 4.337% in IVOL quintile one to 21.941% in IVOL quintile five. The percentage of firms that can be considered as IPO increases from 8.3% in IVOL quintile one to about 36.8% in IVOL quintile five, which is consistent with our prediction of the greater representation of IPO firms in the higher IVOL quintiles. The average market capitalization decreases from $6,443.35 million for IVOL quintile one to that of $554.37 million in IVOL quintile five. Liquidity and price show a similar decrease as we move from the lowest IVOL quintile to the highest IVOL quintile. These characteristics are consistent with our expectation that stocks in the higher IVOL quintiles are small, less liquid, and lower priced.

The last two columns in Table 1.1 provide us with a first look of the idiosyncratic risk puzzle. Consistent with the results from Bali and Cakici (2008), we find the idiosyncratic volatility puzzle primarily in the value-weighted (VW) portfolio returns: the VW return for IVOL quintile five is a minimal 0.365% per month, compared with the 1.110% for IVOL quintile one. On the other hand, there is very little variation in the equally weighted (EW) returns among the different IVOL quintiles.

3. The Empirical Results

In this section, we report the results from our examination of the link between the underperformance of IPOs and the idiosyncratic risk puzzle. Our investigation proceeds as follows. Given our observation of the greater representations of IPO firms in the highest IVOL quintile, we first directly examine whether the IPOs are more volatile than the non-IPOs. Second, we investigate whether the IPO long-run underperformance is a manifestation of the idiosyncratic risk puzzle by comparing the performance between the IPO firms and non-IPO firms while controlling for idiosyncratic volatility. We then move on to examine whether the idiosyncratic risk puzzle exists among non-IPO firms. By comparing the difference in the idiosyncratic risk puzzle between IPO firms and non-IPO firms, we are able to estimate the contribution of IPO underperformance to the idiosyncratic risk puzzle. Finally, we use Fama-Macbeth regression to examine the relation between these two anomalies.

3.1. The Relation between IPOs and the Idiosyncratic Volatility

In this subsection, we directly investigate the relation between IPO stocks and idiosyncratic volatility by conducting two Fama-Macbeth regressions of the idiosyncratic volatility on the dummy variable IPO with or without the natural log of market value and natural log of Book-to-Market ratio. Panel A of Table 1.2 shows that the coefficient of IPO is a statistically significant 3.843, indicating that the idiosyncratic volatility for IPO stocks is 3.843% greater than that of 10.289% for the non-IPO ones. Moreover, even after controlling for size and book-to-market ratio effects, the volatility of IPO stocks is still significantly greater than that of non-IPOs.

The regression results in Panel A of Table 1.2 suggest that, on average, we should expect more IPO firms in the highest IVOL quintile, which is consistent with what we have observed in

Table 1.1. However, given that we are using a calendar time regression approach to investigate the idiosyncratic risk, we need to know that there are sufficient IPO observations in each portfolio. Panel B of Table 1.2 shows the time series average of the number of IPO firms in each IVOL quintile. The average number of IPOs increases from 59 in IVOL quintile one to that of 260 in IVOL quintile five. We notice that there are sufficient numbers of IPOs in different volatile levels. This is important because it ensures that our analysis in the next section is economically meaningful, in which we investigate the relative performance of IPO firms and non-IPO firms across different idiosyncratic volatility quintiles.

3.2. Is the IPO Underperformance Puzzle A Manifestation of the Idiosyncratic Risk Puzzle?

Given the evidence from the previous subsection that IPO firms are generally riskier than non-IPO firms, one might wonder whether IPO underperformance is simply a result of the poor performance of high idiosyncratic risk stocks. To investigate this possibility, we separate stocks in each IVOL quintile into IPO firms and non-IPO firms and compare the performance of IPOs and non-IPOs within each quintile.

In Table 1.3, we report the performance for the entire sample (Column 2), the IPO subsample (Column 3), the non-IPO subsample (Column 4), and the difference in returns between the IPO and non-IPO subsamples (Column 5) for each IVOL quintile as well as the differences in returns between IVOL quintile 1 and IVOL quintile 5.

Turning our attention to Panel A, we observe that the general decreasing trend of raw returns in IVOL quintiles for the entire sample is mostly replicated in both the IPO subsample and non-IPO subsample. For example, the average of raw returns for the IPO firms and non-IPO ones in IVOL quintile one are 1.297% and 1.107%, respectively, much greater than those of 0.031% of

IPO firms and 0.528% for non-IPO firms in quintile 5. Moreover, we see that IPO stocks do not always underperform the non-IPO stocks. In quintile 1, the gross return of IPO (1.297%) stocks is higher than that of non-IPO firms (1.107%). Similarly, the average return for the IPO firms is greater than that of the non-IPO firms in both IVOL quintile 2 and IVOL quintile 3. In other words, there is no sign of underperformance by the IPO firms in the first three IVOL quintiles. If anything, the raw returns for the IPO firms are greater than the non-IPO firms in those quintiles. Moving to IVOL quintile 4, we observe that the average return of IPO firms is 0.801%, 0.237% lower than that of the non-IPO firms (0.1.038%). In IVOL quintile 5, the IPO firms underperform the non-IPO counterparts more severely. The average of raw returns for the IPO sample is close to zero, much lower than the 0.528% for the non-IPO firms. Hence, the underperformance of IPOs is mainly from the abysmally low performance in the most volatile IPO stocks. We do not see a significant difference between the overall gross returns of IPOs and non-IPOs. Therefore, the underperformance of the most volatile IPOs is the reason for the overall poor average performance of IPOs.

The results so far are based on the raw returns and fail to consider the risk characteristics of stocks in different IVOL quintiles. Next, we move to different measures of alpha to control for various risk factors. As shown in Panel B, C, D, and E, the results based on CAPM, FF3, FF4, and FF5 models are similar in spirit to those based on the raw ret urns.

To summarize, we find no underperformance of the IPO firms when the idiosyncratic volatility is low. The underperformance is evident only when idiosyncratic volatility is high. Therefore, the evidence in this section suggests that the average underperformance of IPO stocks arises from the extremely poor performance of most volatile IPOs.

3.3. Is the Idiosyncratic Risk Puzzle a Manifestation of IPO underperformance?

Given the greater representation of IPO firms in the highest IVOL quintile and the long-run underperformance of the IPO firms, one might wonder whether the idiosyncratic risk puzzle is a manifestation of the IPO underperformance. A comparison of the idiosyncratic risk puzzle (last row in Table 1.3 among different columns gives us the answer to this question. As we can observe, the idiosyncratic risk puzzle is prevalent and significant in the entire sample, the IPO sample, and the non-IPO sample. For example, Panel A shows that the idiosyncratic risk puzzle is -0.579% for the non-IPO subsample. Therefore, we can conclude that the idiosyncratic risk puzzle is different from the IPO underperformance puzzle. It remains economically significant even after removing the IPO firms. The underperformance of IPO firms for the highest IVOL stocks, on the other hand, contributes to and magnifies the idiosyncratic risk puzzle for the entire sample, as reflected in the idiosyncratic risk puzzle, at -0.922% for the entire sample.

Similar patterns are observed in Panels B, C, and D in Table 1.3. All information indicates that, while the long-run underperformance of the IPO firms enhances the idiosyncratic risk puzzle for the entire sample, the idiosyncratic risk puzzle is different from the IPO underperformance puzzle. In other words, there are factors in addition to the long-run underperformance of IPO firms contributing to the extremely poor performance for most volatile IVOL firms.

3.4. An Alternative Regression Test

So far, we rely on panel data to examine the relation between the two puzzles. One limitation of such an approach is that it is difficult to simultaneously control for many different potentially confounding factors. In this section, we follow a method developed by Loughran and Ritter (1995) in which they use a Fama-Macbeth approach to control for the size and book-to-market effects in testing the underperformance property of IPO stocks. We extend their model to include additional variables to examine the relation between the IPO underperformance puzzle and the idiosyncratic risk puzzle. We carry out four regressions as follows:

$$r_{it} = \alpha_0 + \alpha_1 lnMV_{it} + \alpha_2 LN(\frac{BV}{MV})_{it} + \alpha_3 IPO_{it} + \varepsilon_{i,t} \quad (3)$$

$$r_{it} = \alpha_0 + \alpha_1 lnMV_{it} + \alpha_2 LN(\frac{BV}{MV})_{it} + \alpha_4 IVOL_{it} + \varepsilon_{i,t} \quad (4)$$

$$r_{it} = \alpha_0 + \alpha_1 lnMV_{it} + \alpha_2 LN(\frac{BV}{MV})_{it} + \alpha_3 IPO_{it} + \alpha_4 IVOL_{it} + \varepsilon_{i,t} \quad (5)$$

$$r_{it} = \alpha_0 + \alpha_1 lnMV_{it} + \alpha_2 LN(\frac{BV}{MV})_{it} + \alpha_3 IPO_{it} + \alpha_4 IVOL_{it} + \alpha_5 IPO_{it} \times IVOL_{it} + \varepsilon_{i,t} \quad (6)$$

where IPO_{it} is a dummy variable indicating whether or not firm i is considered as an IPO in month t, and $IVOL_{it}$ is the idiosyncratic volatility of firm i in month t. MV and BV/MV ratio are defined as in Fama and French (1992). The relation between the two puzzles will be reflected in the coefficients of IPO and IVOL and the interaction term $IPO_{it} \times IVOL_{it}$ in model 6. We report the regression results in Table 1.4.

Looking at columns 2 and 3 in Table 1.4, we find that the coefficient of IPO and IVOL are significantly negative, consistent with both Loughran and Ritter (1995) and AHXZ (2006). For example, the coefficient of IPO in model 1 is -0.316, very close to the value of -0.38 reported in Loughran and Ritter (1995). In model 4, when both IPO and IVOL as well as the interaction term between these two variables are included in the regression, the coefficient of IPO becomes insignificant, suggesting that there is no underperformance of IPOs once we control for idiosyncratic risk. However, the coefficient of IVOL remains significantly negative, indicating that the idiosyncratic risk puzzle dominates the IPO underperformance puzzle. Finally, the significantly negative estimate of the interaction term between IPO and IVOL shows that the underperformance of highly volatile IPOs will magnify the idiosyncratic risk puzzle.

3.5. Robustness

3.5.1. The Results of Equal-weighted Cases

Bali and Cakici (2008) show that the idiosyncratic risk puzzle under the equal-weighted situation is not as significant as that of the value-weighted case. In this session, we check whether the relation between the two puzzles under the equal weighted case display similar properties. The results are shown in Table 1.5. The trends of performance change versus volatility are different from those in Bali and Cakici (2008). We see a significant difference on any level of risk-adjustment between most stable IPO portfolio and non-IPO portfolio. The difference of the results for equal-weighted cases from that in Bali and Vakici (2008) arises after we remove the observation with a price smaller than \$5. Furthermore, all levels of risk-adjusted performance in the most volatile portfolio of IPO stocks is, on average, significantly lower than that of corresponding levels of non-IPO stocks. Additionally, except for the gross return level, the performance of IPOs is not significantly poorer than responding non-IPOs in most stable quintiles. At last, the idiosyncratic risk puzzle of IPO stocks is more serious in IPOs than in non-IPOs; for example, for the gross return level, the number is -0.260% per month. This difference is even higher and more significant at risk-adjusted performance levels. Therefore, the relation between these two puzzles in the equal-weighted case is basically consistent with that in the value-weighted case.

3.5.2. The Results When Considering the Effect of Small size, Illiquid

The IPO stocks are relatively small and illiquid. As in Table 1.1, stocks in quintile 5 have the highest illiquidity, the most IPOs, and the lowest price and market capitalizations. Therefore, we have enough reason to doubt whether the thinly traded small-sized and illiquid IPOs cause the underperformance of the IPOs. This is very important for financial practice. In this section, we consider these effects by removing the stocks with market capitalization smaller than 5% capitalization benchmarks or with an Amihud (2002) illiquidity measurement greater than 95% of the illiquidity benchmarks. After removing the thinly traded observations, the CRSP data set has 1,257,940 month-stock observations; the NYSE data set has 496,265 month-stock observations[9]. We duplicate the calculations in Tables 3 and 5 for the screened CRSP data set. The results are shown in Tables 1.6 and 1.7. We see the same pattern in Table 1.6 as those in Table 1.3. Just one point calls for notice: the idiosyncratic risk puzzles for all three data sets (whole, IPO, and non_IPO) are less serious than those before data screening for illiquid and small-sized stocks, whereas in Table 1.7, the results for the equal-weighted cases display a more significant idiosyncratic risk puzzle than those shown in Table 1.5 when there is no data screening. The relation between two puzzles is well consistent with that in Table 1.4. Therefore, although after screening the data sets to control for illiquid and small-sized effect, a lot of the most volatile IPOs are removed, and we obtain very similar results to those obtained before the thinly traded stocks are removed.

The results in this section show that the idiosyncratic risk puzzle is not equivalent to the IPO underperformance puzzle. Considering that the IPO underperformance is only a manifestation

[9] The results for NYSE data sets with data-screening are available upon request.

of idiosyncratic risk puzzle, we conclude that the idiosyncratic risk puzzle overwhelms the IPO underperformance puzzle even for the sizable, valuable, and liquid enough "normal" stocks.

3.5.3. The NYSE Sample

AHXZ (2006) uses the NYSE sample to test the interaction of the idiosyncratic volatility effect with firm size. We repeat all the analysis in Table 1.1 and 1.3 for NYSE data as well. The results are illustrated in Table 1.8. The NYSE data show a less serious idiosyncratic risk puzzle. Similarly, the difference of the underperformance of IPO stocks to non-IPO stocks between the most stable portfolio and the least stable portfolio is also less significant than in the CRSP universal data set. However, we see that the IPO stocks in the most stable portfolio don't significantly underperform the non-IPO counterpart; those in portfolio 5 do underperform the non-IPO counterparts significantly. Therefore, we conclude that the relation between the two puzzles found in the CRSP sample also hold for the NYSE sample.

3.5.4. The Sub-period Samples

Concerning the IPO stocks, there are three puzzling issues: underpricing, hot-issue market, and long-run underperformance. Hot issue markets are defined as "periods in which the average first month performance (or aftermarket performance) of new issues is abnormally high" (Ibbotson and Jaff, 1975). IPO underpricing is an interchangeable term with high first-day returns. Ritter (2002) and Lowry and Schwert (2002) argue that high IPO activity follows high underpricing. One opinion about the difference between a hot IPO market and a cold IPO market is that the IPO stocks issued during the hot-issue market IPO stocks tend to have poorer after-issuance performance, since many low-quality firms go to the public than during the cold market period

(Loughran and Ritter, 1995). Miller's theory connects the short-term high return and long-run underperformance together as well. Miller (2000) argues that the high divergence of opinion can explain why the long-run underperformance of IPO stocks are those issued with relative high underpricing at the time of issuance. If so, we expect that, during the hot-issue market, the relation between IPO stocks after issuing underperformance and the idiosyncratic risk puzzle displayed in this study will be strengthened during the hot-issue period.

In this section, we aim at testing the difference in the relation between two puzzles during hot and cold IPO markets. Conceptually, IPO hot-issue periods are periods when the number of new issuances is large (Ibbotson and Jaff, 1975, Ritter, 1984). We divide the whole period into two sub-periods. We notice that the number of new issuances is basically continuously large during the period from 11/1982 to 12/2000. Since we concentrate on the long-term underperformance of IPOs, we extend the first sub-period to 12/2002, when we assume that the issuances during the hot period represent most of the IPO observations. The second sub-period is from 01/2003 to 12/2014. We repeat the calculation in Table 3 on the two sub-samples. The results are shown in Tables 9.

We find some different features in the results in these two samples. First, during the hot IPO period, the IPOs underperform the non-IPOs on average at all level risk-adjusted performance; whereas during the cold period, the IPOs even have slightly better performance on average than non-IPOs, even though these differences are not significant. Second, at all risk-adjusted levels, the idiosyncratic risk puzzles during the hot periods are more serious than those during the cold period, whether in the whole sample, the IPO samples, or the non-IPO samples. This finding suggests that during the hot IPO period, the stocks are more volatile, and most volatile stocks perform even more poorly than in the cold IPO period for all types of stocks. Third, the difference between the idiosyncratic risk puzzles in IPOs and in non-IPOs during the hot issuing period is bigger than during the cold issuing period at all risk-adjusted levels except for the FF5 risk-adjusted performance. Forth, after controlling for the idiosyncratic volatility, the performance of IPO stocks is not poorer than that of non-IPO stocks in all cases. Moreover, the idiosyncratic risk puzzle after removing IPO observations remains significant, at a degree lower than that of the corresponding whole samples. Therefore, our sub-period results confirm our arguments about the relationship between two anomalies in the pooled CRSP universe data set. In addition, the comparison of the two periods: the hot issuing period and cold issuing period, enriches the literature of how the hot IPO issuing affects the stock performance.

4. Conclusions

In this paper, we investigate the relation between the idiosyncratic risk puzzle and the IPO long-run underperformance puzzle. Our empirical results suggest that the IPO underperformance puzzle is a manifestation of the idiosyncratic risk puzzle. First, the long-run underperformance of IPOs only exists in highly volatile IPOs. Second, the trend of the IPO performance across idiosyncratic volatility levels is similar to that of the whole data set. On the other hand, the inverse is not tenable. After removing the IPOs, the idiosyncratic risk puzzle remains significant for non-IPO observations.

We use an alternative method, the Fama-Macbeth regression, controlling for size and book-to-market ratio to test the above results. Our regression results confirm our conclusion. We also find that the above results do not attribute to the small and illiquid thinly-traded IPOs. Our other robustness tests divide the sample into subsamples according to the time period, and NYSE stock also basically confirms our conclusions.

In summary, the idiosyncratic risk puzzle dominates the IPO underperformance puzzle, and the IPO underperformance puzzle enhances the magnitude of the idiosyncratic risk puzzle.

Reference:

Aggarwal, R., Leal, R., and Hernandez, L. 1993. The aftermarket performance of initial public offerings in Latin America. Financial Management, 42-53.

Agarwal, S., Liu, C., and Rhee, S. G. 2008. Investor demand for IPOs and aftermarket performance: Evidence from the Hong Kong stock market. Journal of International Financial Markets, Institutions and Money, 18, 176-190.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Ang, A., Hodrick, J.R., Xing, Y. and Zhang, X., 2006. The cross-section of volatility and expected returns. The Journal of Finance 61, 259–299.

Ang, A., Hodrick, J.R., Xing, Y. and Zhang, X., 2009. High idiosyncratic volatility and low returns: international and further US evidence. Journal of Financial Economics 91, 1–23.

Bali, T.G., Cakici, N., 2008. Idiosyncratic volatility and the cross section of expected returns, Journal of Financial and Quantitative Analysis 43, 29–58.

Bali, T.G., Cakici, N., Whitelaw, R.F., 2011. Maxing out: stocks as lotteries and the cross-section of expected returns. Journal of Financial Economics 99, 427–446.

Boehme, R.D., Danielsen, B.R., Kumar, P., Sorescu, S.M., 2009. Idiosyncratic risk and the cross-section of stock returns: Merton (1987) meets Miller (1977). Journal of Financial Markets 12, 438-468.

Boyer, B., Mitton,T., Vorkink, K., 2010. Expected idiosyncratic skewness. Review of Financial Studies 23, 169–202.

Cao, J., Han, B., 2009. Idiosyncratic risk, costly arbitrage, and the cross-section of stock returns.17th Conference on the Theories and Practices of Securities and Financial Markets.

Carhart, M.M., 1997. On persistence in mutual fund performance, The Journal of finance 52, 57–82.

Chen, J., Hong, H., & Stein, J.C., 2002. Breadth of ownership and stock returns. Journal of financial Economics 66, 171-205.

Chen, C., Huang, A. G., &Jha, R., 2012. Idiosyncratic return volatility and the information quality underlying managerial discretion. Journal of Financial and Quantitative Analysis, 47, 873-899.

Chen, J., Chollete, L., Ray, R., 2010. Financial distress and idiosyncratic volatility: An empirical investigation. Journal of Financial Markets 13, 249–267.

Coakley, J., Hadass, L., Wood, A., 2008. Hot IPOs can damage your long-run wealth. Applied Financial Economics18, 1111-1120.

Dimovski, W., Brooks, R., 2004. Initial public offerings in Australia 1994 to 1999: recent evidence of underpricing and underperformance. Review of Quantitative Finance and Accounting 22, 179-198.

Doran, J., Jiang, D., Peterson, D., 2014. Short-sale constraints and the idiosyncratic volatility puzzle: an event study approach. Journal of Empirical Finance28, 36-59.

Duan, Y., Hu, G., McLean, R.D., 2010, Costly arbitrage and idiosyncratic risk: evidence from short sellers. Journal of Financial Intermediation19, 564-579.

Edwards, A.K., Hanley, K.W., 2010. Short selling in initial public offerings. Journal of Financial Economics98, 21-39.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. The Journal of Finance47, 427-465.

Fama, E.F., Kenneth R.F, 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics 33, 3–56.

Fama. E.F. and Kenneth R.F., 2015. A five-factor asset pricing model. Journal of Financial Economics, 116, 1-22.

Fan, S., and Yu, L. 2013. Accrual anomaly and idiosyncratic risk: international evidence. The International Journal of Business and Finance Research, 7, 63-75.

Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. Journal of Financial Economics 91, 24–37.

Gao, Y., Mao, C.X., Zhong, R., 2006. Divergence of opinion and long-term performance of initial public offerings. Journal of Financial Research29, 113-129.

George, T. J., and Hwang, C. Y. 2013. Analyst Coverage and Two Puzzles in the Cross Section of Returns. In AFA 2012 Chicago Meetings Paper.

Grinblatt, M. and Hwang, C.Y., 1989. Signaling and the pricing of new issues. *The Journal of Finance*, *44*(2), 393-420.

Guo, H., and Qiu, B. 2014. Options-implied variance and future stock returns. Journal of Banking & Finance 44, 93-113.

Han, Y., and Lesmond, D., 2011. Liquidity biases and the pricing of cross-sectional idiosyncratic volatility. Review of Financial Studies 24, 1590–1629.

Helwege, J., Liang, N., 2004. Initial public offerings in hot and cold markets. Journal of Financial and Quantitative Analysis 39, 541-569.

Hou, K., & Loh, R. K. 2016. Have we solved the idiosyncratic volatility puzzle?. *Journal of Financial Economics*, *121* 167-194.

Houge, T., Loughran,T., Yan, X., Suchanek, G., 2001. Divergence of opinion, uncertainty, and the quality of initial public offerings. Financial Management.30, 5–23.

Huang, L, Liu, Q., Rhee, S.G., Zhang, L., 2010. Return reversals, idiosyncratic risk, and expected returns. Review of Financial Studies 23, 147–168.

Ibbotson, R.G., 1975. Price performance of common stock new issues. Journal of financial economics 2, 235–272.

Levis, M., and Thomas, D.C., 1995. Investment trust IPOs: Issuing behaviour and price performance Evidence from the London Stock Exchange. Journal of Banking & Finance, 19(8), 1437-1458.

Li, X., Sullivan, R., 2011. The limits to arbitrage revisited: The accrual and asset growth anomalies. Financial Analysts Journal74, 50-66.

Loughran, T., Ritter, J.R. 1995. The new issues puzzle. The Journal of Finance 50, 23–51.

Lowry, M., and Schwert, G. W. (2002). IPO market cycles: Bubbles or sequential learning?. The Journal of Finance, 57(3), 1171-1200.

Malkiel, B.G., Xu, Y., 2002. Idiosyncratic Risk and Security Returns. University of Texas at Dallas. Working Paper.

Mashruwala, C., Rajgopal, S., Shevlin, T., 200. Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. Journal of Accounting and Economics42, 3-33.

