Essays on Applied Microeconomics

A dissertation presented

by

Hoan Soo Lee

to

The Department of Business Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

April 2013

Dissertation Advisors:
Professor Josh Lerner
Professor Oliver Hart

Author:
Hoan Soo Lee

Essays on Applied Microeconomics

Abstract

Empirical and theoretical topics in applied microeconomics are discussed in this dissertation. The first essay identifies and measures managerial advantages from access to high-quality deals in venture capital investments. The underlying social network of Harvard Business School MBA venture capitalists and entrepreneurs is used to proxy availability of deal access. Random section assignment of HBS MBA graduates provides a key exogenous variation for identification. Being socially connected to peer venture capital firms and private equity seeking startups leads to more deal flow, larger asset under management and better performance in the inaugural funds of HBS-executive run venture capital firms. The second essay presents a two-stage model of competing ad auctions. Search engines attract users via Cournot-style competition. Meanwhile, each advertiser must pay a participation cost to use each ad platform. Advertiser entry strategies using symmetric Bayes-Nash equilibrium that lead to the Vickrey-Clarke-Groves outcome of the ad auctions are derived. Consistent with the model predictions, empirical evidence shows that multi-homing advertisers are larger than single-homing advertisers. Comparative statics on consumer choice parameters, quality, and user welfare are used to analyze the prospect of joining auctions to mitigate participation costs. The analysis provides conditions when such joins do and do not increase welfare. The third essay develops and computes a dynamic model of search in internet advertising. Micro-level browsing data from Microsoft's Bing.com (formerly known as Live.com) is used for structural estimations. The model predicts that users do not click on any ad with weak signals due to accumulating search cost and monotonicity of the value function. Rational search reveals a cascading pattern: the user clicks on a sufficiently high, highest-signal ad first, then moves on to the ad with the next highest conditionally expected probability of match once his assessment on the current ad degrades over time. The user exits when maximum assessment of likelihood of match over all ads is below a threshold value. The essay provides a novel approach to understanding rational herding