Merton, R.C., 1987.A simple model of capital market equilibrium with incomplete information. The Journal of Finance 42, 483–510.

Miller, E.M., 1977. Risk, uncertainty, and divergence of opinion. The Journal of Finance32, 1151-1168.

Miller, E. M., 2000. Long run underperformance of initial public offerings: an explanation. Department of Economics and Finance Working Papers, 1991-2006.

Mitchell, M., Pulvino, T., Stafford, E., 2002. Limited arbitrage in equity markets. The Journal of Finance57, 551-584.

Newey, W.K., and West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.Journal of the Econometric Society, 703–708.

Ofek, E., Richardson, M., Whitelaw, R.F., 2004. Limited arbitrage and short sales restrictions: Evidence from the options markets. Journal of Financial Economics74, 305-342.

Oesterle, D.A. 2006. High Cost of IPOs Depresses Venture Capital in the United States. Entrepreneurial Bus. LJ1, 369.

Pontiff, J., 1996. Costly arbitrage: Evidence from closed-end funds. Quarterly Journal of Economics111, 1135–1152.

Pontiff, J. 2006, Costly arbitrage and the myth of idiosyncratic risk. Journal of Accounting and Economics42, 35-52.

Rajgopal, S., Venkatachalam, M., 2011. Financial reporting quality and idiosyncratic return volatility. Journal of Accounting and Economics 51, 1-20.

Ritter, J.R., 1984. The "hot issue" market of 1980. *Journal of Business*, 215-240.

Ritter, J.R., 1991.The long-run performance of initial public offerings. The Journal of Finance 46, 3–27.

Ritter, J. R., and Welch, I. 2002. A review of IPO activity, pricing, and allocations. The Journal of Finance, 57, 1795-1828.

Stambaugh, R.F., Yu, J., Yuan, Y., 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. The Journal of Finance 70, 1903–1948.

Schultz, P., 2003. Pseudo Market Timing and the Long-Run Underperformance of IPOs. The Journal of Finance 68, 483-517.

Shleifer, A., Vishny, R. W., 1997. The limits of arbitrage. The Journal of Finance52, 35-55.

Teoh, S. H., Welch, I., Wong, T.J., 1998. Earnings management and the long-run market performance of initial public offerings. The Journal of Finance53, 1935-1974.

Wurgler, J., Zhuravskaya, E., 2002, Does Arbitrage Flatten Demand Curves for Stocks?. The Journal of Business75, 583-608.

Figures and Tables

Table 1.1: Descriptive statistics of volatility quintiles for CRSP data set

The means of some key variables for each idiosyncratic volatility quintile are presented. Monthly quintile portfolios are constructed by sorting the stocks according to the idiosyncratic volatility relative to the Fama-French (1993) three factor model. The CRSP data is sorted by idiosyncratic volatility breakpoints. The stock monthly idiosyncratic volatilities are the production of the standard deviation of the regression error of the in-month daily returns on Fama-French (1993) three common factors and the square root of the number of the observations in that month. The firm-months when there are less than 15 stocks are deleted while the idiosyncratic risk is calculated. Monthly idiosyncratic volatilities are computed using the data over the previous month. Quintile 1 (5) is the portfolio of stocks with the lowest (highest) idiosyncratic volatility. Monthly stocks are defined as IPO when the firms have been issued for less than 60 months.

"IVOL" is the idiosyncratic volatility; IPO indicates that the stocks are IPOs or non-IPOs; "ME" represents for market capitalization; "Illiquidity" is calculated from the daily return data with the model in Amihud (2002). EW-return is the average of equal-weighted return and VW-return is the value-weighted return average.

IVOL Quintiles	IVOL (%)	IPO	ME (Million $)	Illiquidity (10^{-5} $\$^{-1}$)	Price ($)	EW-return (%)	VW-return (%)
1 (Low)	4.337	0.083	6443.350	0.609	36.255	1.230	1.110
2	7.076	0.136	3349.310	0.518	29.269	1.263	1.020
3	9.640	0.211	1790.540	0.564	23.415	1.267	0.921
4	13.043	0.289	1023.830	0.676	18.981	1.101	0.972
5(High)	21.941	0.368	554.370	1.279	14.903	0.485	0.365

Table 1.2: The relation between IPO and idiosyncratic volatility

Panel A: The Relationship between IPOs and the Idiosyncratic Risk

This table reports the Fama-Macbeth (1973) regression coefficients of the explanatory variables: IPO, lnMV (natural log of stock value) and lnBV/MV (natural log of book-to-value ratio) for the explained variable idiosyncratic volatility of the CRSP ordinary data set. The simple Fama-Macbeth regression model is

$$R_{it} = \alpha 0 + \alpha_1 * IVOL_t + \varepsilon_{it} \quad (1)$$

$$R_{it} = \alpha 0 + \alpha_1 * \ln MV_{it} + \alpha_2 * \ln(BV / MV)_{it} + \alpha_4 * IVOL_t + \varepsilon_{it} \quad (2)$$

Newey-West (1987) adjusted t-statistics are reported in parentheses. ***, **, and * represent significance at the1%, 5%, and 10% level, respectively.

Model	Intercept	IPO	lnMV	lnBV/MV
1	10.289***	3.843***		
	[41.28]	[19.51]		
2	21.523***	0.830***	-2.038***	0.021
	[75.81]	[13.02]	[-58.87]	[0.08]

Panel B: The numbers of IPO stocks in each volatility quintile

This table shows the frequency of IPO stocks in each volatility quintile sorted according to IVOL breakpoints.

IVOL quintiles	1(Low)	2	3	4	5(High)
Numbers of IPOs	59	96	149	205	260

Table 1.3: The portfolio performance of CRSP whole data set, IPO and non-IPO subset under value-weighted situations and without data screening

The table reports the performance of whole CRSP data set, IPOs and non-IPOs subsets. All the data sets are grouped according to the whole CRSP stock IVOL breakpoints. We calculate the average gross returns for each portfolios in each month and then calculated the means of the portfolio return average over the periods.

One group models used to calculate the alphas are as following:

$$R_{i,t} - r_{f,t} = \alpha 0 + \sum b_j * control_{j,t} + \varepsilon_{i,t} \quad (1)$$

Another group models we use is as following:

$$Diff_t = \alpha_0 + \sum b_j * control_{j,t} + \varepsilon_t \quad (2)$$

where $Diff_t$ refers respectively to the difference of the portfolio gross returns between IPO portfolios and non-IPO portfolios on month t when calculating the performance difference between IPO and non-IPOs in each quintile, that between the highest idiosyncratic volatility portfolio and the lowest idiosyncratic volatility portfolio on month t when calculating the performance difference between portfolio 5 and portfolio 1, and that difference between the portfolio 5 and 1 gross return difference of IPO and non-IPOs. We don't use control variables for gross return difference; for CAPM model, the control variable is the market premium; for FF3 model, the control variables are market premium, SMB (Small-Minus-Big factor) and HML (High-Minus-Low factor); for Carhart (Carhart,1997) four-factor model, besides the FF3 factors, the fourth control variable is MOM (Momentum); for Fama-French five factor model, besides FF3 factors, the other two controls variables are RMW (Robust Minus Weak) and CMA (Conservative Minus Aggressive).[10]

We report α_0 in both model 1 and model 2.

Newey-West (1987) adjusted t-statistics are reported in parentheses. ***, **, and * represent significance at the1%, 5%, and 10% level, respectively.

IVOL quintiles	Whole	IPO	NonIPO	IPO-NonIPO
Panel A: Gross returns				
1(Low)	1.202***	1.297***	1.107***	0.190
2	1.200***	1.395***	1.005***	0.390***
3	1.063***	1.215***	0.910***	0.305
4	0.920***	0.801*	1.038***	-0.237
5(High)	0.279	0.031	0.528	-0.497*
Overall	0.958***	0.893***	1.023***	-0.130
5-1	-0.922***[-2.77]	-1.266***[-2.76]	-0.579**[-2.02]	-0.687**[-2.1]

[10] We also use the following model to calculate the difference between the performance of portfolio 5 and portfolio 1, or the difference between IPO and non-IPO, or the difference of the portfolio 5 and 1 difference between IPO and non-IPO: $R_{i,t} - r_{f,t} = \alpha 0 + \alpha * group5 + \sum b_j * control_{j,t} + \sum c_j * control_{j,t} * group5 + \varepsilon_{i,t}$. We report α. We obtain similar results as model 2.

Panel B: CAPM Alpha				
1(Low)	0.331***	0.415**	0.247***	0.168
2	0.157*	0.310**	0.004	0.306**
3	-0.096	0.005	-0.198**	0.203
4	-0.361**	-0.528**	-0.195	-0.333*
5(High)	-1.123***	-1.483***	-0.763***	-0.720***
Overall	-0.151*	-0.348**	0.045*	-0.393**
5-1	-1.454***[-5.12]	-1.899***[-4.55]	-1.010***[-4.27]	-0.889***[-2.66]
Panel C: FF3 Alpha				
1(Low)	0.283***	0.390**	0.177***	0.213
2	0.135	0.320**	-0.049	0.369**
3	-0.028	0.150	-0.206**	0.357**
4	-0.171	-0.241	-0.101	-0.140
5(High)	-0.887***	-1.090***	-0.685***	-0.405*
Overall	-0.041	-0.102	0.021	-0.123
5-1	-1.171***[-5.17]	-1.480***[-4.2]	-0.862***[-4.21]	-0.619**[-2.08]
Panel D: FF4 Alpha				
1(Low)	0.227**	0.336*	0.118*	0.218
2	0.082	0.227	-0.063	0.290*
3	-0.032	0.066	-0.131	0.197
4	-0.132	-0.238	-0.027	-0.211
5(High)	-0.728***	-0.978***	-0.479***	-0.499**
Overall	-0.059	-0.141	0.023	-0.165
5-1	-0.955***[-4.13]	-1.314***[-3.56]	-0.597***[-2.94]	-0.717**[-2.24]
Panel E: FF5 Alpha				
1(Low)	0.271***	0.392**	0.150**	0.243
2	0.092	0.235	-0.050	0.285*
3	0.003	0.198	-0.192**	0.390**
4	-0.143	-0.189	-0.098	-0.091
5(High)	-0.842***	-0.981***	-0.703***	-0.278
Overall	-0.015	-0.047	0.016	-0.063
5-1	-1.113***[-4.6]	-1.374***[-3.78]	-0.853***[-3.71]	-0.521*[-1.79]

Table 1.4: Average parameter values from monthly cross-sectional regressions of stock returns on market value, book-to-market ratio, IPO dummy and IVOL for CRSP data

The table shows the Fama-Macbeth regression coefficients of the stock returns on the natural logs of stock market value (lnMV), book-to-market ratio (lnBV/MV), IPO dummy and IVOL. The definition of MV and BV/MV follows Fama and French (1992). The data screening follows Ritter (1995). Three models are processed:

$$R_{it} = \alpha 0 + \alpha_1 * \ln MV_{it} + \alpha_2 * \ln(BV / MV)_{it} + \alpha_3 * IPO_t + \varepsilon_{it} \quad (1)$$

$$R_{it} = \alpha 0 + \alpha_1 * \ln MV_{it} + \alpha_2 * \ln(BV / MV)_{it} + \alpha_4 * IVOL_t + \varepsilon_{it} \quad (2)$$

$$R_{it} = \alpha 0 + \alpha_1 * \ln MV_{it} + \alpha_2 * \ln(BV / MV)_{it} + \alpha_3 * IPO_t + \alpha_4 * IVOL_{it} + \varepsilon_{i,t} \quad (3)$$

$$R_{it} = \alpha 0 + \alpha_1 * \ln MV_{it} + \alpha_2 * \ln(BV / MV)_{it} + \alpha_3 * IPO_t + \alpha_4 * IVOL_{it} + \alpha_5 * (IPO * IVOL)_{it} + \varepsilon_{i,t} \quad (4)$$

***, **, and * represent significance at the1%, 5%, and 10% level, respectively.

	Model 1	**Model 2**	**Model 3**	**Model 4**
Intercept	1.414*** [5.08]	2.207*** [9.26]	2.218*** [9.24]	2.116*** [8.57]
IPO	-0.3159*** [-2.57]		-0.191* [-1.90]	0.197 [1.33]
IVOL		-0.053*** [-5.76]	-0.051*** [-5.86]	-0.044*** [-4.90]
IPO*IVOL				-0.028 [-3.32]
lnMV	-0.025 [-0.79]	-0.084** [-2.64]	-0.086*** [-2.68]	-0.081*** [-2.52]
lnBV/MV	0.079* [1.93]	0.059 [1.60]	0.049 [1.43]	0.051 [1.51]
Average adjusted R^2	0.018	0.024	0.028	0.029

Table 1.5: The portfolio performance of CRSP whole data set, IPO and non-IPO subsets under equal-weighted situation and without data screening

The models in this table are the same as in Table 1.3. However, the equal weighted portfolio returns are considered. Newey-West (1987) adjusted t-statistics are reported in parentheses. ***, **, and * represent significance at the1%, 5%, and 10% level, respectively

IVOL quintiles	Whole	IPO	NonIPO	IPO-NonIPO
Panel A: Gross returns				
1(Low)	1.147***	1.039***	1.256***	-0.216*
2	1.245***	1.206***	1.284***	-0.078
3	1.224***	1.127***	1.321***	-0.194*
4	1.050***	0.914**	1.186***	-0.272*
5(High)	0.449	0.211	0.687*	-0.477***
Overall	0.972***	0.769**	1.176***	-0.407**
5-1	-0.698***[-2.42]	-0.829**[-2.25]	-0.568**[-2.28]	-0.260[-1.22]
Panel B: CAPM Alpha				
1(Low)	0.367***	0.275	0.459***	-0.185
2	0.263**	0.184	0.342***	-0.158
3	0.155	0.008	0.301**	-0.293***
4	-0.128	-0.339*	0.082	-0.421***
5(High)	-0.813***	-1.149***	-0.478***	-0.671***
Overall	-0.119	-0.428**	0.190	-0.618***
5-1	-1.180***[-4.95]	-1.424***[-4.64]	-0.937***[-4.46]	-0.487**[-2.31]
Panel C: FF3 Alpha				
1(Low)	0.238**	0.197	0.278***	-0.082
2	0.122	0.085	0.160**	-0.076
3	0.077	-0.012	0.167***	-0.179*
4	-0.098	-0.225**	0.028	-0.253**
5(High)	-0.681***	-0.904***	-0.458***	-0.446***
Overall	-0.124**	-0.309***	0.060	-0.369***
5-1	-0.918***[-5.26]	-1.100***[-4.29]	-0.736***[-5.03]	-0.364*[-1.8]
Panel D: FF4 Alpha				
1(Low)	0.220**	0.175	0.266***	-0.091
2	0.129	0.070	0.188***	-0.118
3	0.115	0.018	0.212***	-0.194**
4	-0.049	-0.189*	0.091	-0.280***
5(High)	-0.563***	-0.776***	-0.350***	-0.426***
Overall	-0.075	-0.256***	0.107**	-0.363***
5-1	-0.783***[-4.31]	-0.951***[-3.53]	-0.616***[-4.32]	-0.335[-1.59]
Panel E: FF5 Alpha				
1(Low)	0.243***	0.232	0.254***	-0.022
2	0.151*	0.150	0.151**	-0.001
3	0.121	0.069	0.173***	-0.104
4	-0.028	-0.122	0.066	-0.188*
5(High)	-0.590***	-0.759***	-0.420***	-0.339***
Overall	-0.068	-0.202***	0.066	-0.268***
5-1	-0.833***[-4.92]	-0.991***[-4.03]	-0.675***[-4.64]	-0.317[-1.61]

Table 1.6: The portfolio performance of CRSP whole data set, IPO and non-IPO subsets under value-weighted situation and with data screening

When screening the data, we remove the stocks with the price smaller than $5/share, or with market capitalization smaller than 5% capitalization benchmarks, or with the Amihud (2002) illiquidity measurement bigger than 95% illiquidity benchmarks. The models in this table are the same as in Table 1.3. However, the equal weighted portfolio returns are considered. Newey-West (1987) adjusted t-statistics are reported in parentheses. ***, **, and * represent significance at the1%, 5%, and 10% level, respectively.

IVOL quintiles	Whole	IPO	NonIPO	IPO-NonIPO
Panel A: Gross returns				
1(Low)	1.213***	1.316***	1.111***	0.205
2	1.180***	1.382***	0.978***	0.404***
3	1.053***	1.167***	0.940***	0.227
4	0.942***	0.842*	1.042***	-0.200
5(High)	0.317	0.095	0.540	-0.446
Overall	0.961***	0.898***	1.023***	-0.125
5-1	-0.896***[-2.69]	-1.222***[-2.7]	-0.571**[-2]	-0.651**[-2.06]
Panel B: CAPM Alpha				
1(Low)	0.348***	0.443***	0.253***	0.190
2	0.139	0.300**	-0.022	0.322**
3	-0.103	-0.043	-0.163*	0.120
4	-0.339**	-0.486**	-0.193	-0.293
5(High)	-1.082***	-1.414***	-0.750***	-0.664***
Overall	-0.149*	-0.344*	0.045*	-0.389**
5-1	-1.430***[-5.03]	-1.857***[-4.52]	-1.003***[-4.2]	-0.854***[-2.64]
Panel C: FF3 Alpha				
1(Low)	0.303***	0.423***	0.183***	0.240
2	0.118	0.311**	-0.075	0.386***
3	-0.038	0.089	-0.164*	0.254
4	-0.149	-0.189	-0.109	-0.081
5(High)	-0.837***	-1.016***	-0.658***	-0.358
Overall	-0.038	-0.097	0.021	-0.118
5-1	-1.140***[-5.08]	-1.439***[-4.18]	-0.841***[-4.02]	-0.598**[-2.07]
Panel D: FF4 Alpha				
1(Low)	0.238**	0.352*	0.124*	0.228
2	0.073	0.231	-0.085	0.316*
3	-0.035	0.011	-0.080	0.090
4	-0.135	-0.225	-0.046	-0.179
5(High)	-0.680***	-0.893***	-0.467***	-0.425*
Overall	-0.057	-0.137	0.023	-0.160
5-1	-0.918***[-4.04]	-1.245***[-3.45]	-0.591***[-2.87]	-0.654**[-2.11]
Panel E: FF5 Alpha				
1(Low)	0.288***	0.422***	0.155**	0.267*
2	0.082	0.240	-0.076	0.315*
3	-0.016	0.119	-0.151*	0.270
4	-0.121	-0.129	-0.112	-0.017
5(High)	-0.798***	-0.913***	-0.684***	-0.229
Overall	-0.013	-0.042	0.016	-0.059
5-1	-1.086***[-4.56]	-1.334***[-3.77]	-0.838***[-3.61]	-0.496*[-1.77]

Table 1.7: The portfolio performance of CRSP whole data set, IPO and non-IPO subsets under equal-weighted situation and with data screening

When screening the data, we remove the stocks with the price smaller than $5/share, or with market capitalization smaller than 5% capitalization benchmarks, or with the Amihud (2002) illiquidity measurement bigger than 95% illiquidity benchmarks. The models in this table are the same as in Table 1.3. However, the equal weighted portfolio returns are considered.
Newey-West (1987) adjusted t-statistics are reported in parentheses. ***, **, and * represent significance at the1%, 5%, and 10% level, respectively.

IVOL quintiles	Whole	IPO	NonIPO	IPO-NonIPO
Panel A: Gross returns				
1(Low)	1.187***	1.102***	1.272***	-0.171
2	1.264***	1.250***	1.278***	-0.029
3	1.221***	1.147***	1.296***	-0.149
4	1.100***	0.998***	1.202***	-0.203
5(High)	0.499	0.275	0.722*	-0.447***
Overall	0.999***	0.814**	1.185***	-0.371**
5-1	-0.688**[-2.27]	-0.827**[-2.22]	-0.550**[-2.01]	-0.277[-1.35]
Panel B: CAPM Alpha				
1(Low)	0.395***	0.322*	0.468***	-0.146
2	0.273**	0.217	0.329**	-0.112
3	0.138	0.021	0.255*	-0.234**
4	-0.100	-0.269	0.069	-0.338**
5(High)	-0.812***	-1.107***	-0.517***	-0.591***
Overall	-0.113	-0.403**	0.177	-0.579***
5-1	-1.207***[-4.9]	-1.429***[-4.65]	-0.985***[-4.39]	-0.444**[-2.11]
Panel C: FF3 Alpha				
1(Low)	0.269***	0.251	0.286***	-0.035
2	0.141	0.131	0.150*	-0.019
3	0.064	0.006	0.122*	-0.117
4	-0.059	-0.144	0.025	-0.169
5(High)	-0.668***	-0.847***	-0.488***	-0.359***
Overall	-0.112**	-0.271***	0.047	-0.318***
5-1	-0.936***[-5.18]	-1.098***[-4.21]	-0.774***[-5.05]	-0.324[-1.61]
Panel D: FF4 Alpha				
1(Low)	0.243**	0.215	0.271***	-0.056
2	0.148*	0.119	0.177**	-0.059
3	0.104	0.037	0.170***	-0.133
4	-0.012	-0.111	0.087	-0.197
5(High)	-0.550***	-0.726***	-0.373***	-0.354***
Overall	-0.063	-0.221***	0.095**	-0.316***
5-1	-0.793***[-4.22]	-0.941***[-3.44]	-0.644***[-4.28]	-0.298[-1.44]
Panel E: FF5 Alpha				
1(Low)	0.274***	0.288*	0.259***	0.029
2	0.163*	0.191	0.135*	0.057
3	0.106	0.085	0.127**	-0.042
4	0.002	-0.046	0.051	-0.097
5(High)	-0.582***	-0.702***	-0.463***	-0.239**
Overall	-0.060	-0.166**	0.046	-0.212***
5-1	-0.856***[-4.83]	-0.990***[-3.97]	-0.722***[-4.65]	-0.268[-1.39]

Table 1.8: Results of the NYSE data set

Panel A shows the descriptive statistics of volatility quintiles for NYSE data sample. The NYSE data is sorted by NYSE volatility breakpoints. The definitions of variables are same as in table 1.

Panel B-E are the portfolio performance of NYSE whole data set, IPO and non-IPO sub data– value-weighted cases without data screening similar to table 3 for CRSP data set. The models in this table are the same as in Table 1.3. Newey-West (1987) adjusted t-statistics are reported in parentheses. ***, **, and * represent significance at the1%, 5%, and 10% level, respectively.

Panel A: Descriptive statistics of volatility quintiles for CRSP data set

IVOL Quintiles	IVOL (%)	IPO	ME (Million $)	Illiquidity (10^{-5} $\$^{-1}$)	Price ($)	EW-return (%)	VW-return (%)
1 (Low)	3.896	0.039	10116.940	0.062	43.556	1.149	1.091
2	5.730	0.070	6922.220	0.065	38.718	1.232	1.054
3	7.373	0.107	4620.030	0.093	33.114	1.191	0.904
4	9.568	0.151	2934.560	0.112	27.285	1.279	1.015
5(High)	15.501	0.196	1701.930	0.164	20.776	0.869	0.791

IVOL quintiles	Whole	IPO	NonIPO	IPO-NonIPO
Panel B: Gross returns				
1(Low)	1.084***	1.079***	1.090***	0.001
2	1.200***	1.354***	1.046***	0.309*
3	1.116***	1.342***	0.890***	0.452***
4	1.094***	1.187***	1.000***	0.187
5(High)	0.687**	0.522	0.852**	-0.331
Overall	1.007***	1.015***	1.000***	0.015
5-1	-0.407[-1.57]	-0.576[-1.64]	-0.238[-0.89]	-0.336[-1.22]
Panel C: CAPM Alpha				
1(Low)	0.249**	0.220	0.278***	-0.055
2	0.270**	0.427**	0.113	0.314*
3	0.070	0.261	-0.121	0.381**
4	-0.026	0.046	-0.097	0.144
5(High)	-0.538***	-0.699***	-0.377**	-0.322
Overall	-0.004	-0.070	0.062	-0.132
5-1	-0.787***[-3.54]	-0.921***[-2.75]	-0.654***[-3.18]	-0.271[-0.94]
Panel D: FF3 Alpha				
1(Low)	0.158	0.121	0.196***	-0.073
2	0.164	0.326*	0.003	0.323*
3	-0.036	0.190	-0.261***	0.451***
4	-0.067	0.074	-0.208**	0.282
5(High)	-0.576***	-0.704***	-0.448***	-0.256
Overall	-0.045	-0.067	-0.022	-0.045
5-1	-0.735***[-3.41]	-0.826**[-2.38]	-0.645***[-3.45]	-0.186[-0.64]
Panel E: FF4 Alpha				
1(Low)	0.091	0.041	0.141**	-0.098
2	0.130	0.294	-0.033	0.327
3	-0.038	0.174	-0.249***	0.422***

4	-0.045	-0.009	-0.082	0.074
5(High)	-0.435***	-0.584**	-0.286**	-0.298
Overall	-0.062	-0.104	-0.020	-0.085
5-1	-0.526***[-2.44]	-0.626*[-1.78]	-0.427**[-2.31]	-0.202[-0.67]
Panel F: FF5 Alpha				
1(Low)	0.148	0.114	0.183***	-0.066
2	0.090	0.196	-0.016	0.212
3	-0.088	0.076	-0.251***	0.326
4	-0.067	0.122	-0.256**	0.378*
5(High)	-0.569***	-0.713***	-0.426***	-0.287
Overall	-0.073	-0.108	-0.039	-0.069
5-1	-0.718***[-3.22]	-0.828**[-2.33]	-0.609***[-3.13]	-0.224[-0.76]

Table 1.9: The results of CRSP whole data set, IPO and non-IPO subsets under value-weighted situation during two sub periods: 01/1983~12/2002 and 01/2003~12/2014

The models in this table are the same as in Table 1.3. Panel A-E are the portfolio performance of CRSP whole data set, IPO and non-IPO sub data–value-weighted cases without data screening similar to table 4 for CRSP data set. In each panel, the results of two different sub-periods: 01/1983~12/2002 and 01/2003~12/2014 are presented. Newey-West (1987) adjusted t-statistics are reported in parentheses. ***, **, and * represent significance at the1%, 5%, and 10% level, respectively.

Panel A: Gross Returns

IVOL quintiles	Whole	IPO	NonIPO	IPO-NonIPO
01/1983~12/2002				
1(Low)	1.307***	1.379***	1.235***	0.143
2	1.181***	1.304***	1.058***	0.247
3	1.015***	1.063**	0.966***	0.097
4	0.780	0.584	0.975**	-0.390
5(High)	-0.044	-0.346	0.259	-0.605
Overall	0.903***	0.709	1.098***	-0.389
5-1	-1.350***[-3.01]	-1.725***[-2.81]	-0.976***[-2.87]	-0.749*[-1.7]
01/2003~12/2014				
1(Low)	1.024***	1.159***	0.889***	0.270
2	1.231***	1.547***	0.915**	0.632**
3	1.144**	1.473**	0.815	0.658**
4	1.157**	1.168*	1.146*	0.022
5(High)	0.826	0.669	0.982	-0.313
Overall	1.051***	1.205**	0.897**	0.308
5-1	-0.198[-0.43]	-0.490[-0.77]	0.093[0.19]	-0.583[-1.25]

Panel B: CAPM Alpha

01/1983~12/2002				
1(Low)	0.377***	0.442*	0.313***	0.129
2	0.102	0.177	0.028	0.149
3	-0.167	-0.172	-0.162	-0.010
4	-0.537**	-0.804***	-0.270	-0.533**
5(High)	-1.451***	-1.888***	-1.015***	-0.873***
Overall	-0.258***	-0.591***	0.075**	-0.666***
5-1	-1.829***[-4.69]	-2.330***[-4.13]	-1.327***[-4.23]	-1.002**[-2.31]
01/2003~12/2014				
1(Low)	0.249	0.364	0.134*	0.230
2	0.244	0.545*	-0.056	0.601*
3	0.020	0.315	-0.274**	0.588*
4	-0.034	0.003	-0.071	0.074
5(High)	-0.551*	-0.742*	-0.361	-0.380
Overall	0.053	0.116	-0.010	0.126
5-1	-0.801**[-2.28]	-1.106**[-2.12]	-0.496[-1.66]	-0.610[-1.26]

Panel C: FF3 Alpha

IVOL quintiles	Whole	IPO	NonIPO	IPO-NonIPO
01/1983~12/2002				
1(Low)	0.221**	0.291	0.151*	0.140
2	0.022	0.114	-0.070	0.184
3	-0.044	0.062	-0.149	0.211
4	-0.188	-0.312	-0.063	-0.249
5(High)	-0.980***	-1.146***	-0.814***	-0.332
Overall	-0.087	-0.202	0.028	-0.230
5-1	-1.201***[-4.31]	-1.437***[-3.56]	-0.965***[-3.57]	-0.472*[-1.69]
01/2003~12/2014				
1(Low)	0.252	0.357	0.146**	0.211
2	0.248*	0.552*	-0.056	0.608*
3	-0.008	0.280	-0.295***	0.575**
4	-0.070	-0.048	-0.093	0.045
5(High)	-0.612**	-0.820**	-0.404*	-0.417
Overall	0.039	0.086	-0.008	0.094
5-1	-0.864***[-2.74]	-1.178**[-2.38]	-0.550*[-1.92]	-0.628[-1.31]

Panel D FF4 Alphas

IVOL quintiles	Whole	IPO	NonIPO	IPO-NonIPO
01/1983~12/2002				
1(Low)	0.155	0.243	0.066	0.177
2	-0.063	-0.056	-0.071	0.015
3	-0.035	-0.087	0.017	-0.104
4	-0.147	-0.361*	0.067	-0.427*
5(High)	-0.748***	-0.964***	-0.533***	-0.431
Overall	-0.120	-0.285**	0.044	-0.329**
5-1	-0.903***[-3.11]	-1.207***[-2.71]	-0.599**[-2.34]	-0.608*[-1.94]
01/2003~12/2014				
1(Low)	0.240	0.343	0.137**	0.206
2	0.239	0.543*	-0.065	0.607*
3	-0.015	0.266	-0.295***	0.561**
4	-0.061	-0.034	-0.088	0.054
5(High)	-0.584**	-0.811**	-0.357	-0.454
Overall	0.036	0.082	-0.010	0.092
5-1	-0.824***[-2.64]	-1.154**[-2.32]	-0.494*[-1.71]	-0.660[-1.37]

Panel E: FF5 Alpha

IVOL quintiles	Whole	IPO	NonIPO	IPO-NonIPO
01/1983~12/2002				
1(Low)	0.177	0.231	0.123	0.108
2	-0.044	-0.017	-0.072	0.055
3	-0.006	0.113	-0.125	0.238
4	-0.172	-0.275	-0.069	-0.206
5(High)	-0.978***	-1.072***	-0.885***	-0.186
Overall	-0.072	-0.167	0.024	-0.191
5-1	-1.155***[-3.85]	-1.302***[-3.06]	-1.008***[-3.34]	-0.295[-0.94]
01/2003~12/2014				
1(Low)	0.300*	0.487*	0.112*	0.375
2	0.245*	0.536*	-0.047	0.583*
3	0.008	0.315	-0.299***	0.614**
4	-0.009	0.042	-0.060	0.102
5(High)	-0.463**	-0.608*	-0.317	-0.291
Overall	0.085	0.185	-0.014	0.199
5-1	-0.763***[-2.5]	-1.096**[-2.29]	-0.430[-1.62]	-0.666[-1.33]

ESSAY2 THE CAPITALIZATION OF SCHOOL QUALITY INTO HOUSE PRICES WHEN QUALITY IS UNCERTAIN: THEORY AND EMPIRICAL EVIDENCE

1. Introduction

Most economists and policy advocates believe that parents care about the quality of their children's education. While there is a large and growing empirical literature estimating the relationship between school quality and housing prices, these studies thus far overlook the modeling issues arising when quality indicators are imperfect or quality itself is uncertain to the market participants. School quality is a multidimensional concept that is difficult for parents or objective observers to quantify. As a consequence, observable measures of quality are noisy signals of underlying quality-and market participants understand this. Therefore, even if empirical analysts' measures of school quality closely match parents' quality measures, test scores or state-administered school performance grades provide at best noisy signals of underlying school quality. This paper develops an uncertainty framework to examine the extent to which housing markets capitalize school quality and quality uncertainty into prices in the spatial market.

This paper makes several contributions to the literature. It makes a theoretical contribution, extending the theory of urban household behavior to address how uncertainty in location amenities like public school quality affects housing and location demands. There is a significant theoretical literature introducing housing and location in the urban area (Turnbull 1995). All of these frameworks are similar in that housing or non-housing consumption are uncertain, differing from the household's ex ante consumption plans. In the case of school quality, however, different realizations of quality outcomes do not introduce uncertainty in housing and non-housing consumption per se, but affect ex post utility through a third channel-education consumption-which is influenced solely by the household's choice of location (i.e., which schools to attend). Thus, the structure of this model dealing with location-specific uncertainty like school quality uncertainty differs from earlier frameworks and leads to more clear-cut capitalization predictions. The theory shows that greater school quality increases housing prices and steepens the price gradient. At the same time, though quality risk decrease housing prices and flatten the price gradient. Our results confirm that the bit-rent property does extend to the uncertainty environment like public school quality uncertainty.

School quality uncertainty arises from two different sources. First, there may be variability in school performance over time or uncertainty about the information content of official school signals like exam scores or governmentally provided school ratings. We label this school quality volatility. Second, some houses in the school district are reassigned periodically from one school attendance zone to another as attendance boundaries are adjusted over time to accommodate changes in the spatial distribution of school age children as family's age or move about in the metropolitan area. We label this attendance zone instability. Frequent school zone change can both increase the school quality uncertainty and the neighborhood environment instability. Both types of school quality uncertainties and the neighborhood environment uncertainty resulting from the latter reduce house prices.

The second contribution is empirical. This paper provides the first estimates of how uncertain school quality risk and attendance zone uncertainty ach affect housing prices and the extent to which they modify capitalization estimates of expected quality into housing prices. We use geocoded house sales data in Orange County, Florida, over the period 2001-2012 to develop empirically relevant measures of school quality volatility and attendance zone instability. The data provide both full sample estimates and as well as estimates based on the boundary matched sample approach to separate school quality and quality risk from local neighborhood effects. The sample period covers the expansionary phase before June 2007 and the subsequent market crash and initial recovery, which allows us to examine how market conditions influence performance and risk capitalization estimates.

Our empirical results are consistent with the conclusions from our theory. For both the pooled and boundary samples spanning the full period, when we consider the school average score only, its estimates are significant and positive; when the school related risk variables are added into the model, the estimates for school average scores decrease slightly. However, when we add the school peer effect variable the percentage of students enjoying free lunches into the model, the estimates of school score drop significantly; when further adding student/teacher ratio, we receive insignificant school score estimates. In all the above models, estimates for two school risk measures are significantly negative. We also find that the school peer effect variable, on average, displays strong negative effects on house prices over the full period. The results from the two sub-periods help to understand results for the full period samples. In the rising market, school average score and score quality volatility display a strong effect on house pricing; whereas the school peer effect variable the percentage of students enjoying free lunch and input variable student/teacher ratio only show a small effect. However, in the falling market, the estimates of school score and school quality volatility are insignificant, whereas the peer effect variable displays extremely strong negative effect - the estimates are about ten times of those in the rising market. Zabel (2015) shows a significant negative estimate for school test score in the downturn market during 2006-2012 in Boston MSA whereas school scores display positive effect before this crisis. Our results show that during the turbulently falling market period, the buyers still care more about the school quality; however, the parents are concerned more with the student composition rather than the school scores. Our results also consistent with some researchers' findings that foreclosure and ROE neighbors negatively affect the price of normal houses[11]. Low income families tend to hold subprime mortgage and be more likely experience foreclosure during the falling period, therefore it is not surprising that the buyers pay more attention to the percentage of students enjoying free lunches. We also show that in higher income neighborhoods, the buyers are more concerned with the percentage of students having free lunch, especially during house price falling period. We don't receive significant estimates for the variable of school zone change in almost all regressions, showing that a one-time school zone change won't affect house price significantly. That is because those parcels were assigned to stable new school areas.

The paper is organized as follows. Section 2 presents the existing evidence of school quality capitalization; section 3 is the theoretical derivation; section 4 shows empirical work; section 5 presents our main conclusions.

[11] See the work ofKobie, and Lee, 2010,Daneshvary and Clauretie, 2012, Ihlanfeldt, and Mayock, (2013)

2. Existing Evidence

Nguyen-Hoang and Yinger (2011) provide a recent review of the literature on local public school capitalization. The premise that greater public school quality leads to higher housing prices is confirmed by most of the studies they survey. According to their Table 2.2, Nguyen-Hoang and Yinger (2011) report that 34 out of 50 papers obtain completely positive results. Although the hedonic studies basically confirm a positive relation between the house price and public education service, the results are not always consistent across studies and not always stable within studies.

Early studies tended to use input-based measures of education quality, like perpupil spending. Oates (1969) finds that, per pupil expenditures are reflected in house price. And some later studies, such as Brasington (1999), Black (1999), Downes and Zabel (2002), Barrow and Rause (2004), Brasington and Haurin (2006), similarly find significant positive relationship between per-pupil expenditures and house price. Crone (2006), however, does not find significant effect of district level per-pupil expenditure on house value while Mathur (2008) concludes that higher per-pupil expenditures only increase the value of high quality houses.

The tie between input-based variables and resultant education quality is intuitive, but imprecise. Spending on inputs does not measure the quality of outcomes (Rosen and Fullerton 1977) even before the era of state grants-in-aid designed to equalize school spending across districts. Two highly cited papers by Hanushek published in 1996 and 2007 show that school inputs have no apparent impact on student achievement. Later studies turned away from input measurement and now tend to use output or student performance measures to indicate school quality. Most rely on test scores, state government schools grades (typically based on standardized test scores), and value added test scores. Of course, there is much variation in these quality measures across studies, reflecting factors specific to different states and locales. Generally though, the school grades (based on test scores) are usually letter variables or other discrete performance indicators. Value added test scores are the difference of test scores during two continuous years for the same student cohort and can be school level or district level. Brasington (1999) offers a different measure of value added as the difference between individual school test scores and state average scores. Using this rationale, all normalized scores can be alternatively interpreted as measures of value added.

The coefficients of the school outputs variables are not always significantly positive either. Brasington (1999) tests various school quality measurements, including the math, reading, science, writing passing rate in 4th , 8th, and 12th grades and all the corresponding value added scores, for six Ohio metropolitan areas using both tradition hedonic regression and spatial correlation models. His results are mixed, although he finds significant estimates for most of the models for his 4th grade math test measure. Black (1999) is often credited with popularizing the boundary sample approach. She uses the three year average of fourth grade relative math and reading scores the in Massachusetts Educational Assessment Program (MEAP) system as the measure of school quality. In traditional hedonic regression, the coefficient of test score shows that a 5% increase in test scores increases house prices by about 4.9%, all else being equal. She focuses on houses near attendance zone boundaries in order to distinguish the school quality premium from unobserved neighborhood attribute effects and finds test score coefficients that are much smaller than in the pooled sample. Bogart and Cromwell (2000) find that third grade reading scores are negatively related to house price in Shaker Heights, Ohio, over 1983-1994. They conclude that the test scores are mainly serving as a proxy for unobservable heterogeneity among schools and neighborhoods and not school quality.

Kane et al. (2003) use mean scaled math and reading composite scores to measure of elementary school quality. They find that only the long run average test scores over time affect the house price significantly. Their short-term measure of school test performance, the latest test score, does not significantly affect house price. Apparently real-estate markets ignore short run variations in test scores but respond to long run levels. They also conclude that some of the effect of elementary schools appears to be due to middle school and high school assignments.

Figlio and Lucas (2004) use the house price data from the Gainesville, Florida, metropolitan area to examine the effects of state assigned grades for school performance. They include the elementary school average reading test score, the average math test scores, and the state assigned letter grade in their models. The average math test score and the state assigned letter grade are significantly positive in the price equation. The coefficients for average reading test scores, however, are insignificant. Brasington and Haurin (2006) test whether parents value test scores or score-based value added. They find both expenditures and test scores significantly affect house price; at the same time, the value added of schooling is not capitalized into house value.

Zahirovic-Herbert and Turnbull (2009) examine how housing markets capitalize school quality information. They consider both the school performance score and the school performance improvement. The results show that the former is not systematically capitalized into house price; however, the latter significantly increases the house price. In a different vein, Zahirovic-Herbert and Turnbull (2008) argue that the housing market is a search market in which capitalization can occur along two margins, price and liquidity. Their results support this notion, showing that school quality is capitalized into house liquidity or difficulty of sale when price capitalization is weak.

Yinger (2010) includes a wide variety of school quality measures in the hedonic price regressions: the relative elementary test scores, which are the individual school test scores minus the district average scores, the high school passing rate, the district elementary value added, and the quadratic value added. The coefficients of the elementary relative scores, the high school passing rate and the elementary value added are significantly positive while the others are not. Chiodo, et al (2011) also use test scores, using the fourth grade math school-level index generated by the Missouri Department of Elementary and Secondary Education, which is calculated from school test scores. In a recent paper, Bogin and Phuong (2014) study the effect of the No Child Left Behind (NCLB) policy on house prices for schools designated as failing. They find that even after controlling the school quality, houses assigned to failing NCLB schools have significantly lower prices. They include the 3th, 4th, 5th grade math and reading tests passing percentages in their models and obtain mixed results for these performance measures.

The broad expectation is that housing markets capitalize public school quality into house values. This review illustrates that studies use a variety of performance or quality measures. It is not clear the extent to which the variety of results can be attributed to differences in school quality measures. It is possible that the empirical measures sometimes do not reflect the overall quality of the schools or, as suggested by Kane et al (2003), parents ignore short run variation in measured outcomes. In contrast with this perspective, the next section shows that households do not ignore short run variation in school performance. Short run variation in measured outcomes affect house prices when market participants recognize that short run variation in test scores or other quality measures arise because they are noisy signals of underlying school quality. In this case, short run variation in performance provides some insight into the noise level in the quality signal.

3. School Quality, Risk, and House Prices

This section extends urban household theory to the setting with uncertain school quality. Previous studies examine household behavior in a variety of risky settings; see Turnbull (1995) for a survey of seminal papers adapting multidimensional expected utility analysis to the urban spatial environment. Ex post household utility $u(x, y, q)$ is a strictly concave function of housing x, nonhousing y, and school quality q, where the concavity reflects the assumption that the household is risk averse. The household chooses x, y before the school quality q is realized. The expected quality, however, is distributed with mean $E[q] = \mu > 0$ and finite variance $V(q) > 0$. We deconstruct realized school quality into the expected and risky components, $q = \mu + \sigma\varepsilon$, where σ a risk is shift parameter ($\sigma = 1$ initially) and ε is stochastic, distributed $E[\varepsilon] = 0$, $V(\varepsilon) > 0$. Clearly, $V(q) = \sigma^2 V(\varepsilon)$ so that $d\sigma > 0$ yields an increase in mean-preserving-spread: $dE[q]/d\sigma = 0$ and $dV(q)/d\sigma = 2\sigma V(\varepsilon) > 0$.

In order to derive housing price capitalization effects, we must first derive the underlying nonspatial Hicksian demands as the solution to the following problem at each given location (indexed by t, suppressed for the time being):

$$\min_{x,y}\{Px + y\}\, s.t.\ \ E[u(x, y, \mu + \sigma\varepsilon)] = U$$

The relevant first order conditions can be expressed as the uncertainty counterpart to the usual consumer marginal conditions

$$MRS_{x,y} = P$$

$$U - E[\mu(x, y, \mu + \sigma\varepsilon)] = 0$$

where $MRS_{x,y} = E[u_x]/E[u_y]$ is the marginal rate of substitution between housing and other goods, the absolute slope of the expected utility indifference curve U^1 depicted in figure 2.1. It turns out that indifference curve maps in $x - y$space for expected utility levels do not have neoclassical properties when either x or y are stochastic (Turnbull 1991, 1994, 1995). Neither x nor y are stochastic in this model, however, so well behaved indifference curves do exist for the uncertainty environment considered here. As a consequence, these marginal conditions have the usual interpretation, requiring that the solution, the Hicksian demands $\{x(P,U,\mu,\sigma), y(P,U,\mu,\sigma)\}$, occurs where the iso-expenditure line aa is tangent to the expected utility constraint at α in figure

The appendix shows that the usual properties of Hicksian demand functions hold for these Hicksian demands under education quality uncertainty, including the substitution theorem (law of demand), $\partial x/\partial P < 0$, and housing demand monotonic increasing in expected utility, $\partial x/\partial U > 0$, when housing is a normal good. In addition, the appendix shows that totally differentiating the first order conditions and solving for the comparative static results in the usual way reveal housing demand is a function of expected education quality μ and quality risk, as reflected by the mean-preserving-spread parameter σ, satisfying the following

$$\frac{\partial x}{\partial \mu} = -\left(\frac{\partial x}{\partial U}\right)E[u_q] - \frac{E[u_y]}{|\mathrm{B}_2|}\left(\frac{dMRS}{d\mu}\right) \tag{1}$$

$$\frac{\partial x}{\partial \sigma} = -\left(\frac{\partial x}{\partial U}\right)E[u_q\varepsilon] - \frac{E[u_x]E\lfloor u_{yq}\varepsilon\rfloor - E[u_y]E\lfloor u_{xq}\varepsilon\rfloor}{|\mathrm{B}_2|} \tag{2}$$

where $|\boldsymbol{B}_2|< 0$ is the second order bordered Hessian determinant, which is negative for concave utility (i.e., multidimensional risk aversion).

The first result (1) shows that the total effect of greater mean education quality on Hicksian housing demand comprises two separate effects. The first right hand side term is unambiguously negative since $\partial x / \partial U > 0$. Intuitively, this term captures the fact that greater expected school quality increases expected utility, which reduces the housing consumption required to attain the given level of expected utility. It is through this channel that greater expected school quality reduces the Hicksian demand for housing, leading to $\partial x / \partial U < 0$. At the same time, however, the second term in (1) shows that greater school quality may alter the shape of the indifference map in x, y space, which generates an additional effect on housing demand. This housing taste effect takes the same sign as $dMRS / d\mu$. If greater school quality does not affect the strength of the household's preference for housing relative to non-housing consumption then $dMRS / d\mu = 0$ and an increase in expected school quality shifts the indifference curve for the given expected utility level from U^1 to the solid curve U^2 in Figure 2.1. (It is not necessary that U^1 be a radial expansion of U^2, i.e., constructed such that the slopes of the two curves are the same along any ray from the origin.) In this case, the second term, the housing taste effect, is zero in (1) and the total effect of expected school quality on housing demand is unambiguously negative. Figure 2.1 illustrates this case by the movement from α to β. When greater school quality weakens the household's preference for housing relative to non-housing consumption, $dMRS / d\mu < 0$ and the second term in (1) is negative, reinforcing the first term and leading to an unambiguous negative effect of expected school quality on Hicksian housing demand. If depicted in a graph like Figure 2.1, greater school quality would make the indifference curve for the given expected utility level shallower, which by itself decreases Hicksian housing demand. In the case in which $dMRS / d\mu > 0$, greater school quality strengthens the household's taste for housing relative to non-housing consumption, steepening the indifference curve for the given level of expected utility so that the indifference curve with greater expected school quality shifts to a curve like the dashed curve U^3 in Figure 2.1. The household's Hicksian equilibrium is now at point δ in the figure and the movement from β to δ reflect the second term in (1). Clearly, in this case the taste effect of school quality offsets the direct expected utility effect, leaving the total effect of expected quality on Hicksian housing demand (1) ambiguous. If expected school quality has a strong enough taste effect on $MRS_{x,y}$ then it may be possible that the Hicksian demand for housing increases with school quality. But the taste effect in this case would have to be quite strong, and this outcome is less likely the stronger the normality of housing demand in the traditional sense (which leads to a larger $\partial x / \partial U$ term). Summarizing, then, we have the following: $dMRS / d\mu \leq 0$ is sufficient for $\partial x / \partial U < 0$. And even if $dMRS / d\mu > 0$ as long as it is sufficiently small or $\partial x / \partial U$ sufficiently large then $\partial x / \partial U < 0$ as well.

The second comparative static result for Hicksian housing demand (2) shows that the effect of school quality risk also has two separate effects on Hicksian housing demand. As shown in the appendix, the first term on the right hand side is positive when housing is normal; greater school quality risk increases Hicksian housing demand through this channel. Intuitively, greater school quality risk reduces expected utility, which requires that the household consume more housing (and other goods) to attain the original level of expected utility. The appendix also shows that the second term on the right hand side of (2) is unambiguously positive the effect of school quality on the taste for housing does not increase with school quality, an economically intuitive restriction that parallels diminishing MRS in the $x - y$ plane. However interpreted, this is sufficient for increasing school quality risk to increase the Hicksian demand for housing, $\partial x / \partial U > 0$.

As a final step before deriving the house price function properties under uncertainty, the expenditure function under school quality uncertainty is defined in the usual way using the Hicksian demands $\{x(P,U,\mu,\sigma), y(P,U,\mu,\sigma)\}$ derived above.

$$e(P,U,\mu,\sigma) \equiv Px(P,U,\mu,\sigma) + y(P,U,\mu,\sigma) \tag{3}$$

It is straightforward to show that the standard certainty expenditure function properties hold in this context, including increasing in housing price (from Shepherd.s lemma), $(\partial e/\partial P) = x(P,U,\mu,\sigma) > 0$, and increasing in expected utility, $(\partial e/\partial U) = \lambda > 0$, The additional properties related to school quality also hold, as shown in the appendix.

$$\frac{\partial e}{\partial \mu} = -\frac{E[u_q]}{\lambda} < 0 \tag{4}$$

$$\frac{\partial e}{\partial \sigma} = -\frac{E[u_q \varepsilon]}{\lambda} > 0 \tag{5}$$

where the sign of the second result follows from the application of similar-dissimilar orderings when $u_{qq} < 0$ under risk aversion. Greater average education quality increases expected utility, which means that the household requires less spending on housing and other goods to attain a given expected utility level (4). Similarly, a greater mean-reserving-spread lowers expected utility under risk aversion, which means that the household requires more spending on housing and other goods to compensate for the loss in expected utility (5).

Having established the essential expenditure function properties under education quality uncertainty, we use the function to derive the spatial housing price function as follows. In location equilibrium, housing price at each location t satisfies the condition

$$e(P, U, \mu, \sigma) = m - c(m, t) \tag{6}$$

where m is household money income and c(m, t) commuting cost, with the standard properties $c_t > 0$, $c_{tt} \leq 0$, $0 < c_m < 1$. The housing bid price function is $P(t, U, \mu, \sigma, m)$ the implicit solution to the above condition. Implicit differentiation yields the usual properties of the housing price function as declining and convex in distance,

$$\frac{\partial P}{\partial t} = -\frac{c_t}{x} < 0 \tag{7}$$

$$\frac{\partial^2 P}{\partial t^2} = -\frac{x c_{tt} - c_t (\partial x / \partial P)/(\partial P / \partial t)}{x^2} > 0 \tag{8}$$

Turning to school quality effects, the following new results pertain to expected school quality and change in school quality risk. Implicitly differentiating (6) and directly differentiating (7) yield the expected school quality effects on the housing price function level and slope as, respectively,

$$\frac{\partial P}{\partial \mu} = -\frac{e_\mu}{x} > 0 \tag{9}$$

$$\frac{\partial}{\partial \mu}\left(\frac{\partial P}{\partial t}\right) = \left(\frac{\partial x}{\partial \mu} + \frac{\partial x}{\partial P}\frac{\partial P}{\partial \mu}\right)\frac{c_t}{x^2} < 0 \tag{10}$$

An increase in expected school quality increases housing price at all locations and steepens the housing price function slope (9), as illustrated by the shift from aa to bb in Figure 2.2. An increase in school quality risk, as an increase in mean-preserving-spread of the distribution of school quality outcomes, lowers housing price and makes the housing price function shallower (10), as in the shift from bb to aa in Figure 2.2. The school quality risk effects on the price function level and slope are similarly derived as

$$\frac{\partial P}{\partial \sigma} = -\frac{e_\sigma}{x} < 0 \tag{11}$$

$$\frac{\partial}{\partial \sigma}\left(\frac{\partial P}{\partial t}\right) = \left(\frac{\partial x}{\partial \sigma} + \frac{\partial x}{\partial P}\frac{\partial P}{\partial \sigma}\right)\frac{c_t}{x^2} > 0 \tag{12}$$

The signs of the price function shift results are unambiguous, following directly from the expenditure function properties (4) and (5). The price function slope result (10) holds under the assumption that education quality and housing are not strong complements in demand in the sense described earlier, a reasonable restriction ensuring $\partial x / \partial \mu < 0$. The price function slope result (12) holds under the non-increasing $dMRS / d\mu$ assumption, which is sufficient for $\partial x / \partial \sigma > 0$.

These last four results establish the main empirical implications of the theory: greater expected school quality increases house price while steepening the price gradient for this household type; school quality risk, however, reduces house price while fattening the price gradient. The expected school quality price shift result (9) demonstrates the standard result under certainty that has been widely examined in the school quality capitalization literature. The price function slope result (10) is new and demonstrates that even the simplest monocentric urban model with a single household type predicts weaker school quality capitalization in the urban periphery than in interior locations. Brasington (2002) argues that a greater supply of vacant parcels within the urban area nearer the periphery leads to weaker price capitalization for sites farther from the CBD; our results show that this pattern holds regardless of the vacant land supply. The new risk results (5) and (12) demonstrate that spatially invariant risk also leads to systematic spatial capitalization effects in the static neoclassical land use model without vacant land. In any case, the risk results are opposite the expected quality results, an intuitively appealing prediction.

Most of these results are new. Unfortunately, we cannot fully test the broader spatial implications of the theory here, since doing so requires data drawn from multiple school districts across the entire urban area. The remainder of the paper does, however, test the capitalization predictions associated with the housing price function shifts identified above.

4. The Empirical Analysis

4.1 The Data

House price and characteristics data are drawn from the sales of the Orange County Property Appraiser (OCPA) records of properties and sales occurring from August, 2001, through August, 2012. We focus on arm's length sales of single family detached houses with school attendance zones (by year) and census tracts and blocks. We clean the OCPA data to delete sales with missing values, selling prices of $1000 or less (which indicate administrative transactions), quit claims and other transactions not designated arm's length, and consecutive transactions occurring within 6 months that exhibit unusual prices. Finally we trim the sample by deleting the observations that lie outside 3 standard deviations in a robust regression on the base model without school related variables. The census tracts and year combinations that have fewer than 20 sales are removed as well. The resultant complete pooled sample covers 127,120 separate transactions; the elementary school boundary sample, explained below, covers 36,607 separate transactions.

The dependent variable is the natural log of house selling price (PRICE). House characteristics include house age (AGE), heated area in 1000 ft2 (AREA), net area, which equals to total area minus heated area, in 1000 ft2 (NET AREA), number of bedroom (BEDS), number of bathrooms (BATHS), and an index of building quality (CONDITION) provided by the OCPA.

We obtain the school test grades from the Florida Education Department and school attendance zone maps for each year from the Orange County Public school System (OCPS). We use several measures of school quality. The output based measure is based on annual school performance in the state standardized math test. We use the percentage of students who perform at level 3 or higher in math (level 3 is a state defined benchmark). Florida did not fundamentally change the method for grading schools over the sample period, but it did change the test in the last two years of our sample period. Most schools experienced large declines in the math test scores under the new test. Therefore, in order to control these effects as well as other more subtle changes in the tests or testing procedures over time, we normalize each school's math test performance measure using the school math grade divided by the average grade for all the schools in the Orange County Public School District for each year. Our final school score (TEST SCORE) in the regressions is the average of the school's normalized scores across the entire sample period. Performance risk or quality volatility (VOLATILITY) for each school is measured by standard deviation of the normalized test performance across the entire sample period. When calculating school quality volatility for new schools, we require that the new school operates long enough to report at least 3 annual test grades through 2012.

Some houses experience only one school attendance zone change. It turns out that these cases reflect the assignment of households to newly constructed schools in otherwise stable zones. Because these houses are not in unstable school zones per se, we separate houses that experience one school zone change during the sample period from those that experience multiple school zone changes. We set the attendance zone instability variable (INSTABILITY) equal to 0.1 when the number of school attendance zone changes for the subject house during the twelve year sample period exceeds one and zero otherwise. The CHANGE dummy variable is equal to 0.1 if there is a change of school zone and zero otherwise. (We use 0.1 instead of 1 simply to rescale reported coefficient estimates for INSTABILITY and CHANGE.) When used in conjunction with INSTABILITY, the variable CHANGE indicates a house assigned to a newly constructed school. Table 2.1 lists the new schools added during 2000-2012. The table shows that 25 new elementary schools were opened over the 12 years, an addition of more than 25% to the number operating in 2000. Not surprising, the new schools are built near the outskirts of the Orlando metro area in Orange County. To better distinguish the effects adding new schools from their geographic location, we include the distance from the Orlando central business district (DISTANCE) in the model.

The school enrollment information and student/teacher ratio are from the National Center for Education Statistics. We include the percentage of students enjoying free lunch (FREE LUNCH) as school peer effect variable and student/teacher ratio (S/T RATIO) as an input-based quality measure in several models.

Neighborhood socioeconomic data are from the Census. Subject houses are matched to census blocks to construct neighborhood demographic variables. Observations up through 2005 use the 2000 Census data; observations after 2005 use the 2010 Census data. We use the census neighborhood variable median household income (INCOME) to build stratify sub-samples by income level in some of the analysis. In that analysis, high income neighborhoods are defined as census block groups with median household income greater than the average of block groups in the sample; low income neighborhoods are defined as block groups with median household income less than the block group average.

In addition to the variable measuring distance to CBD, we include census tract fixed effects as more refined location controls. The empirical models use two samples. The first sample is simply the pooled sample of all single family house transactions within Orange County. The second sample follows what has become a popular approach in school quality studies by focusing on houses that lie on both sides of a school attendance zone boundary. This approach approximates a matched-sample approach to controlling for localized neighborhood effects not fully captured by the census tract fixed effect in the models (Black 1999). We construct the boundary sample for elementary schools, including houses up to 0.15 mile from the elementary zone boundary. The resultant elementary boundary sample covers 36,607 observations.

Table 2.2 reports the relevant variables and their description. Table 2.3 presents summary statistics of the main variables for the pooled and boundary sample.

4.2 The Empirical Analysis

Our empirical model is as following:

$$\ln(P_i) = \sum_{h=1}^{6} \gamma_h H_{ih} + \sum_{c=1}^{q_m} \vartheta_c SH_{ci} + \sum_{k=1}^{n} \tau_k TR_{ki} + \sum_{t=1}^{12} \rho_t T_{ti} + \varepsilon_i \qquad (13)$$

where P_i is price of sale i; *SH* is the school related variable vectors in model *m* (*m*=1,2,3,4); *TR* is the tract fixed effects vector; *T* is the year fixed effects vector; q_m is the number of school related variables in model *m*; and *H* is the house characteristics vector; ε_i is the error item.

Table 2.4 presents the regression results for the pooled sample spanning the full period. The coefficients on house characteristics, gross property tax and distance to CDB exhibit no surprising price effects and are stable across models.

Looking at the school quality variables, Model 1 includes TEST SCORE as the sole quality related variable. This coefficient estimate is 0.2495 and is significant. Model 2 adds the three school quality risk variables discussed earlier, VOLATILITY, CHANGE and INSTABILITY. Interestingly, the TEST SCORE coefficient estimate declines but remains significant when these risk variables are included. VOLATILITY and INSTABILITY coefficients are significantly negative but the CHANGE coefficient is not significant. Model 3 includes the peer effect FREE

LUNCH, which reduces the school quality effect by more than two thirds, although it remains significant. In contrast, including the input based measure of school quality, S/T RATIO, in models 4 yields an insignificant TEST SCORE coefficient. These additional variables, FREE LUNCH and S/T RATIO, have significant negative price effects. The peer effect and input based quality measure appear to fully capture variation in school quality across the sample to the exclusion of the output based test score variable.

Drawing together the results in Table 2.4, school quality, as measured by peer effects and inputs, has a positive effect on house price while quality risk from performance volatility and attendance zone instability tend to reduce house price. The CHANGE coefficient estimate is insignificant in all cases, indicating that parents are not very sensitive to a one time school zone change, mainly because these changes are in otherwise stable zones and reflect assignment to newly constructed schools.

Table 2.5 results for the boundary sample over the entire period. Recall that, by focusing on the sales that lie on both sides of a school attendance zone boundaries, this approach control for any localized neighborhood effects not fully captured by the census tract fixed effects in the pooled sample. Table 2.5 and subsequent tables report only key estimates. Table 2.5 reports similar patterns as found for the pooled sample in Table 2.4. Except for INSTABILITY, estimates for other school related variables are generally smaller than in the pooled sample. This is consistent Black's (1999) argument that the boundary sample does a better job of controlling for neighborhood effects. However, once we apply the tract fixed effect to the pooled sample as well, the estimate differences of TEST SCORE are not as great as found by Black (1999). Aside from this unsurprising difference, the pooled sample conclusions are robust for the boundary sample as well.

House prices in Orange County experienced a continuous and substantial drop from June 2007 through 2011, providing us a natural experiment to study school quality capitalization during different market phases. The FHFA Purchase Only house price index indicates that the Orlando MSA housing market peaked in the second quarter of 2007. Therefore, we use June 2007 to partition the period into rising market and declining market subsamples, dividing the pooled and boundary samples each into rising and falling market phases. Table 2.6 reports the key estimates for these sub-samples. It is evident that the school related variables display different effects on house prices across the two periods. During the rising market, school quality, quality volatility, and school zone instability are all significantly capitalized into house price, whereas during the falling period, these factors do not significantly affect house price when simultaneously considering FREE LUNCH and S/T RATIO. Peer effects appear to matter more in the declining market than in the rising market. The change in performance of output-based measures of school quality over the market cycle may seem counterintuitive at first blush, but we note that Zabel (2015) also finds differences across market phases. We will return to this point shortly. Nonetheless, our results suggest that even in the midst of the housing market turbulence created by the subprime mortgage crisis during which price discovery was clearly disrupted, the market still capitalized school quality, albeit somewhat differently than before.

Tables 2.7-2.9 report key estimates for higher and lower income neighborhood subsamples. The table reports estimates for the full period sample, the rising market sample and the declining market subsamples. The results reveal interesting systematic patterns across neighborhoods.

According to table 2.7, households in high income neighborhoods respond more consistently to peer effects, school zone instability and student-teacher ratios than do households in low income neighborhoods, whether for the pooled or boundary samples. Table 2.8 shows that this pattern persists in the rising market as well. In contrast, the effect of school quality and quality volatility are significant for the low income neighborhoods during the rising market period; during the declining market period, school quality capitalization in low income neighborhoods is mainly reflected in the negative peer effects observed for the percentage of students receiving free lunch. And once again, point estimates indicate somewhat weaker peer effect capitalization rates in lower income neighborhoods than in higher income neighborhoods.

As noted earlier, our test score results resemble Zabel (2015) to some extent; he finds a significant positive school test score price effect in the rising market but a significantly negative test score effect in the post-crisis falling market during 2006-2012 for the Boston MSA. It is not surprising that capitalization effects differ during the housing market collapse and the period immediately thereafter. While we do not find Zabel's (2015) surprising reverse capitalization, we do find that parents in all neighborhoods appear to rely less on the output based measures of school performance like test scores and apparently rely more on observable peer effects and student/teacher ratios during this period. Of course, the peer effects variable, proportion of students receiving free lunches, may be correlated with the spatial distribution of foreclosures and short sales arising during the market collapse.[12] But table 2.9 clearly shows that lower income neighborhoods where foreclosures are more prevalent do not exhibit stronger peer effects in the

[12]Hanson, et al. (2012) show that households sort spatially by credit quality, creating spatially clustering foreclosure risks. See Kobie, and Lee (2010), Daneshvary and Clauretie (2012), and Ihlanfeldt and Mayock, (2013) for evidence of foreclosure effects on the prices of neighboring houses.

declining market than observed in higher income neighborhoods. This pattern indicates that it is peer effects and not foreclosures that are driving the declining market results.

As a final robustness check, we consider the extent to which access to private schools, magnet schools or charter schools affect the capitalization estimates for traditional public schools. We re-estimate the models with controls to capture access to these alternative schools. One approach controls for the number of private, magnet and charter schools in the same census tract of the subject property. A second approach controls for the distances to the nearest school of each type. Neither approach changes our capitalization conclusions; while households value access to alternative schools, the availability of these alternatives does not dilute the capitalization associated with traditional schools with designated attendance zones.

5. Conclusion

It seems reasonable to presume that parents care about the education their children receive. Observable measures of public school quality, however, provide noisy signals of underlying quality to parents. Accordingly, this paper examines school quality and quality risk capitalization when school quality is uncertain. Extending residential bid rent theory to this uncertainty environment, greater expected school quality increases housing prices and steepens the price gradient, whereas the quality risk decreases the housing prices and flattens the price gradient.

The empirical models consider two fundamentally different sources of school quality risk. The first is due to stochastic variation in the quality measure or to stochastic production of quality and uncertainty in school quality. The second is due to school attendance zone instability over time. House transactions data for Orange County, Florida, over the period September, 2001, through August, 2012, provide empirical results for elementary schools that exhibit anticipated patterns; better schools increase house prices and school quality related risks reduce house prices. But point estimates are sensitive to sample periods. For example, test score capitalization, school quality volatility capitalization and school instability capitalization are stronger in the rising market than in the declining market; however, the school peer student composition affects house prices more strongly during the post-crash period than during the rising period. In addition, the instability of school attendance zones, student peer effects and the student-teacher ratio matter more in higher income neighborhoods than in lower income neighborhoods. Test scores and score volatility exhibit mixed capitalization estimates across different income level neighborhoods and vary across the market cycle. Further, our examination of other education options available to families reveals that the capitalization patterns do not appear to be driven by greater access to private schools, charter schools and magnet programs as alternatives to traditional public schools. The results instead suggest that the difference in capitalization effects we observe across neighborhoods reflects the underlying normality of education as a household consumption good and stronger consumption risk aversion with greater real income.

Appendix

Derive the non-spatial Hicksian housing and non-housing demands $\{x(P,U,\mu,\sigma), y(P,U,\mu,\sigma)\}$ as the solution to the household's problem

$$\min_{x,y}\{Px + y\} \qquad s.t. \qquad E\left[u(x, y, \mu + \sigma\varepsilon)\right] = U \tag{A.1}$$

Define the Lagrangian multiplier λ. The first order conditions for the above problem are

$$P - \lambda E[u_x] = 0 \tag{A.2}$$

$$I - \lambda E[u_y] = 0 \tag{A.3}$$

$$U - E[ux, y, \mu + \sigma\varepsilon)] = 0 \tag{A.4}$$

The bordered Hessian matrix for the optimization problem is **B**, where the second order condition $|\mathbf{B}_2| < 0$ is fulfilled for the risk averse household with strictly concave utility $u(x, y, q)$.

The total differential of the FOC is

$$\begin{bmatrix} -\lambda E[u_{xx}] & -\lambda E[u_{xy}] & -\lambda E[u_x] \\ -\lambda E[u_{yx}] & -\lambda E[u_{yy}] & -\lambda E[u_y] \\ -E[u_x] & -\lambda E[u_y] & 0 \end{bmatrix} \begin{bmatrix} dx \\ dy \\ d\lambda \end{bmatrix} = \begin{bmatrix} 0 & \lambda E[u_{xq}] & \lambda E[u_{xq}\varepsilon] \\ 0 & \lambda E[u_{yq}] & \lambda E[u_{yq}\varepsilon] \\ 1 & E[u_q] & E[u_q\varepsilon] \end{bmatrix} \begin{bmatrix} dU \\ d\mu \\ d\sigma \end{bmatrix}$$

Find the comparative static properties of the Hicksian housing demand using Cramer's Rule and the FOC to simplify:

$$\frac{\partial x}{\partial P} = \frac{E[u_y]^2}{|\mathbf{B}_2|} < 0 \tag{A.5}$$

$$\frac{\partial x}{\partial U} = \frac{\lambda^2}{|\mathbf{B}_2|}\left(E[u_{xy}]E[u_y] - E[u_{yy}]E[u_x]\right) > 0 \tag{A.6}$$

$$\frac{\partial x}{\partial \mu} = -E[u_q]\left(\frac{\partial x}{\partial U}\right) - \frac{E[u_y]}{|\mathbf{B}_2|}\left(\frac{dMRS_{x,y}}{d\mu}\right) \tag{A.7}$$

$$\frac{\partial x}{\partial \sigma} = -E[u_q \varepsilon]\left(\frac{\partial x}{\partial U}\right) + \frac{E[u_y]}{|\mathbf{B}_2|} E[u_{xq}\varepsilon - Pu_{yq}\varepsilon] \qquad \text{(A.8)}$$

Result (A.5) is the usual substitution theorem thatHicksian housing demand decreases with higher price. The sign of (A.6) follows from the assumption that housing is a normal good in the usual sense (i.e., Marshallian demand increases with household income). The second term in (A.7) uses the marginal rate of substitution between housing and nonhousing consumption along an expected utility indifference curve,

$$MRS_{x,y} = \frac{E[u_x]}{E[u_y]} \qquad \text{(A.9)}$$

to find how the change in expected school quality at any given point along the indifference curve alters the slope of the indifference curve

$$\frac{dMRS_{x,y}}{d\mu} = \frac{E[u_{xq}] - MRS_{x,y}E[u_{yq}]}{E[u_y]} \qquad \text{(A.10)}$$

Now consider the third comparative static property (A.8) pertaining to risk effects. Decomposing the first right hand side term yields $E[u_q\varepsilon] = COV[u_q, \varepsilon] < 0$ where the covariance sign follows from $u_{qq} < 0$. Therefore, the first term in (A.7) is positive when housing is normal. Looking at the second term, again decompose the product expectations using $E[\varepsilon] = 0$ to obtain

$$E[u_{xq}\varepsilon - Pu_{yq}\varepsilon] = COV[u_{xq} - Pu_{yq}, \varepsilon] \qquad \text{(A.11)}$$

Applying similar/dissimilar orderings, this term hence the second term in (A.8) takes the sign of $u_{xqq} - Pu_{yqq}$. To evaluate this expression, denote the marginal rate of substitution between housing and nonhousing consumption along an ex ante indifference curve as $mrs_{x,y} = u_x / u_y$. Using this definition, we have

$$\frac{d^2 mrs}{dq^2} = \frac{u_{xqq} - (mrs) u_{yqq}}{u_y} \tag{A.12}$$

Continuity ensures there exists a $\{x, y\}$ such that $mrs = P$. i.e., along the ex-ante indifference curve that is tangent to the consumer isocost line, where

$$\frac{d^2 mrs}{dq^2} = \frac{u_{xqq} - P u_{yqq}}{u_y}$$

This implies

$$\frac{E[u_y]}{|\mathrm{B}_2|} E[u_{xq}\varepsilon - P u_{yq}\varepsilon] \gtrless 0 \qquad as \qquad \frac{d \;\; mrs}{dq^2} \gtrless 0 \tag{A.13}$$

Thus, non-increasing $dmrs/dq$ ensures that the second term in (A.8) is positive or zero so that $\partial x/\sigma > 0$ unambiguously. Intuitively, what this requires is that changes in the slope of the ex-ante indifference curves from increases in (realized) school quality diminish at greater and greater levels of school quality. Put somewhat differently, this requires that the effect of school quality on the taste for housing does not increase with school quality.

The expenditure function under school quality uncertainty is defined as

$$e(P,U,\mu,\sigma) \equiv Px(P,U,\mu,\sigma) + y(P,U,\mu,\sigma) \tag{A.14}$$

It is straightforward to show that the standard certainty expenditure function properties hold in this context, including increasing in housing price (from Shepherd.s lemma),

$\partial e/\partial P = x(P, U, \mu, \sigma) > 0$ and increasing in expected utility, $\partial e/\partial U = \lambda > 0$. The additional properties related to school quality also hold

$$\frac{\partial e}{\partial \mu} = -E[u_q] < 0 \tag{A.15}$$

$$\frac{\partial e}{\partial \sigma} = -\frac{E[u_q \varepsilon]}{\lambda} = -\frac{COV[u_q, \varepsilon]}{\lambda} > 0 \tag{A.16}$$

To derive (A.15), differentiate (A.14), substitute from (A.2)-(A.3) to obtain

$$\frac{\partial e}{\partial \mu} = \lambda E[u_x]\frac{\partial x}{\partial \mu} + \lambda E[u_y]\frac{\partial y}{\partial \mu} \tag{A.17}$$

Substitute the Hicksian demands into (A.4) and differentiate with respect to μ to get

$$\lambda E[u_x]\frac{\partial x}{\partial \mu} + \lambda E[u_y]\frac{\partial y}{\partial \mu} = -E[u_q] \tag{A.18}$$

substituting (A.18) into (A.17) yields (A.15).

Follow the same procedure to derive (A.16). Differentiate (A.14) with respect to σ, evaluate the result at $\sigma = 1$, and substitute from (A.2) - (A.3) to get

$$\frac{\partial e}{\partial \sigma} = \lambda E[u_x]\frac{\partial x}{\partial \sigma} + \lambda E[u_y]\frac{\partial y}{\partial \sigma} \tag{A.19}$$

Substitute the Hicksian demands into (A.4) and differentiate to with respect to σ get

$$\lambda E[u_x]\frac{\partial x}{\partial \sigma} + \lambda E[u_y]\frac{\partial y}{\partial \sigma} = -E[u_q \varepsilon] \tag{A.20}$$

Substitute (A.20) into (A.19) yields (A.16).

The intuition for these expenditure function properties is straightforward. Greater average education quality by itself increases expected utility so that the household requires less spending on housing and other goods to attain a given expected utility level. Similarly, a greater mean-preserving-spread lowers expected utility under risk aversion so that the household requires more spending on housing and other goods to compensate for the loss in expected utility. These results are key to deriving the effects of expected school quality and risk on the shape of the equilibrium housing price function.

References

Barrow, L., Rouse, C.E. 2004. Using market valuation to assess public school spending. Journal of Public Economics 88 (9), 1747–1769.

Black, S.E. 1999. Do better schools matter? Parental valuation of elementary education. Quarterly Journal of Economics 114 (2), 577–599.

Bogart, W.T., Cromwell, B.A., 2000. How much is a neighborhood school worth? Journal of Urban Economics 47 (2), 280–305.

Bogin, A., & Nguyen-Hoang, P. 2014. Property Left Behind: An Unintended Consequence of a No Child Left Behind "Failing" School Designation. Journal of Regional Science, 54(5), 788-805.

Brasington, D.M. 1999. Which measures of school quality does the housing market value? Spatial and non-spatial evidence. Journal of Real Estate Research 18 (3), 395–413.

Brasington, D.M. 2001. Capitalization and community size. Journal of Urban Economics 50 (3), 385–395.

Brasington, D.M., Haurin, D.R. 2006. Educational outcomes and house values: a test of the value-added approach. Journal of Regional Science 46 (2), 245–268.

Brasington, D.M., Haurin, D.R. 2009. Parents, peers, or school inputs: which components of school outcomes are capitalized into house value? Regional Science and Urban Economics 39 (5), 523–529.

Chiodo, A., Hernández-Murillo, R., & Ouyang, M.T. 2010. Nonlinear effects of school quality on house prices. Federal Reserve Bank of St. Louis Review, 92 (May/June 2010).

Clark, D.E., Herrin, W.E. 2000. The impact of public school attributes on home sale prices in California. Growth and Change 31 (3), 385–407.

Clauretie, T.M., Neill, H.R. 2000. Year-round school schedules and residential property values. Journal of Real Estate Finance and Economics 20 (3), 311–322.

Crone, T. 2006. Capitalization of the quality of local public schools: What do home buyers value? Working paper. Federal Reserve Bank of Philadelphia.

Dills, A.K., 2004. Do parents value changes in test scores? High stakes testing in

Texas. Contributions to Economic Analysis & Policy 3 (1), 1–32.

Downes, T.A., Zabel, J.E., 2002. The impact of school characteristics on house prices: Chicago 1987–1991. Journal of Urban Economics 52 (1), 1–25.

Figlio, D.N., Lucas, M.E., 2004. What's in a grade? School report cards and the housing market. American Economic Review 94 (3), 591–604.

Hanushek, E. A. 1986. The economics of schooling: Production and efficiency in public schools. Journal of Economic Literature, 1141-1177.

Hanushek, E. A. 1997. Assessing the effects of school resources on student performance: An update. Educational Evaluation and Policy Analysis, 19(2), 141-164.

Hughes Jr, W. T., & Turnbull, G. K. 1996. Uncertain neighborhood effects and restrictive covenants. Journal of Urban Economics, 39(2), 160-172.

Kane, T. J., Staiger, D. O., Samms, G., Hill, E. W., & Weimer, D. L. 2003. School Accountability Ratings and Housing Values [with Comments].Brookings-Wharton Papers on Urban Affairs, 83-137.

Mathur, Shishir, 2008. Impact of transportation and other jurisdictional level infrastructure and services on housing prices. Journal of Urban Planning and Development 134 (1), 32–41.

Nguyen-Hoang, P., & Yinger, J. 2011. The capitalization of school quality into house values: A review. Journal of Housing Economics, 20, 30-48.

Oates, Wallace E. 1969. The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the Tiebout hypothesis. Journal of Political Economy, 957-971.

Rosen, H. S., & Fullerton, D. J. 1977. A note on local tax rates, public benefit levels, and property values. Journal of Political Economy, 433-440.

Turnbull, G. K. 1991. The spatial demand for housing with uncertain quality. Journal of Real Estate Finance and Economics, 4(1), 5-17.

Turnbull, G. K. 1994. Location and housing demand with endogenous consumption risk. Regional Science and Urban Economics, 24(5), 543-563.

Turnbull, G. K. 1995. Urban Consumer Theory, Washington D.C.: Urban Institute Press.

Yinger, J. 2010. Hedonic Markets and Explicit Demands: Bid-Function Envelopes for Public Services and Neighborhood Amenities. Center for Policy Research Working paper, Syracuse University.

Zabel, J. 2015. The hedonic model and the housing cycle. Regional Science and Urban Economics, *54*, 74-86.

Zahirovic-Herbert, V., Turnbull, G.K. 2008. School quality, house prices and liquidity. Journal of Real Estate Finance and Economics 37 (2), 113–130.

Zahirovic-Herbert, V., Turnbull, G.K. 2009. Public school reform, expectations, and capitalization: what signals quality to homebuyers? Southern Economic Journal 75 (4), 1094–1113.

Figures and Tables

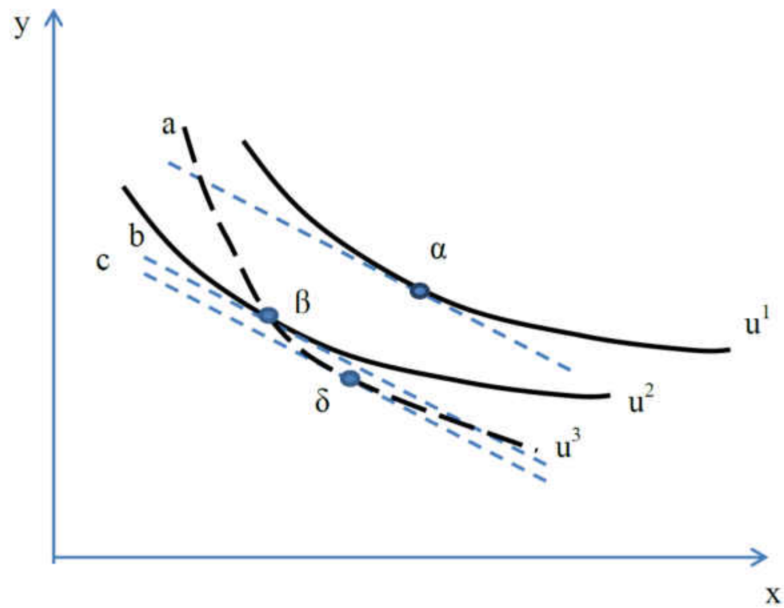

Figure 2.1: School quality risk effect on household choice of housing, x, and nonhousing consumption, y. Greater quality risk when no effect on MRS shifts equilibrium iso-expenditure from a to b. If MRS increases with school quality then equilibrium iso-expenditure shifts to c, offsetting effect of a and b on housing demand x.

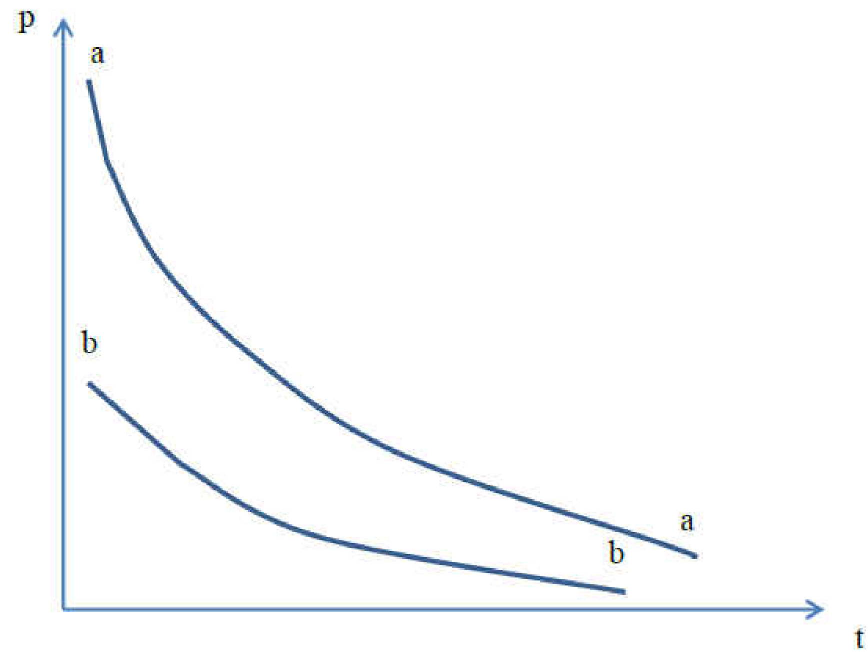

Figure 2. 2: Greater expected school quality at all locations t shifts house function from bb to aa, increasing price at gradient at all t. Greater school quality risk at all locations shifts house price function from aa to bb, decreasing price and gradient at all t.

Table 2.1: The new schools added during 2001 to 2012

	Elementary	**Middle**	**High**
Number of non-chartered public schools in 2000-2001	98	26	12
Number of school added during 2001-2012	25	9	7
2001-2002	Avalon Elementary Camelot Elementary Three Points Elementary	Odyssey Middle	Olympia High Timber High
2002-2003	Thornebrooke Elementary		
2003-2004			Freedom High
2004-2005	West Oaks Elementary Eagle's Nest Elementary West Creek Elementary		
2005-2006	Andover Elementary East Lake Elementary Whispering Elementary	Freedom Middle Lagacy Middle	Ocoee High
2006-2007	Bay Meadows Elementary Castle Creek Elementary Stone Lakes Elementary Vista Lakes Elementary Wolf Lake Elementary Wyndham Lakes Elementary	Avalon Middle South Creek Middle Wolf Lake Middle	
2007-2008	Millennia Elementary Moss Park Elementary Sun Park Elementary	Bridgewater Middle	Wekiva High
2008-2009	Timber Elementary Westbrooke Elementary		
2009-2010	Keene's Crossing Elementary	Lake Nona Middle	East River High Lake Nona High
2010-2011			
2011-2012	Blankner Wetherbee Elementary Forsyth Woods Elementary	Blankner	

Table 2.2: Variable descriptions

Variable	Description
Dependent variable:	
PRICE	Natural logarithm of house price
House characteristics	
AGE	The age of house
AREA	heated area of the house, unit in thousand square feet
NET AREA	NET AREA of a house equals the total area minus (heated) area, unit in thousand square feet
BATHS	Number of bathrooms
BEDS	Number of bedrooms
CONDITION	House quality index with 6 levels
Community variables	
TAX	Gross tax
DISTANCE	Distance to CBD, unit in miles
School related variables	
TEST SCORE	Mean value of normalized [a] elementary math scores since 00-01 school year
VOLATILITY	Standard deviation of normalized math scores
CHANGE	Dummy variable for school zone change (0.1 or 0)
INSTABILITY	Dummy of school zone changes one more time or not (0.1 or 0)
FREELUNCH	Portion of students enjoying free or reduced lunch
STRATIO	Student/teacher ratio/10
Variables from census at block group level	
INCOME	Median income in thousand dollars

Note: [a] normalized math score for a school equals the score of that school divided by the average score of the whole county

Table 2.3: Data Summary Statistics

Panel A: pooled sample				
Variable	**Mean**	**Std. Dev.**	**Minimum**	**Maximum**
Dependent variable:				
PRICE	12.0246	0.5525	10.5967	13.8878
House characteristics				
AGE	19.0062	14.5052	0	200
AREA	1.8262	0.7031	0.3040	6.7510
NET AREA	0.5782	0.2962	0	4.9760
BATHS	2.1389	0.6487	1	5
BEDS	3.3155	0.7319	1	6
CONDITION	3.4086	0.9352	1	6
Community variables				
TAX	18.1142	0.9092	15.5201	22.2056
DISTANCE	8.6874	3.9625	0.2971	28.2263
School related variables				
TEST SCORE	1.0440	0.1827	0.5619	1.4354
VOLATILITY	0.1045	0.0411	0.02297	0.2722
CHANGE	0.0217	0.0412	0	0.1
INSTABILITY	0.0047	0.0212	0	0.1
FREELUNCH	0.5335	0.2151	0	0.9971
STRATIO	1.5331	0.1696	0.9300	1.9400
Variables from census at block group level				
INCOME	54.5245	19.2972	12.5001	210.1796
BLACK	14.6987	17.9749	0	98.2100
HISPANIC	21.9524	15.4672	0	87.0400
Panel B: Elementary school boundary sample				
Variable	**Mean**	**Std. Dev.**	**Minimum**	**Maximum**
Dependent variable:				
PRICE	11.9551	0.5391	10.5984	13.8832
House characteristics				
AGE	19.9997	14.8514	0	200
AREA	1.7259	0.6570	0.4400	6.751
NET AREA	0.5390	0.2752	0	3.460
BATHS	2.0669	0.6202	1	5
BEDS	3.2688	0.7199	1	6
CONDITION	3.3356	0.9017	1	6
Community variables				
TAX	18.1353	0.8999	15.5201	20.845
DISTANCE	7.9897	3.4950	0.7174	18.2643
School related variables				
TEST SCORE	1.0240	0.1946	0.5619	1.3651
VOLATILITY	0.1077	0.0421	0.0230	0.2722

CHANGE	0.0262	0.0440	0	0.1
INSTABILITY	0.0048	0.0214	0	0.1
FREELUNCH	0.5509	0.2277	0	0.9959
STRATIO	1.5327	0.1705	0.9300	1.9400
Variables from census at block group level				
INCOME	51.9541	17.9972	12.5498	174.1692

Note: Summary statistics for dummy variables of sale year and tracts are not reported.

Table 2.4: The estimates of the pooled sample over the period 2001-2012

Explanatory variable	Dependent variable ln(price) Model 1	Model 2	Model 3	Model 4
Constant	10.5079*** (0.1856)	10.5425*** (0.1814)	10.7589*** (0.1940)	10.9164*** (0.1988)
AGE	-0.0055*** (0.0003)	-0.0055*** (0.0003)	-0.0055*** (0.0003)	-0.0055*** (0.0003)
AREA	0.3139*** (0.0061)	0.3141*** (0.0061)	0.3138*** (0.0062)	0.3133*** (0.0063)
NEW AREA	0.2456*** (0.0087)	0.2451*** (0.0085)	0.2437*** (0.0083)	0.2438*** (0.0084)
BATHS	0.0392*** (0.0046)	0.0391*** (0.0045)	0.0387*** (0.0044)	0.0393*** (0.0043)
BEDS	0.0027 (0.0023)	0.0028 (0.0023)	0.0029 (0.0023)	0.0028 (0.0023)
CONDITION	0.0275*** (0.0032)	0.0276*** (0.0032)	0.0274*** (0.0032)	0.0277*** (0.0032)
TAX	0.0206** (0.0087)	0.0204** (0.0087)	0.0211** (0.0085)	0.0212** (0.0086)
DISTANCE	-0.0141*** (0.0041)	-0.0139*** (0.0037)	-0.0128*** (0.0038)	-0.0133*** (0.0040)
TEST SCORE	0.2495*** (0.0360)	0.2328*** (0.0362)	0.0778* (0.0435)	0.0581 (0.0449)
VOLATILITY		-0.2252* (0.1251)	-0.2367** (0.1072)	-0.2696** (0.1121)
CHANGE		0.0155 (0.0846)	0.0028 (0.0851)	0.0074 (0.0813)
INSTABILITY		-0.2847* (0.1569)	-0.3295** (0.1405)	-0.3211** (0.1352)
FREELUNCH			-0.1401*** (0.0306)	-0.1620*** (0.0309)
STRATIO				-0.0780*** (0.0195)
Number of Observations Used	127120	127120	127120	127120
Sale Year Fixed Effect	yes	yes	yes	yes
Tract Fixed Effect	yes	yes	yes	yes
Number of Clusters	167	167	167	167
R-Square	0.8627	0.8628	0.8633	0.8635

Note: This table shows the regression results for the pooled sample spanning 2001-2012. Model 1 only considers the school quality effect; model 2 includes school quality and school quality risk, school zone change effect, and school

zone uncertainty risk in order to examine the effects of school related risks on house price and the effect of school related risk on school quality capitalization; besides variables in model 2, models 3 adds school peer effect: free lunch to discriminate the school output and school peer effect on house price; beyond model 3, model 4 also includes student/teacher ratio to further separate school input effect and school output effects in house pricing. Clustered standard errors are presented in parentheses. Coefficients for dummy variables for sale year and location controls based on census tracts are not reported. Stars denote statistical significance:
*** P<0.01; **P<0.05;*P<0.10.

Table 2.5: The estimates of the boundary sample over the period 2001-2012

Key explanatory variable	Dependent variable ln(price)			
	Model 1	**Model 2**	**Model 3**	**Model 4**
TEST SCORE	0.2166*** (0.0438)	0.2061*** (0.0442)	0.0606 (0.0521)	0.0459 (0.0526)
VOLATILITY		-0.1269 (0.1106)	-0.1572* (0.0920)	-0.1714* (0.0969)
CHANGE		0.0195 (0.0774)	0.0137 (0.0761)	0.0178 (0.0734)
INSTABILITY		-0.2859** (0.1374)	-0.2958** (0.1252)	-0.2821** (0.1217)
FREELUNCH			-0.1388*** (0.0389)	-0.1533*** (0.0390)
STRATIO				-0.0584*** (0.0219)
Number of Observations Used	36607	36607	36607	36607
Year Fixed Effect	yes	yes	yes	yes
Tract Fixed Effect	yes	yes	yes	yes
Number of Clusters	123	123	123	123
R-Square	0.8592	0.8593	0.8599	0.8601

Note: This table shows the estimates for the elementary school boundary sample spanning 2001-2012. Clustered standard errors are presented in parentheses. Coefficients for dummy variables for sale year and location controls based on census tracts are not reported. Stars denote statistical significance:
*** P<0.01; **P<0.05;*P<0.10.

Table 2.6: The estimates during the rising and falling market

Key explanatory variable	Dependent variable ln(price)			
	08/2001-06/2007		07/2008-12/2012	
	Pooled sample	**Boundary sample**	**Pooled sample**	**Boundary sample**
TEST SCORE	0.1761***	0.1064**	-0.0925	-0.0467
	(0.0396)	(0.0461)	(0.1042)	(0.1233)
VOLATILITY	-0.2698***	-0.1996**	0.0412	0.0813
	(0.1041)	(0.0943)	(0.1779)	(0.1958)
CHANGE	-0.0095	0.0398	-0.0458	-0.0339
	(0.0785)	(0.0678)	(0.0951)	(0.1024)
INSTABILITY	-0.3794**	-0.4462**	-0.3339*	-0.0799
	(0.1770)	(0.2190)	(0.1823)	(0.1706)
FREELUNCH	-0.0359*	-0.0607***	-0.3300***	-0.2877***
	(0.0187)	(0.0190)	(0.0695)	(0.0708)
STRATIO	-0.0385**	-0.0169	-0.0448*	-0.0499*
	(0.0180)	(0.0199)	(0.0235)	(0.0293)
Number of Observations Used	84472	25239	42648	11368
Sale Year Fixed Effect	yes	yes	yes	yes
Tract Fixed Effect	yes	yes	yes	yes
Number of Clusters	165	121	145	73
R-Square	0.8835	0.8780	0.8396	0.8372

Note: This table presents the coefficients of samples during two sub-periods 2001-2007 and 2008-2012, examining the effects of market condition on school quality capitalization. Clustered standard errors are presented in parentheses. Coefficients for dummy variables for sale year and location controls based on census tracts are not reported. Stars denote statistical significance:
*** P<0.01; **P<0.05;*P<0.10.

Table 2.7: The estimates during 2001-2012 considering income effect

Key explanatory Variables	**Dependent variable ln(price)**			
	High income		**Low income**	
	Pooled sample	**Boundary sample**	**Pooled sample**	**Boundary sample**
TEST SCORE	-0.0366 (0.0845)	-0.0677 (0.0993)	0.1236*** (0.0468)	0.0809 (0.0553)
VOLATILITY	-0.2737 (0.2162)	-0.0428 (0.2105)	-0.2473** (0.1085)	-0.1331 (0.0902)
CHANGE	-0.0158 (0.1343)	-0.0198 (0.1147)	-0.0066 (0.0667)	0.1138 (0.0797)
INSTABILITY	-0.4539** (0.1818)	-0.3237** (0.1456)	-0.1013 (0.1345)	-0.2333 (0.1823)
FREELUNCH	-0.2535*** (0.0500)	-0.2623*** (0.0471)	-0.0615** (0.0247)	-0.0709** (0.0279)
STRATIO	-0.1058*** (0.0263)	-0.0756** (0.0301)	-0.0474** (0.0239)	-0.0399* (0.0224)
Number of Observations Used	63547	16700	63573	19907
Year dummies included	yes	yes	yes	yes
class level	99	65	132	99
Number of Clusters	99	65	132	99
R-Square	0.8603	0.8588	0.8366	0.8358

Note: This table presents the coefficients of pooled and boundary subsamples of high income and low income during the period 2001-2012, examining the effects of residents' income on school quality capitalization. Clustered standard errors are presented in parentheses. Coefficients for dummy variables for sale year and location controls based on census tracts are not reported. Stars denote statistical significance:
*** $P<0.01$; ** $P<0.05$; * $P<0.10$.

Table 2.8: The estimates in the rising market considering income effect

Key explanatory Variable	Dependent variable ln(price)			
	High income		Low income	
	Pooled sample	Boundary sample	Pooled sample	Boundary sample
TEST SCORE	0.2013*** (0.0738)	0.0560 (0.0616)	0.1530*** (0.0486)	0.0828 (0.0591)
VOLATILITY	-0.5657*** (0.2069)	-0.2740 (0.1791)	-0.2009** (0.1018)	-0.1743* (0.0985)
CHANGE	0.0038 (0.1585)	0.0185 (0.1222)	-0.0325 (0.0732)	0.1017 (0.0763)
INSTABILITY	-0.6363*** (0.2027)	-0.6038** (0.2351)	-0.0396 (0.1360)	-0.1993 (0.1557)
FREELUNCH	-0.0630 (0.0423)	-0.1363*** (0.0385)	-0.0250 (0.0187)	-0.0282 (0.0191)
STRATIO	-0.0937*** (0.0274)	-0.0952*** (0.0288)	-0.0247 (0.0245)	-0.0021 (0.0266)
Number of Observations Used	35027	9465	49445	15774
Year dummies included	yes	yes	yes	yes
class level	95	59	131	96
Number of Clusters	95	59	131	96
R-Square	0.8784	0.8745	0.8448	0.8381

Note: This table presents the coefficients of pooled and boundary subsamples of high income and low income during the period 2001-2007, examining the effects of residents' income on school quality capitalization in the housing rising market. Clustered standard errors are presented in parentheses. Coefficients for dummy variables for sale year and location controls based on census tracts are not reported. Stars denote statistical significance:
*** P<0.01; **P<0.05;*P<0.10.

Table 2.9: The estimates in the falling market considering income effect

Key explanatory variable	Dependent variable ln(price)			
	High income		Low income	
	Pooled sample	Boundary sample	Pooled sample	Boundary sample
TEST SCORE	-0.0870 (0.1327)	-0.0920 (0.1804)	-0.0371 (0.1102)	0.0071 (0.1286)
VOLATILITY	0.1305 (0.2333)	0.3048 (0.3290)	-0.0699 (0.2014)	0.1156 (0.2159)
CHANGE	-0.0829 (0.1137)	-0.0102 (0.1263)	0.1238 (0.1106)	0.1318 (0.1987)
INSTABILITY	-0.3351 (0.2132)	-0.1496 (0.2120)	-0.0014 (0.3225)	-0.0428 (0.4979)
FREELUNCH	-0.3497*** (0.0818)	-0.3156*** (0.0907)	-0.2525*** (0.0902)	-0.2352** (0.1083)
STRATIO	-0.0725** (0.0323)	-0.0273 (0.0426)	-0.0128 (0.0365)	-0.0634 (0.0413)
Number of Observations Used	28520	7235	14128	4133
Sale Year Fixed Effect	yes	yes	yes	yes
Tract Fixed Effect	yes	yes	yes	yes
Year dummies included	yes	yes	yes	yes
class level	91	49	98	47
Number of Clusters	91	49	98	47
R-Square	0.8275	0.8222	0.7902	0.7831

Note: This table presents the coefficients of pooled and boundary subsamples of high income and low income during the period 07/2007-2012, examining the effects of residents' income on school quality capitalization in the housing falling market. Clustered standard errors are presented in parentheses. Coefficients for dummy variables for sale year and location controls based on census tracts are not reported. Stars denote statistical significance:
*** P<0.01; **P<0.05; *P<0.10.

ESSAY3 A META-ANALYSIS OF SCHOOL QUALITY CAPITALIZATION IN HOUSE PRICES

1. Introduction

Even though it seems reasonable to expect public school quality to be capitalized into housing prices, empirical estimates of these capitalization effects vary considerably not only across studies but also within studies. Differences are reflected in coefficient values, statistical significance, and even the direction of capitalization. The reasons for the variety of results remain unclear; is it due to studies using different school quality measures, control variables, geographic regions, sample time periods, estimation techniques or some combination of these factors? Or is there something in the housing market price discovery process that precludes consistent patterns of capitalization? This paper applies meta-analysis to evaluate the relationship between different capitalization conclusions across and within studies and the characteristics of those studies. The results provide an overview of the state of the school quality capitalization literature and a useful adjunct to existing reviews of the literature.[13]

This study surveys school quality capitalization estimates from studies appearing in 1968 and thereafter. We use meta-analysis to empirically evaluate which of the variety of school quality measures used in the empirical literature most closely relate to households' perceptions of quality as well as the extent to which capitalization results are influenced by data sample characteristics, variable definitions, econometric methods and other factors. In order to capture the sign of

[13] See Ross and Yinger (1999) and Nguyen-Hoang and Yinger (2011) for interpretive reviews of the empirical capitalization literature.

capitalization effects and their statistical significance in individual studies, we use a measure based the t-statistics for school quality variables as dependent variables in the meta regression model. The right hand side variables in the meta model include indicators for the type of school quality measure used in the original study and whether the study includes controls for school peer effects, neighborhood characteristics, or other relevant variables. Our meta model also includes variables controlling for the nature of the data both by level of aggregation and geographic region as well as aspects of the econometric model and estimation method.

Most scholars and policy makers believe that parents care about their children's educational quality, but the formal measures of school quality used in empirical studies may or may not accurately reflect parents' perceptions of school quality. It may be that the reported quality measures are noisy signals of underlying quality or that the quality of education produced by individual schools is inherently stochastic (Turnbull, et al. 2016). Either way, the empirical measures of school quality are imperfect signals and it is reasonable to expect households to recognize this fact, although the degree to which they do so is not known. Meta-analysis can help sort out the extent to which the characteristics of the quality measure itself is associated with specific capitalization conclusions from the extent to which the price discovery process in illiquid markets for heterogeneous goods like housing may inhibit full capitalization in house values. Meta-analysis also offers insights into which aspects of individual studies influence empirical conclusions.

This is a crucial step in sorting out the nature of measuring school quality from the underlying complications ignored in many capitalization studies. Because housing markets are both local and spatial within localities, at some point the empirical literature is going to have to address the fact that school quality price effects are likely to vary across neighborhoods even when quality is accurately measured. First, housing market conditions vary across locations in metropolitan areas and may yield different capitalization at different locations (Brasington 1999). Second, even in the absence of varying market conditions across locations, household uncertainty over school quality by itself implies spatial variation in capitalization in bid rent theory under uncertainty (Turnbull, et al. 2016). Third, education is a normal good so the strength of parental willingness and ability to pay for higher quality varies with household income or wealth. And fourth, households in ethnic or cultural enclaves in U.S. urban areas also exhibit different preferences for educational quality. Any combination of these four factors may drive capitalization patterns to differ across neighborhoods.

The results are useful from a narrower modeling perspective as well. Ever since the first empirical studies, the tendency has been to introduce a greater variety of school quality measures into the hedonic house price model and to rely on more complex econometric methods and models in order to obtain better estimates. The result has been anything but consensus in terms of types of data or models used across studies. Meta-analysis offers the chance to step back for a moment to evaluate what the long line of existing research offers in terms of guidance on these issues. The air pollution and environmental quality literature has already moved in this direction (Harrison and Rubinfeld 1978, Smith and Huang 1995, Nelson and Kennedy 2009). The existing body of empirical educational quality capitalization studies has developed to the point where it, too, can benefit from the same of introspection.

2. Literature Review

2.1 Trends in the School Quality Capitalization Literature

Researchers have been studying school quality capitalization in house prices for more than 50 years. Various school related variables are used to measure school quality, which we divide into three categories: input based measures, including per pupil expenditure, student/teacher ratio, and all other types of expenditures; peer effects variables, including the percentage of minority students, the percentage of students who enjoy subsidized lunches, etc.; and output based measures, including mean test scores over time, recent test scores, and value-added based on changes in test scores.

In the early years from 1968 to about 1985, researchers tend to rely on per pupil expenditure to represent public school service quality. In our list of sources in Table 1, these studies include Orr (1968), Oates (1969), Heinberg and Oates (1970), Pollakowski (1973), Hyman and Pasour (1973), Edel and Sclar (1974), Meadows (1976), Schnare and Struyk (1976), Gustely (1976), McMillan and Carlson (1977), Rosen and Fullerton (1977), Harrison and Rubinfeld (1978), Brueckner (1979), Gurwitz (1980), and Cushing (1984). Investigators originally concentrate on the tax rate effect on house price, later gradually turning to the effect of per pupil expenditure. Because of the nature of available data and the existing state of data processing, these early studies use aggregate or average house prices. The data are drawn from the US census and sample sizes tend to be small. Table 1 illustrates that the results from these studies vary considerably.

Li and Brown (1980) is the first study using transaction data in our sample of papers. They examine the effect of three types of micro-neighborhood variables like aesthetic attributes, pollution levels and proximity on house price. The elementary school test score is included in the model as a measure of public service. The increasing availability of school related variables in subsequent studies corresponds with a greater focus on the effects of school quality on house prices. Brasington (1999) illustrates this focus and provides a microcosm of the mixed results found in the literature before and after his study. He carries out 444 hedonic regressions to test the capitalization significance of 37 school quality measures with 2 models, OLS and mixed spatial autocorrelation, using house sales in six Ohio MSAs. The 37 school district level variables include input variables, test scores and value-added based on test scores. His results show that no school measures are capitalized in house prices significantly in all six MSAs using the two models. The best performance comes from the elementary school math test score, which has positive significant estimates for all six MSA samples when using the spatial autocorrelation model and five out of six MSA samples when using OLS. The relative significance of quality capitalization between OLS and autocorrelation models is not clear. Value-added based on test scores systematically yield even lower significance.

Most of the later studies include various school quality measures in one regression, although the main purpose of some of them is not to compare the capitalization significance of those variables. These include Jud (1981), Walden (1990), Hayes and Taylor (1996), Clauietie and Neill (2000), Clark and Herrin (2000), Weimer and Wolkoff (2001), Downes and Zabel (2002), Crone (2006), Kane, Riegg, and Staiger (2006), Brasington and Haurin (2006), Clapp, Nanda, and Ross (2008), Sedgley, Williams, and Derrick (2008), Zahirovic-Herbert and Turnbull (2008, 2009), Seo and Simons (2009), Dhar and Ross (2012), and Turnbull, Zahirovic-Herbert, and Zheng (2016) in Table 1. Among these studies, Jud (1981) finds a strong positive estimates for elementary school test scores, but not the minority student ratio; Hayes and Taylor (1996) and Downes and Zabel

(2002) show that school test scores outperform the input based variable expenditure per pupil; similarly, Crone (2006) and Seo and Simons (2009) argue that output based variables perform better. In contrast, Clauietie and Neill (2000) conclude that test scores are not significant among the various school quality measures they use; Clark and Herrin (2000) argue that input based measures outperform output based measures; and Clapp, et al. (2008) show that the percentage of Hispanic students has consistently strong effects on house price while middle school average math exam scores price effects are mixed. In the midst of the above disagreements among studies, there is at least one point of agreement: value-added appears less powerful in capitalization than other school quality measures.

School quality capitalization research focuses more on elementary schools than middle and high schools. In Table 3.1, 31 out of 44 papers, or 29 out of 31 papers published after 1980, use elementary school quality measures. However, Walden (1990), Weimer and Wolkoff (2001), Crone (2006), and Sedgley, et al. (2008) argue that middle school and/or high school quality matter more than the elementary school quality in house pricing. Adding more variation to the body of existing conclusions, Downes and Zabel (2002) and Crone (2006) disagree with whether neighborhood school quality or school district quality is more important: the former argues that school quality measures outperform the district quality measures while the latter comes to the opposite conclusion on this question.

Another branch of the literature focuses on distinguishing the school quality effect from other neighborhood amenities influencing house prices. The authors use three methods: adding neighborhood variables to regression models, controlling for location, and using instrumental variables methods. Many papers include extra neighborhood variables in regressions. Pollakowski

(1973), Downes and Zabel (2002), Brasington and Haurin (2006), and Bayer, Fernando, and Mcmillan (2007) consider the effects of adding neighborhood variables and find that school variables estimates are affected by adding neighborhood variables to the model. With respect to location controls, there are four approaches: adding distance to the CBD, using local jurisdiction fixed effects, using neighborhood fixed effects and including school boundary fixed effects. Twenty two out of 45 papers include distance to CBD. Many studies use school boundary fixed effects (BFE) after the method was popularized by Black (1999), including Kane, et al. (2003), Crone (2006), Bayer, et al. (2007), Zahirovic-Herbert and Turnbull (2009), Dhar and Ross (2012) and Imberman and Lovenheim (2013) in our sample. All of these papers report strong evidence that the impact of school quality on price declines when using boundary samples and BFE. There is some disagreement over the appropriate methods for introducing neighborhood controls. Black (1999) shows that adding neighborhood variables to boundary sample does not reduce school quality capitalization as much as using the BFE model. Crone (2006) finds that including neighborhood variables in the school boundary samples reduces school quality capitalization more than using BFE alone.

Kane, et al. (2003), Clapp, et al. (2008), Zahirovic-Herbert and Turnbull (2009), Dougherty, et al. (2009), and Turnbull, et al. (2016) use neighborhood fixed effects. Weimer and Wolkoff (2001), Brunner, Murdoch, and Thayer (2002), Kane, et al. (2003), Crone (2006), Clapp, et al. (2008), Zahirovic-Herbert and Turnbull (2009) use local jurisdiction fixed effects. Clapp (2008) shows that using jurisdiction fixed effects plus census tract variables yields roughly the same reduction in school capitalization as using neighborhood fixed effects.

Weimer and Wolkoff (2001) in our sample use median house value for neighborhood quality. Downes and Zabel (2002) consider the endogeneity of tax rates, per pupil expenditures and test scores. They use the proportion of the tax base that is residential, per pupil assessed value, the proportion renting, and the proportion of the school aged population as instrumental variables. This approach, however, has not been popular in recent years.

In our sample, two papers include children private school attendance rates in their hedonic models. Hayes and Taylor (1996) include the share of the elementary school population attending private school and find that it positively impacts house prices. Clark and Herrin (2000) similarly include the percentage of students enrolled in private school in census tracts as a school attribute variable. In contrast with Hayes and Taylor (1996), they find that private school attendance reduces house price. Neither paper includes results comparing including and excluding private school attendance. Further, there is an insufficient number of papers in our sample including private schools so we do not include this factor in our meta-analysis.

Finally, school quality capitalization may differ before and after the recent housing market crash. Zabel (2015) finds that school test score significantly reduce house prices in the downturn during 2006-2012 in the Boston MSA even though test scores generate a positive effect before the crisis. Turnbull, et al. (2016) show that different measures of school quality generate meaningful capitalization effects during the crises than in the previous rising market in Orange County, FL; they find that test scores are not significantly capitalized in house prices during the downturn but peer effects are. The results are consistent with Zabel (2015) in terms of test scores but does not clarify why test scores do not correlate with households' perceptions about school in the downturn while peer effects appear to do so. Not enough time has passed to provide a large enough number of studies looking specifically at capitalization during the market crash to include the market phase as a factor in our meta-analysis.

2.2 Meta-Analysis Overview

Meta-analysis is a statistical method that contrasts and synthetizes results from different studies to identify patterns among studies in order to explain factors associated with disparities among results. The method is widely used in experimental medical, psychological and educational research and has spread to other fields, including fields relying on non-experimental research. The method has been applied in the housing literature as well. For example, Smith and Huang (1996) use meta-analysis to evaluate estimates of marginal willingness-to-pay (MWTP) for improving air quality. Their results show that the market conditions and the procedures used to implement the hedonic models are important to the resulting MWTP. Schipper, Nijkamp, and Rietveld (1998) apply meta-analysis to explain the variation among noise depreciation index (NDI) in hedonic models. They conclude that variables such as timing, sample wealth, and specification of the original studies significantly contribute to the variation of NDI. Nelson (2004) studies airport noise effects on house prices with meta-analysis, using the estimated noise discount as the effect size (i.e., the dependent variable in meta regression). He finds that the country and model specification have some effects on the measured noise discount. Sirmans, MacDonald, and Macpherson, and Zietz (2006) perform meta-analysis on hedonic house price models, looking at nine house characteristic variables. Geographic location, time, household wealth (median income), and size of the hedonic model are considered as candidates affecting the estimates. Debrezion, Pels, and Rietveld (2007) use meta-analysis to explain the variation in the results from studies of the impact of railway stations on residential and commercial property values. Sirmans, MacDonald, and Macpherson (2010) look at how factors like controlling for year of sale, income, model-specification, and location influence time on market (TOM) coefficients in hedonic price models. Braden, Feng, and Won (2011) study waste site effects on property value with meta-analysis.

3. Data and Variables

This study draws results from articles related to school quality capitalization from 1968 and after. The main objective is to identify factors that affect the sign and significance of public school quality impacts on house prices, so we extract the t-statistics for school quality measures in the hedonic regressions to construct our meta dependent variable. Coefficient estimates of school quality variables are not appropriate because there are many school quality measures differing in magnitudes and concepts. Considering that some quality measures, like the student/teacher ratio, have negative expected effects on quality hence house price, we define an effective t-statistic as having the same sign as the original t-statistic when we expect the sign of the coefficient estimates of school quality measures to be positive and minus one multiplied to the t-statistic when we expect the sign of the coefficient estimates to be negative. This approach allows the meta-analysis to identify whether individual factors lead to unexpected capitalization effects as well as statistically insignificant effects.

To maintain comparability across studies, we do not include all model results reported in the surveyed papers. Because foreign public school systems are different from that in United States, we exclude papers using non-US data. Some authors use subsamples to investigate structural differences when the magnitudes of some variables differ. For example, Clauretie and Neill (2000) divide their sample into six subsamples according to the number of bedrooms. We do not use subsample regressions because the subsample structures are not consistent or systematic across studies. We include regressions estimated using individual school attendance zone or school district boundary subsamples because these represent standard subsample partition methods and there is a sufficient number of papers using these types of samples.

We also exclude any regressions using interaction effect items between a school quality variable and other variables because the interaction effects make calculating an effective t-statistic dependent on data values. Similarly, we do not include results for school quality nonlinear effects because there is not a large enough number of studies using this method.

For papers using models with attendance zone boundary fixed effects, we include the results pertaining to only one definition of such fixed effects. Finally, we delete observations with an absolute t-statistic larger than 20, as these appear to be extreme values in the set of surveyed studies.

The meta-analysis sample draws from 45 different papers and includes 368 observations. There are 210 observations using output based school quality measures like test scores; the rest rely solely on input based measures like spending or student/teacher ratios. We conduct the meta-analysis for both the pooled samples of 368 observations and the subsample of 210 observations pertaining to output based measures.

Table 3.2 reports some important characteristics of each control variable used in this exercise as they relate to capitalization results, including the frequency of the treatment in our pooled sample, the average, maximum, minimum effective t-statistics related to capitalization when the indicated control variable is present in the regression, and the number of equations using the indicated control variable. This table offers an overview of the relative importance of various control variables in the capitalization literature.

The following discussion describes the main control variables pertaining to each question addressed in the meta model.

3.1 School Quality Measures

Perhaps the most important variables are the school quality measures themselves. We take two alternative approaches to classifying the way these variables are used in various empirical studies. The first approach defines seven school different quality measures. Of these seven quality measures, three are output measurement variables based on standardized tests administered in elementary, middle or high schools. Our TEST SCORE variable indicates that a test score is used in the original study, TEST MEAN indicates that an average test score over a period is used typically as a measure of long term school quality, and VAL ADDED indicates that a value-added measure based on test score is used in the original study. The first approach also includes two input based performance variables, EXP which indicates per pupil expenditure is used as a variable in the original study and the indicator for student/teacher ratio S/T RATIO. The literature uses another variable often defined as an input based quality measure, peer effects. These typically reflect the racial or ethnic composition of the student body or some other measure of socioeconomic status. We indicate the presence of any of the group variables measuring the percentage of Hispanic, black or generally minority students by the MINORITY indicator variable. Our FREE LUNCH dummy variable indicates that the original model includes the percentage of students enjoying free or reduced lunches as a peer effects measure.

From Table 3.2, we see that the expenditure per pupil, long term school quality, and recent test score are often found to be capitalized into house prices, with average t-statistics greater than 3. The student /teacher ratio has a t-statistic implying 10% significance on average. Other school quality related measures—the percentage of students receiving free lunches, the percentage minority students, and value added measures of quality—do not appear to be regularly capitalized into house prices.

The second approach to identifying individual school quality measures categorizes quality measures into three broad categories: input, output, and peer effect variables. Table 3.2 reveals that, on average, input and output based quality measurement variables yield significant capitalization estimates while peer effect variables do not always yield significant capitalization estimates.

3.2 Neighborhood Effects

Strictly speaking, school quality is a neighborhood amenity. Researchers have struggled with how to separate school quality from other neighborhood effects. One way of controlling for non-school neighborhood effects is to include as many relevant neighborhood characteristics variables as possible. We use three variables to indicate results obtained by studies using this approach. The first dummy variable INCOME indicates that the original model includes neighborhood average income level; the second dummy variable TAX indicates that the original model includes the property tax rate; the third variable NUM NEIGH VAR indicates the number of other neighborhood control variables in the original study besides income or property tax rate. We expect to find that including more neighborhood variables lowers the significance of school quality capitalization as the additional variables control for correlated effects unrelated to school quality. We treat the income level and tax rate separately since one or the other or a combination of both is frequently used in the school quality capitalization studies. Table 3.2 shows that the inclusion of these two variables decreases the significance of school quality capitalization estimates.

The second way of controlling for neighborhood effects other than school quality is to introduce location control, either fixed effects or distance from job centers, into the hedonic price equation. We use the variable CBD DIST to indicate that the original study includes distance to CBD or similar employment center. Since the correlation between the distance of CBD and school quality are uncertain, we have no expectation regarding the effect of including in this variable in hedonic school quality models. An alternative way to control for neighborhood effects is to use house transactions that occur only within a narrow band next to school district or attendance zone boundaries and/or use boundary fixed effects to identify observations near specific school boundaries. We indicate these approaches with BOUNDARY, a dummy variable showing the data sample is school boundary sample or not, and BOUNDARY*BFE, showing the case when using boundary fixed effects are also used for the boundary subsample. Another method of controlling for neighborhood characteristics is to use local jurisdiction fixed effects when the sample of the original study crosses jurisdictional boundaries. We indicate this approach with the JURISDICT FIXED EFF indicator variable. We use the indicator NEIGH FIXED EFF for studies using census tract, subdivision, census block group or neighborhood fixed effects. Since census tracts, blocks and block groups are constructed to be relatively homogeneous, we also expect that including these types of neighborhood controls in an hedonic house price model reduces the size of the school quality capitalization, although the effect on significance is difficult to anticipate.

3.3 Sample Characteristics

Since the housing market is local and school evaluation systems have changed over time, we use three groups of control variables to representing these features of data samples in the meta-analysis.

The first group is based on US geographic regions using the Census Bureau definition of regions. There are four regions, or three control variables in each Meta regression in the pooled sample. From Table 3.2, in all four regions, the average effective t-statistics are greater than 2, showing that school quality significantly impacts house prices when these variables are included in the hedonic price model. The average effective t-statistic in the South (S) is the lowest, followed by Northeast (NE), Midwest (MW) and West (W). The output based school measures sample uses the single dummy variable SOUTH to indicate studies of locales in that region. There are few observations for the West and Midwest in this subsample and the average effective t-statistic for the South is considerably lower than the other three regions.

The second group of meta-analysis variables indicates the geographic level of the sample, whether state, MSA, county or city. We use two geographic level dummy variables in each meta regression. The STATE indicator is associated with unexpectedly large average capitalization t-statistics that do not appear to fit the pattern observed for other geographic level indicators. Since there are only three papers in which the data from an entire state are used, this summary of capitalization effective t-statistics may not be all that informative.

The third set of variables indicates the sample period. The earliest data used draws from 1930, but this is used in only in one study. Studies before 1980 usually use aggregative house values from census data, mostly with expenditure per pupil as the measure of school quality.

Therefore, we use 1980 as a time benchmark. We use 1999 as a second time period benchmark for two reasons. First, the United States inflation adjusted house price is relatively stable over 1980 to 1999 and not so much after 1999. Second, and more importantly, Black's (1999) paper signals a change in the direction of subsequent research. As a final note, although house prices in the US have dropped abruptly since 2007, there are only two papers using data after 2006 in our sample, and moreover, the degrees of capitalization significance reported in these two papers are very different. We group these two papers with the other studies using data after 1999.

Our meta-analysis using the entire or pooled sample identifies these three sample period indicator variables. In contrast, because there are very few observations in the output based measures subsample before 1980, the meta regression using the output based measures subsample identifies only two sample periods, before and after 1999. According to Table 2, for the pooled sample the average capitalization t-statistics indicate significance for all three periods, with the largest observed for regressions run using data from the 1980-1999 period.

Finally, researchers sometimes concentrate their studies on single family housing. Therefore, we also include the indicator SINGLE FAMILY to see whether this focus on one type of residential property can affect the capitalization significance. From Table 3.2, it does not appear that focusing on single family residences has much effect on the significance of capitalization effects in the original study.

3.4 Econometric methods

Five categories of meta-analysis variables relate to econometric methods. The first category indicates the method used to estimate the standard error. OLS represents the standard approach and is indicated by OLS in the meta model. OLS is sometimes used as a baseline model in papers that include estimates using other error estimation methods. The dummy variables ROBUST and CLUSTER indicate studies using robust and spatially clustered standard errors, respectively. Since both the clustered and robust errors tend to lead to larger error estimates relative to OLS, we expect that these two methods will dampen the significance of school quality capitalization. Table 3.2, however, reveals that the average capitalization t-statistics are larger when using clustered errors. This is not expected, but the meta regression analysis may lend further insight into this relationship as it controls for the simultaneous influence of the entire range of factors on the effective capitalization t-statistic.

The second category of meta variables indicates whether the study uses two stage least squares or not (2SLS). We include in this category all efforts to deal with endogenous right hand side variables in the hedonic model. This includes methods to distinguish the effects of school quality from those of missing neighborhood variables on house prices.

The third category in the meta regression indicates that the original study uses time fixed effects. The fourth and fifth categories are associated with using logarithmic forms for the dependent and independent variables, respectively.

3.5 School Levels and School Units

There are three school levels: elementary school, middle school and high school. Because attendance zones for the schools typically indicate that lower level schools feed into specific higher level schools, some capitalization studies include measures of school quality at different levels while others use aggregate measures across levels. The variable ALL indicates studies using a comprehensive set of school output based performance variables. Other indicators pertain to studies focusing on only one level of school. We construct another group of control variables to indicate whether the original study measures quality at the district level or for individual schools.

The results reported in Table 3.2 reveal that high school quality appears to exhibit weaker house price effects than elementary school quality. Further, Table 3.2 shows no great difference in capitalization conclusions whether measuring school quality for the district as a whole or for individual neighborhood schools.

In addition to these variables indicating how school quality is measured, we include DIST TO SCHOOL to pick up any effect from including the variable distance to school in the capitalization model and NUM SCH VAR to control the number of different school related variables in the original hedonic model. In our analysis of the subsample of studies relying primarily on output based measures of school quality, we also include controls for the presence of input based measures or peer effects variables in the model. Peer effects were discussed earlier. The corresponding dummy variables are INPUT and PEER, respectively

4. Meta-Analysis Results

The summary of capitalization t-statistics associated with the presence of the variety of variables identified in Table 3.2 provides an overview of simple association between the variables and the capitalization results. It does not control for the simultaneous influence of other variables in the set. The meta regression approach taken here does.

The effective t-statistic pertaining to the school quality capitalization coefficient in each regression equation reported in the studies included in our data set serves as the dependent variable in all of the meta-analysis regression models. The independent variables include the control variables explained in section 3. Table 3.3 reports the full set of meta-analysis estimates; the results in this table pertain to the pooled sample of empirical studies, pooling those that rely on output based school performance measures like test scores with those that use input based school performance measures like student-teacher ratios or expenditures.

Table 3.3 reveals that adding more school quality variables to the hedonic price model lowers the significance of individual school quality measures. One rationale for this result is that people evaluate public schools from different aspects so increasing the number of school quality attribute measures separates the individual contributions of each. Another reason is related to the researcher's regression strategy; when researchers do not find expected significant estimates, they tend to add new school related variables to the model.

Table 3.3 also shows that including the distance to school in the original model enhances the school quality capitalization significance. This is also reasonable since people may choose some houses in a lower education quality area because it is close to the school, the added convenience damping the effect of poor school quality. Controlling for distance to school will remove this factor's effect on location demand, leaving the underlying quality difference to be reflected in prices. These results suggest that the effects of being close to school in even in low quality school zones may be different from those in high quality school zones: buyers may tend to buy houses near schools when school quality is low, but not necessarily when school quality is high.

The performance of other indicator variables in the meta-analysis are discussed in the following sub-sections. For each set of variables we re-estimate the meta regression using different categorical variables as the omitted base case in order to make is easier to compare the relative effects of each without calculating differences in coefficients and testing for significance in differences.

4.1 Different School Quality Measures

In order to rank the significance of different school quality measures capitalization, we regress the effective t-statistic on three types of school measures (or seven major school quality measure variables) along with all other control variables as shown in Table 3.4. The key estimates for this sub-section are reported in Table 3.4. The results in Panel A show the relative importance of three types of school measures. Considering two models together, we can see that peer effect variables lead to less significant estimated capitalization effects when compared with input based measurements and output based measurements, when the influence of all of the other listed factors in Table 3.3 are taken into account. At the same time, though, there does not appear to a significant difference between the input based and output based quality variables.

Panel B presents more detail about these variables. Perhaps the most important is that one of the output based measures, value added based on test scores, leads to lower capitalization significance than all other school measures considered in these studies. In addition, it is clear that the peer effect measured by percentage of minority students leads to lower capitalization significance than expenditure per pupil, free lunch, and long-term and short term output based measures. Although the estimated coefficient for the student/teacher ratio is negative in the four meta regressions, not one is significant. This implies that the presence or absence of student/teacher ratio as an input based measure of school quality has no differential effect on capitalization significance across studies. On the other hand, although the short term output measure, recent test score, appears more important than other measures, the effect on house prices is not significantly different from expenditure per pupil, free lunch, long-term average of test score, or student/teacher ratio. The literature does not offer consistent evidence that one or more of the school quality measures dominates the others.

4.2 Separating School Quality from Neighborhood Characteristics

Recall that the first method we identified for separating the school quality and other neighborhood amenity effects is to add additional neighborhood demographic variables to the hedonic price model. Table 3.3 shows that all INCOME, TAX, and NUM NEIGH VAR coefficient estimates are negative, albeit none is significant. Apparently, adding neighborhood variables has no systematic effect on school quality capitalization estimates. If this strategy is used to remove additional neighborhood effects, either the neighborhood effects are largely uncorrelated with school quality effects or this method of control does not successfully distinguish the effect of better schools from other neighborhood features in house pricing.

As for location controls, Table 3.3 reports no significant estimate for CBD DIST; including this variable or excluding it from capitalization equations does not matter in the statistical sense. Turning to other methods, authors frequently use school attendance zone or school district boundary fixed effects to further control for neighborhood characteristics when using boundary subsamples. We see that the interaction effect of BOUNDARY*BFE is significantly negative at a level less than 1%. Clearly, boundary fixed effects can distinguish the school quality impact on houses prices from those of other neighborhood amenities: test score significance in capitalization studies are smaller with BFE than those without BFE in the hedonic regressions. The meta regression estimates in Table 3 also reveal that neighborhood fixed effects offers an effective location control method.

4.3 Sample Characteristics

We now look into the nature of the samples used in capitalization studies. First, of 45 papers, 12 papers consider only single family house sales data. From Table 3.3, it appears that relying solely on single family sales does not significantly influence school quality capitalization estimates.

Panel A of Table 3.5 illustrates how different geographic levels (MSA, COUNTY, or STATE) affect capitalization estimates. We compare the difference between the MSA and COUNTY. We do not see significant difference of school capitalization.

Panel B looks at how the different sample regions affect school capitalization. We find that data extracted from the South yields lower school capitalization significance than data extracted from the Northeast and West. The reason for this regional difference remains unclear. It may be that the long history of federal court involvement in public school systems in much of the South has shattered the public's confidence in both school performance and official reports regarding school performance. Or, it may be that regional differences in local governments affect the ease of real estate development in these areas, with the resulting elastic housing supply generating systematically lower capitalization effects. This rationale extends Brasington's (2002) rationale for Ohio to the regional context considered here. Testing this notion formally, however, lies outside the intended scope of this study.

Looking at Panel C, it is clear that early period data (before 1980) leads to lower capitalization significance than more recent data. Note that this result takes into account evolution in the data properties from aggregate to individual house transactions and changes in estimation methods that have occurred over time.

4.4 Different Econometric Methods

Table 3.6 reports the results of our analysis concerning how econometric methods influence capitalization estimates. Panel A shows that using robust error methods leads to lower capitalization significance than found when using OLS errors. We do not find significant differences between clustered and OLS error methods. Similarly, 2SLS yields no significant difference when compared with OLS. Including time fixed effects also does not change capitalization significance. As for using logarithmic dependent variables, we find no effect. Using logarithms for key independent variables, however, reduces capitalization significance.

4.5 School District versus Individual School

Table 3.7 reports the results comparing the level of school performance observation. Panel A shows that elementary school level measures lead to greater capitalization significance than middle school level measures or measures using averages of all levels of schools. Perhaps not surprising, most hedonic studies rely on elementary school level measures. (Elementary school measures are used in 30 of 45 papers and 202 of 374 observations.) Since studies tend to report significant results, the relatively stronger performance of elementary school measures identified here should not be surprising. Panel B, however, clearly shows no significant difference between using the school district average measures and individual neighborhood school quality measures.

4.6 Analysis of Subsample Using Output Based Measures of Quality

The premise underlying capitalization theory is that parents really care their children's academic achievement. This naturally leads to relying on output based measures of school quality using test scores or other student performance indicators. Nonetheless, educational inputs and peer effects influence student achievement. When the output based measures are not available or do not yield significant estimates, researchers typically turn to the two latter types of variables. In our sample, there are 210 observations relying on output based measures. Their prominence in the empirical literature means that they deserve special attention. Tables 3.8 to 12 repeat the meta-analysis for this output based measurement subsample.

The output based measures subsample requires that we modify the meta-analysis regression model. First, we use three indicator variables for school quality measures: TEST MEAN, TEST SCORE, and VALUE ADDED. Second, we include two dummy variables, PEER and INPUT, to control for the appearance of input based measurement and peer effect variables, respectively, in the original hedonic regression. Third, since the papers using data before 1980 mostly use the input based quality measure expenditure per pupil, there are very few observations before 1980 in the output based measurement subsample. As a result, we only use one time period partition for the output based measures subsample, studies before and after 1999. And finally, the distribution of observations across regions for this subsample only permits two regional distinctions, the South and all other regions.

Tables 3.8 to 3.12 reveal patterns consistent with those reported in Tables 3 through 7. In particular, adding more school quality variables into the house price regression equation dampens the significance of any one school quality variable's estimate. Using boundary fixed effects or neighborhood fixed effects reduces the significance of school quality capitalization as well. The significance of value-added quality measures is much lower than that of other output based quality measures. Finally, the South once again exhibits lower significance in school capitalization rates than other regions.

At the same time, the output based measurement subsample yields some results that differ from the pooled sample. Unlike the pooled sample, DIST TO SCHOOL and JURISDICT FIXED EFF are no longer significant. Including the tax rate in hedonic models now reduces the significance of school output variable capitalization. And unlike the pooled sample, we now find no differences in capitalization significance when using logarithmic transformations on school quality measures in hedonic models. Including high school output based quality variables in the original models increases capitalization significance more than measures for middle schools or all school level averages, something else not found in the pooled sample.

Finally, adding input and peer effect variables in hedonic regressions does not affect the significance of the output based quality variable capitalization on average. Moreover, we do not see significant coefficients of MSA and COUNTY in both pooled and output based samples.

5. Conclusion

The conclusions of the public school quality capitalization research vary considerably, not only across different studies, but also within individual studies. This paper takes a closer look at the likely sources of these variations, using formal meta-analysis to identify factors associated with variations in the significance of school quality capitalization estimates.

We define the meta regression effective size, the dependent variable in meta regressions, as the effective t-statistic in the hedonic models such that a positive coefficient on the indicator variable in the meta regression equation shows greater significance in the capitalization estimate in the expected direction in the original model. Our pooled sample uses 374 observations from 45 different papers spanning from 1968 to 2016. The output based quality measurement subsample uses 202 observations drawn from estimated models reported in 30 papers. We conduct meta-analysis for both the pooled sample from the results reported in all of the surveyed studies and the subsample of results pertaining only to output-based school quality measures.

The results show that the way school quality is measured matters. This is not surprising. Nonetheless, the specific patterns are interesting. Peer effect measures yield less significant capitalization estimates than either output based measures like test scores or input based measures like student/teacher ratios or expenditures. Within the three main output based performance measures, the value-added test score approach exhibits a much lower level of significance than either the short term or long run test scores.

Both boundary fixed effects and neighborhood fixed effects approaches appear to control for the impact of neighborhood amenities, thereby reducing capitalization significance. Surprisingly, econometric methods do not appear to be driving results. And even more surprising, using either school districts or neighborhood schools as the units of observation does not make a significant difference on quality capitalization estimates. The widely accepted notion that less aggregated data contains more information about housing market performance, the rationale underlying the shift to higher quality disaggregated data as it has become increasingly available, is not born out here.

Another unanticipated conclusion is that region matters. The school quality capitalization significance is much lower in the South than in other regions, for both of the pooled sample and the output based quality subsample. The reason for this result is unclear. It may be attributable to the legacy effects of cultural differences, a history of federal government control of local school systems, or Brasington's (1999) hypothesis that more elastic housing supply reduces capitalization effects since metropolitan areas in the South have been areas that have enjoyed rapid housing development. Although beyond the scope of this analysis, the reason underlying this persistent regional result deserves additional study, especially in light of the fact that the South as a region exhibits distinctive empirical results in other economic contexts as well, ranging from systematic differences is property law (Baker, et al. 2002) to the size of the state and local public sector (Turnbull and Salvino 2009).

The meta-analysis also suggests directions for additional study. First, while we focus on US housing markets, there is a growing literature looking at similar questions for housing markets in other countries. Differences in housing markets and government function across countries means that the modeling lessons found here need not transfer to other countries. While the body of literature for other countries does not yet appear to be large enough for separate formal meta-analysis, a less formal qualitative comparative study could provide additional guidance for that growing branch of the literature. For similar reasons, the small sample of research using post-2006 data means that we cannot yet systematically investigate the influence of the housing market crash on capitalization estimates. Since this was a period in which price discovery was clearly impeded, it would be useful to understand the extent to which crisis market conditions affect capitalization estimates in order to ascertain capitalization effects in the broad sense and stimulate theoretical consideration of the capitalization process in periods of market stress.

Second, the results reported here differ somewhat across the pooled sample and output based measures subsample. For example, we find very different coefficients for tax rate and distance to school across the two samples. It is unclear what is driving these differences.

In closing, we note that the meta-analysis yields some unexpected results that are immediately useful. For example, we do not find that using clustered errors reduces school quality capitalization significance, even though clustered error estimates are generally larger than their OLS counterparts. Also, econometric efforts to control for endogeneity do not seem to influence capitalization conclusions, an interesting lesson in light of the notable trend over the past decade toward relying more on instrumental variables methods in urban and real estate empirical modeling.

References

Bayer, P, F. Ferreira, and R. McMillan. 2007. A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *National Bureau of Economic Research* w13236.

Black, S.E. 1999. Do Better Schools Matter? Parental Valuation of Elementary Education. *Quarterly Journal of Economics* 114: 577-599.

Bogin, A. and P. Nguyen-Hoang. 2014. Property Left Behind: An Unintended Consequence of a No Child Left Behind "Failing" School Designation. *Journal of Regional Science* 54: 788-805.

Braden, J.B., X. Feng, and D.H. Won. 2011. Waste Sites and Property Values: a Meta-Analysis. *Environmental and Resource Economics* 50: 175-201.

Brasington, D. 1999. Which Measures of School Quality Does the Housing Market Value?, *Journal of Real Estate Research* 18: 395-413.

Brasington, D., and J.R. Haurin. 2006. Educational Outcomes and House Values: a Test of the Value Added Approach. *Journal of Regional Science* 46: 245-268.

Brueckner, J.K.1979. Property Values, Local Public Expenditure and Economic Efficiency. *Journal of Public Economics* 11: 223-245.

Brunner, E.J., J. Murdoch, and M. Thayer. 2002. School Finance Reform and Housing Values: Evidence from the Los Angeles Metropolitan Area. Public Finance and Management 2: 535-565.

Chiodo, A.J., R. Hernández-Murillo, and M.T. Owyang. 2010. Nonlinear Effects of School Quality on House Prices. *Federal Reserve Bank of St. Louis Review* 92: 185-204.

Clapp, J.M., A. Nanda, and S.L. Ross. 2008. Which School Attributes Matter? The Influence of School District Performance and Demographic Composition on Property Values. *Journal of Urban Economics* 63: 451-466.

Clark, D.E., and W.E. Herrin. 2000. The Impact of Public School Attributes on Home Sale Prices in California. *Growth and Change* 31: 385-407.

Clauretie, T.M., and H.R. Neill. 2000. Year-Round School Schedules and Residential Property Values. *Journal of Real Estate Finance and Economics* 20: 311-322.

Crone, T.M. 2006. Capitalization of the quality of local public schools: what do home buyers value?, Working paper, Federal Reserve Bank of Philadelphia,.

Cushing, B.J. 1984. Capitalization of Interjurisdictional Fiscal Differentials: An Alternative Approach. *Journal of Urban Economics* 15: 317-326.

Dougherty, J., J. Harrelson, L. Maloney, D. Murphy, R. Smith, M. Snow, and D. Zannoni. 2009. School Choice in Suburbia: Test Scores, Race, and Housing Markets. *American Journal of Education* 115: 523-548.

Debrezion, D., E. Pels, and P. Rietveld. 2007. The Impact of Railway Stations on Residential and Commercial Property Value: a Meta-Analysis. *Journal of Real Estate Finance and Economics* 35: 161-180.

Dhar, P. and S.L. Ross. 2012. School District Quality and Property Values: Examining Differences along School District Boundaries. *Journal of Urban Economics* 7: 18-25.

Downes, T.A. and J.E. Zabel. 2002. The Impact of School Characteristics on House Prices: Chicago 1987–1991. *Journal of Urban Economics* 52: 1-25.

Edel, M. and E. Sclar. 1974. Taxes, Spending, and Property Values: Supply Adjustment in a Tiebout-Oates Model. *Journal of Political Economy* 82: 941-954.

Gurwitz, A.S. 1980. The Capitalization of School Finance Reform. *Journal of Education Finance* 5: 297-319.

Gustely, H.D. 1976. Local Taxes, Expenditures and Urban Housing: A Reassessment of the Evidence. Southern Economic Journal 42: 659-665.

Harrison, D. and D.L. Rubinfeld. 1978. Hedonic Housing Prices and the Demand for Clean Air, *Journal of Environmental Economics and Management* 5: 81-102.

Haurin, D.R. and D. Brasington. 1996. School Quality and Real House Prices: Inter-and Intrametropolitan Effects. *Journal of Housing Economics* 5: 351-368.

Hayes, K.J. and L.L. Taylor. 1996. Neighborhood School Characteristics: What Signals Quality to Homebuyers?. *Economic Review-Federal Reserve Bank of Dallas*: 2-9.

Heinberg, J.D. and W.E. Oates. 1970. The Incidence of Differential Property Taxes on Urban Housing: A Comment and Some Further Evidence. *National Tax Journal* 23: 92-98.

Hyman, D.N. and E.C. Pasour. 1973. Real Property Taxes, Local Public Services, and Residential Property Values. *Southern Economic Journal*: 601-611.

Imberman, S.A. and M.F. Lovenheim. 2015. Does the Market Value Value-added? Evidence from Housing Prices after a Public Release of School and Teacher Value-added. *Journal of Urban Economics* 91: 104-121.

Johnson, M.S. and M.J. Lea. 1982. Differential Capitalization of Local Public Service Characteristics. *Land Economics* 58: 189-203.

Jud, D.J. and J.M. Watts. 1981. Schools and Housing Values. *Land Economics* 57: 459-470.

Jud, D.J. 1985. A Further Note on Schools and Housing Values. *Real Estate Economics* 13: 452-462.

Kane, T.J., D.O. Staiger, S. Samms, E.W. Hill, and D.L. Weimer. 2003. School Accountability Ratings and Housing Values [with Comments]. In *Brookings-Wharton Papers on Urban Affairs*: 83-137.

Kane, T.J., S.K. Riegg, and D.O. Staiger. 2006. School Quality, Neighborhoods, and Housing Prices. *American Law and Economics Review* 8: 183-212.

Li, M.M. and H.J. Brown. 1980. Micro-neighborhood Externalities and Hedonic Housing Prices. *Land Economics* 56: 125-141.

Longstreth, M., A.R. Coveney, and J.S. Bowers. 1984. Conservation Characteristics among Determinants of Residential Property Value. *Journal of Consumer Research* 11: 564-571.

McMillan, M. and R. Carlson. 1977. The Effects of Property Taxes and Local Public Services upon Residential Property Values in Small Wisconsin Cities. *American Journal of Agricultural Economics* 59: 81-87.

Meadows, J.R. 1976. Taxes, Spending, and Property Values: A Comment and Further Results. *Journal of Political Economy* 84: 869-880.

Nelson, J.P. 2004. Meta-Analysis of Airport Noise and Hedonic Property Values. *Journal of Transport Economics and Policy* 38: 1-27.

Nguyen-Hoang, P. and J. Yinger. 2011. The Capitalization of School Quality into House Values: A review. *Journal of Housing Economics* 20: 30-48.

Oates, W.E. 1969. The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis. *Journal of Political Economy* 77: 957-971.

Orr, L.L. 1968. The Incidence of Differential Property Taxes on Urban Housing. *National Tax Journal* 21: 253-262.

Pollakowski, H.O. 1973. The Effects of Property Taxes and Local Public Spending on Property Values: A Comment and Further Results. *Journal of Political Economy* 81: 994-1003.

Rosen, H.S. and D.J. Fullerton. 1977. A Note on Local Tax Rates, Public Benefit Levels, and Property Values. *Journal of Political Economy* 85: 433-440.

Ross, S. and J. Yinger. 1999. Sorting and Voting: A Review of the Literature on Urban Public Finance. In *Handbook of Urban and Regional Economics*, ed. Paul Cheshire and Edwin S. Mills 2. 191-200.

Schnare, A.B. and R.J. Struyk. 1976. Segmentation in Urban Housing Markets. *Journal of Urban Economics* 3: 146-166.

Schipper, Y., P. Nijkamp, and P. Rietveld. 1998. Why Do Aircraft Noise Value Estimates Differ? A Meta-Analysis. *Journal of Air Transport Management* 4: 117-124.

Sedgley, N.H., N.A. Williams, and F.W. Derrick. 2008. The Effect of Educational Test Scores on House Prices in a Model with Spatial Dependence. *Journal of Housing Economics* 17: 191-200.

Seo, Y., and R. Simons. 2009. The Effect of School Quality on Residential Sales Price. *Journal of Real Estate Research* 31: 307-327.

Sirmans, S.G., L. MacDonald, D.A. Macpherson, and E.N. Zietz. 2006. The Value of Housing Characteristics: A Meta-Analysis. *Journal of Real Estate Finance and Economics* 33: 215-240.

Sirmans, S.G., L. MacDonald, and D. Macpherson. 2010. A Meta-Analysis of Selling Price and Time-on-the-market. *Journal of Housing Research* 19: 139-152.

Smith, K.V. and J. Huang. 1995. Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models. *Journal of Political Economy* 103: 209-227.

Turnbull, G.K. and R.F. Salvino, 2009. Do Broader Eminent Domain Powers Increase Government Size?. *Review of Law & Economics* 5: 785-806.

Turnbull, G.K., V. Zahirovic-Herbert, and M. Zheng. 2016. The Capitalization of School Quality into House Prices when Quality is Uncertain: Theory and Empirical Evidence. Dr. P. Phillips School of Real Estate Working Paper 1601, University of Central Florida.

Walden, M. 1990. Magnet Schools and the Differential Impact of School Quality on Residential Property Values. *Journal of Real Estate Research* 5: 221-230.

Weimer, D.L. and M.J. Wolkoff. 2001. School Performance and Housing Values: Using Non-Contiguous District and Incorporation Boundaries to Identify School Effects. *National Tax Journal* 54: 231-253.

Yadavalli, A.P. and R.GMS Florax. 2013. The Effect of School Quality on House Prices: A Meta-Regression Analysis. 2013 Annual Meeting, Washington, DC, no. 151291, Agricultural and Applied Economics Association.

Zabel, J. 2015. The Hedonic Model and the Housing Cycle. *Regional Science and Urban Economics* 54: 74-86.

Zahirovic-Herbert, V., and G.K. Turnbull. 2008. School Quality, House Prices and Lquidity. *Journal of Real Estate Finance and Economics* 37: 113-130.

Zahirovic-Herbert, V., and G.K. Turnbull. 2009. Public School Reform, Expectations and Capitalization: What Signals Quality to Homebuyers?. *Southern Economic Journal* 75: 1094-1113.

Figures and Tables

Table 3.1: Effective capitalization t statistics in each paper

Paper ID	Frequency	Mean	Minimum	Maximum
Orr, 1968	1	1.470	1.470	1.470
Oates, 1969	2	2.100	2.100	2.100
Heinberg and Oates, 1970	2	1.500	1.200	1.800
Hyman and Pasour, 1973	1	0.500	0.500	0.500
Pollakowski,1973	6	2.089	0.667	3.826
Edel and Sclar,1974	5	1.240	-0.670	4.080
Gustely,1976	2	-0.802	-1.556	-0.048
Meadows, 1976	2	2.392	1.528	3.256
Schnare and Struyk,1976	1	5.280	5.280	5.280
Mcmillan and Carlson,1977	3	1.190	0.450	2.020
Rosen and Fullerton,1977	8	2.078	-0.390	3.125
Harrison and Rubinfeld, 1978	1	6.210	6.210	6.210
Brueckner, 1979	2	-1.332	-2.015	-0.649
Gurwitz, 1980	1	-1.250	-1.250	-1.250
Li and Brown, 1980	1	1.800	1.800	1.800
Johnson and Li, 1982	1	1.200	1.200	1.200
Cushing, 1984	1	1.430	1.430	1.430
Longstreth, Coveney, Bowers, 1984	1	5.440	5.440	5.440
Jud, 1985	6	1.447	-3.480	4.560
Walden, 1990	6	1.058	-0.471	2.442
Haurin and Brasington, 1996	1	5.778	5.778	5.778
Hayes and Taylor,1996	4	0.520	0.038	1.180
Black, 1999	4	4.461	1.429	8.750
Clark and Herrin, 2000	3	0.127	-5.410	3.640
Clauietie and Neill, 2000	4	2.898	-1.180	8.110
Weimer and Wolkoff, 2001	35	0.259	-2.333	2.704
Brunner, Murdoch, Thayer, 2002	4	7.500	4.000	10.000
Downes and Zabel, 2002	10	4.139	-1.340	12.765
Kane, et.al, 2003	17	3.467	-0.455	5.706
Brasington and Haurin, 2006	5	0.700	-18.800	14.200
Crone, 2006	66	2.545	-2.871	7.172
Kane, Riegg, Staiger, 2006	16	3.087	-0.172	7.219
Bayer, Fernando, McMillian, 2007	6	6.852	2.932	14.727
Clapp, Nanda, Ross, 2008	24	1.295	-8.160	9.650
Sedgley, Williams, Derrick 2008	6	3.683	1.000	8.600
Zahirovic-Herbert and Turnbull, 2008	5	5.394	4.000	9.175

Dougherty,et al., 2009	7	1.726	0.166	3.498
Seo and Simons, 2009	6	6.175	1.660	10.190
Zahirovic-Herbert and Turnbull, 2009	21	-0.041	-2.177	2.113
Chiodo, Murlllo, Owyang, 2010	3	4.050	1.780	7.790
Dhar and Ross, 2012	39	6.365	1.353	17.833
Imberman and Lovenhein, 2013	4	6.976	3.023	10.714
Bogin and Nguyen-Hoang, 2014	1	3.450	3.450	3.450
Turnbull,Zahirovic-Herbert, Zheng, 2016	24	2.791	-0.888	6.931

Note: Define the effective t-statistic as the t value reported in the study when reported coefficients have expected signs and minus one times the t-value when reported coefficients have signs opposite expected sign.

Table 3.2: Summary statistics for effective capitalization t-statistic observations

Groups	Variables	Frequency	Effective t-statistic			Paper # used
			Mean	Min	Max	
Panel A: School quality measures						
SCHOOL QUALITY MEASURES	EXP	68	3.409	-2.015	17.833	20
	FREE LUNCH	33	1.516	-2.222	8.110	6
	TEST MEAN	128	3.432	-8.160	14.727	10
	MINORITY	31	0.826	-6.640	5.976	7
	TEST SCORE	77	3.155	-3.480	14.200	23
	S/T RATIO	17	1.422	-5.410	6.210	6
	VAL ADDED	14	-0.275	-18.800	9.175	6
SCHOOL QUALITY MEASURE TYPE	INPUT	85	3.012	-5.410	17.833	26
	OUTPUT	210	3.225	-18.800	14.727	30
	PEER	73	1.051	-8.160	8.110	10
Panel B: Location controls						
CBD DIST	NO	222	3.061	-18.800	17.833	27
	YES	146	2.263	-3.480	7.357	21
BFE	NO	299	2.528	-18.800	17.833	44
	YES	69	3.681	-1.439	10.714	9
NEIGH FIXED EFF	NO	312	2.892	-18.800	17.833	312
	YES	56	1.921	-1.617	6.931	56
JURISDICT FIXED EFF	NO	214	3.240	-18.800	17.833	214
	YES	154	2.056	-2.871	10.000	154
Panel C: Data samples characteristics						
GOE LEVEL	MSA	64	2.407	-18.800	14.200	20
	COUNTY	240	2.400	-5.410	14.727	21
	STATE	64	4.372	-8.160	17.833	3
USA REGION	MW	31	3.518	-18.800	14.200	9
	NE	206	2.623	-8.160	17.833	19
	S	105	2.357	-2.177	9.175	11
	W	26	4.349	-5.410	14.727	7
DATA PERIOD	<1980	157	3.369	-5.410	17.833	17
	1980-1999	170	2.425	-18.800	14.200	13
	>1999	41	1.676	-2.015	6.210	18
SINGLE FAMILY	NO	291	2.894	-18.800	17.833	35
	YES	77	2.179	-2.333	10.000	11

Panel D: Econometric methods						
2SLS	NO	302	2.957	-18.800	17.833	38
	YES	66	1.772	-2.333	12.765	13
STD ERROR METHOD	CLUSTER	191	3.073	-8.160	17.833	14
	ROBUST	85	2.128	-5.410	7.172	4
	OLS	92	2.632	-18.800	14.200	28
DEP VAR	NON-LOG	51	2.188	-3.480	14.727	19
	LOG	317	2.834	-18.800	17.833	26
INDEP VAR	NON-LOG	313	3.094	-18.800	17.833	41
	LOG	55	0.754	-3.480	4.560	7
TIME FIXED EFFECT	NO	183	2.017	-18.800	14.727	29
	YES	185	3.464	-8.160	17.833	16
Panel E: School features						
SCHOOL LEVEL	ELEMENTARY	196	2.815	-3.480	14.727	29
	HIGH	43	1.872	-5.410	14.200	7
	MIDDLE	52	3.233	-8.160	12.765	5
	ALL	77	2.724	-18.800	17.833	22
SCHOOL UNIT	DISTRICT	203	2.804	-18.800	17.833	31
	SCHOOL	165	2.671	-2.177	14.727	19
Panel F: Other control variables						
INCOME	NO	201	3.016	-5.410	17.833	25
	YES	167	2.417	-18.800	14.200	26
TAX	NO	133	3.063	-5.410	14.727	18
	YES	235	2.564	-18.800	17.833	28

Table 3.3: Meta regression for pooled sample

Independent variable		Model 1	Independent variable		Model 1
INTERCEPT		3.936 (1.63)	DATA PERIOD	<1980	-2.101 (-1.33)
CBD DIST		0.539 (0.93)		1980-1999	0.100 (0.16)
DIST TO SCHOOL		1.808 (1.61)		>1999	
BOUNDARY		0.750 (1.22)	STD ERROR METHOD	CLUSTER	-0.730 (-0.82)
BOUNDARY* BFE		-2.277*** (-3.66)		ROBUST	-2.350** (-2.04)
NEIGH FIXED EFF		-2.047*** (-2.74)		OLS	
JURISDICT FIXED EFF		-1.312* (-1.9)	2SLS	NO	-0.673 (-0.96)
INCOME		-0.547 (-0.93)		YES	
TAX		-0.125 (-0.12)	TIME FIXED EFF	NO	-0.720 (-0.76)
SINGLE FAMILY		-0.461 (-0.61)		YES	
NUM NEIGH VAR		-0.153 (-1.49)	DEP VAR	LOG	1.164 (0.93)
NUM INDEP VAR		0.047 (0.74)		NON-LOG	
NUM SCHOOL VAR		-0.429*** (-2.92)	SCHOOL LEVEL	ELEM	1.962*** (2.92)
SCHOOL QUALITY MEASURE	INPUT	1.967*** (2.74)		HIGH	1.177 (1.38)
	OUTPUT	1.548*** (3.09)		MIDDLE	0.294 (0.34)
	PEER			ALL	
GEO LEVEL	MSA	-2.41* (-1.69)	INDEP VAR	LOG	-1.846** (-2.06)
	COUNTY	-2.032 (-1.11)		NON-LOG	
	STATE		SCHOOL UNIT	DISTRICT	-0.314 (-0.59)
USA REGION	MW	1.776* (1.89)		SCHOOL	
	NE	2.079** (2.43)	NUMBER OF OBSERVATIONS USED		368
	W	3.169*** (3.08)			
	S		ADJ. R-SQUARE		0.330

Notes: This table reports the full set of estimates from regressing effective t values on all control variables. Effective t is defined based on the t values of school quality variables extracted from the existing papers. When the expected t value of the school quality variables are positive, effective t has the same sign as original t; otherwise, the effective t has the opposite sign as the original. Independent variables are dummy variables except NUM NEIGH VAR, NUM INDEP VAR, and NUM SCHOOL VAR. *** p<0.01; ** p<0.05;* p<0.10.

Table 3.4: The relative importance among school quality measures

Panel A: The relative significance among three types of school quality measurements

Variable	Model 1		Model 2	
	coeff	t-stat	coeff	t-stat
INPUT	1.967***	2.74	0.419	0.70
OUTPUT	1.548***	3.09		
PEER			-1.548***	-3.09

Panel B: The relative significance among seven main school quality measurements

Variable	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
EXP			0.21	0.2	0.08	0.1	1.98*	1.7	-0.38	-0.4	0.88	0.7
FREE LUNCH	-0.21	-0.2			-0.13	-0.2	1.77**	2.3	-0.59	-0.9	0.67	0.7
TEST MEAN	-0.08	-0.1	0.13	0.2			1.90***	2.7	-0.46	-0.8	0.80	0.9
MINORITY	-1.98*	-1.7	-1.77**	-2.3	-1.90***	-2.7			-2.35***	-3.6	-1.10	-1.1
TEST SCORE	0.38	0.4	0.59	0.9	0.46	0.8	2.35***	3.6			1.26	1.4
S/T RATIO	-0.88	-0.7	-0.67	-0.7	-0.80	-0.9	1.10	1.1	-1.26	-1.4		
VAL ADDED	-4.85***	-4.4	-4.64***	-2.9	-4.77***	-3.0	-2.88*	-1.8	-5.23***	-3.4	-3.97**	-2.3

Notes: This table investigates the relative importance among different quality measures. The regression control variables are same as Table 3.3. Panel A reports results for three broad categories of quality measures, whereas Panel B reports results for seven school quality variables. The reference dummy variable categories differ across models for ease of comparison. The t statistics of the meta-analysis in parenthesis. *** p<0.01; ** p<0.05;* p<0.10

Table 3.5: The effect of samples on school quality capitalization

Panel A: Data geographic size effect on the significance						
Variable	Model 1		Model 2			
MSA	-2.41*	(-1.69)	-0.378	(-0.39)		
COUNTY	-2.032	(-1.11)				
STATE			2.032	(1.11)		
Panel B: The geographic location effect on significance						
Variable	Model 1		Model 2		Model 3	
MW	1.776*	(1.89)	-1.394	(-1.19)	-0.303	(-0.31)
NE	2.079**	(2.43)	-1.090	(-1.08)		
W	3.169***	(3.08)			1.090	(1.08)
S			-3.169***	(-3.08)	-2.079**	(-2.43)
Panel C: The data period effect on significance						
Variable	Model 1		Model 2			
<1980	-2.101	(-1.33)	-2.202*	(-1.65)		
1980-1999	0.100	(0.16)				
>1999			-0.100	(-0.16)		

Notes: This table investigates the effect of hedonic samples on the significance of school capitalization. The regression control variables are same as Table 3.3 Panel A concentrates on geographic sizes, Panel B concentrates on the geographic locations, Panel C is on the data periods. The reference dummy variable is different in different models for ease of comparison. *** p<0.01; ** p<0.05;* p<0.10.

Table 3.6: The effect of econometric methods on significance of school quality

Panel A: The compute methods of estimate error impact		
Variable	Model 1	Model 2
CLUSTER	-0.730 (-0.82)	1.620* (1.78)
ROBUST	-2.350** (-2.04)	
OLS		2.350** (2.04)
Panel B: The 2SLS method impact		
Variable	Model 1	
NO	-0.673 (-0.96)	
YES		
Panel C: The time fixed effect		
Variable	Model 1	
NO	-0.720 (-0.76)	
YES		
Panel D: The form of dependent variable impact		
Variable	Model 1	
LOG	1.164 (0.93)	
NO LOG		
Panel E: The form of independent variable effect		
Variable	Model 1	
LOG	-1.846** (-2.06)	
NO LOG		

Notes: This table investigates the effect of econometric methods on the significance of school capitalization. The regression control variables are same as Table 3.3. Panel A concentrates on the calculation of estimate error, Panel B concentrates on the effect of 2SLS application, Panel C concentrates on the effect of time fixed effects, and the last two are, respectively, using logarithmic form for house price and school quality variables. The reference dummy variable is different in different models for ease of comparison. *** p<0.01; ** p<0.05;* p<0.10.

Table 3.7: The effect of school characteristics on school quality capitalization

Panel A: The different impact of elementary school, middle school, and high school on significance						
Variable	Model 1		Model 2		Model 3	
ELEMENTARY	1.962***	(2.92)	1.906**	(2.28)	0.787	(1.32)
HIGH	1.177	(1.38)	1.118	(1.16)		
MIDDLE	0.294	(0.34)			-1.118	(-1.16)
ALL			-0.056	(-0.06)	-1.174	(-1.38)
Panel B: The different impact of district level education quality measurements and school level measurements						
Variable	Model 1					
DISTRICT	-0.314	(-0.59)				
SCHOOL						

Notes: This table investigates the effect of school levels and school units on the significance of school capitalization. The regression control variables are same as Table 3.3. Panel A concentrates on four different levels, and Panel B concentrates on two different school units. The reference dummy variable is different in different models for ease of comparison. *** $p<0.01$; ** $p<0.05$;* $p<0.10$.

Table 3. 8: Meta regression for output based quality measures subsample

Independent variable		Model 1	Independent variable		Model 1
INTERCEPT		-0.790 (-0.19)	DATA PERIOD	BEFORE 2000	0.007 (0.01)
CBD DIST		0.014 (0.01)		AFTER 1999	
DIST TO SCHOOL		-0.135 (-0.09)	STD ERROR METHOD	CLUSTER	0.634 (0.49)
BOUNDARY		-0.437 (-0.61)		ROBUST	-3.316** (-2.01)
BOUNDARY* BFE		-1.672** (-2.30)		OLS	
NEIGH FIXED EFF		-2.299** (-2.20)	TSLS	NO	-0.240 (-0.21)
JURISDICT FIXED EFF		0.080 (0.08)		YES	
INCOME		0.291 (0.32)	TIME FIXED EFF	NO	-1.710 (-1.31)
TAX RATE		-1.950 (-1.23)		YES	
SINGLE FAMILY		-1.381 (-1.16)	DEP VAR	LOG	1.008 (0.70)
NUM NEIGH VAR		-0.064 (-0.45)		NON-LOG	
NUM INDEP VAR		0.010 (0.11)	INDEP VAR	LOG	-2.484 (-1.43)
NUM SCHOOL VAR		-0.448** (-1.99)		NON-LOG	
PEER		-0.440 (-0.53)	SCHOOL LEVEL	ELEM	-4.549* (-1.77)
INPUT		0.438 (0.57)		HIGH	-4.896* (-1.89)
SOUTH		-3.735*** (-3.32)		MID	-0.655 (-0.32)
SCHOOL QUALITY	TEST MEAN	8.924*** (2.94)		ALL	
	TEST SCORE	8.889*** (3.37)	SCHOOL UNIT	DISTRICT	0.170 (0.23)
	VAL ADDED			SCHOOL	
GEO LEVEL	MSA	3.167 (1.41)	NUMBER OF OBSERVATIONS USED		210
	COUNTY	5.310* (1.67)			
	STATE		ADJ. R-SQUARE		0.418

Notes: This table shows a full set of estimates from regressing effective t values on all control variables. The dependent variable effective t is defined based on the t values of school quality variables extracted from the existing papers. When the t value of the school quality variables are positive, effective t has the same sign as original; otherwise, the effective t has the opposite sign as the original. Independent variables are dummy variables except NUM NEIGH VAR, NUM INDEP VAR, and NUM SCHOOL VAR. *** p<0.01; ** p<0.05;* p<0.10.

Table 3.9: Relative importance among school quality measures for output quality measure subsample

Variable	Model 1		Model 2	
TEST MEAN	8.924***	(2.94)	0.036	(0.04)
TEST SCORE	8.889***	(3.37)		
VAL ADDED			-8.889***	(-3.37)

Notes: This table investigates the relative importance of different quality measures. The regression control variables are the same as Table 8. The reference dummy variable is different in different models for ease of comparison. *** p<0.01; ** p<0.05;* p<0.10.

Table 3. 10: Sample effect on school quality capitalization for output measures subsample

Panel A: Data location size effect on significance				
Variable	Model 1		Model 2	
MSA	3.167	(1.41)	-2.143	(-1.22)
COUNTY	5.310*	(1.67)		
STATE			-5.31*	(-1.67)
Panel B: The significance difference in different regions				
Variable	Model 1			
SOUTH	-3.735***	(-3.32)		
NON SOUTH				
Panel C: The data period effect on significance				
Variable	Model 1			
BEFORE 2000	0.007	(0.01)		
AFTER 1999				

Notes: This table investigates the effect of hedonic samples on the significance of school capitalization. The regression control variables are the same as in Table 3.8. Panel A concentrates on geographic sizes, Panel B concentrates on the geographic locations, Panel C concentrates on sample periods. The reference dummy variable is different in different models for ease of comparison. *** $p<0.01$; ** $p<0.05$;* $p<0.10$.

Table 3.11: The effect of econometric models on school quality capitalization in output measures subsample

Panel A: The compute methods of estimate error impact		
Variable	Model 1	Model 2
CLUSTER	0.634 (0.49)	3.950*** (2.75)
ROBUST	-3.316** (-2.01)	
OLS		3.316** (2.01)
Panel B: The two step least square method impact		
Variable	Model 1	
NO	-0.240 (-0.21)	
YES		
Panel C: The time fixed effect		
Variable	Model 1	
NO	-1.710 (-1.31)	
YES		
Panel D: The form of dependent variable impact		
Variable	Model 1	
LOG	1.008 (0.70)	
NO LOG		
Panel E: The form of independent variable effect		
Variable	Model 1	
LOG	-2.484 (-1.43)	
NO LOG		

Notes: This table investigates the effect of econometric methods on the significance of school capitalization. The regression control variables are same as Table 3.8. Panel A concentrates on the calculation of estimate error, Panel B concentrates on the effect of 2SLS application, Panel C concentrates on time fixed effects, and the last two are, respectively logarithmic form for house price and school quality. The reference dummy variable is different in different models for ease of comparison. *** $p<0.01$; ** $p<0.05$;* $p<0.10$.

Table 3.12: The effect of school characteristics on school quality capitalization for output measures subsample

Panel A: The different impact of elementary school, middle school, and high school on significance						
Variable	Model 1		Model 2		Model 3	
ELEMENTARY	-4.549*	(-1.77)	-3.893***	(-2.64)	0.348	(0.55)
HIGH	-4.896*	(-1.89)	-4.241***	(-2.77)		
MIDDLE	-0.655	(-0.32)			4.241***	(2.77)
ALL			0.655	(0.32)	4.896*	(1.89)
Panel B: The different impact of district level education quality measurements and school level measurements						
Variable	Model 1					
DISTRICT	0.170	(0.23)				
SCHOOL						

Notes: This table investigates the effect of school levels and school units on the significance of school capitalization. The regression control variables are same as Table 3.8. Panel A concentrates on four different levels, and Panel B concentrates on two different school units. The reference dummy variable is different in different models for ease of comparison. *** p<0.01; ** p<0.05;* p<0.10.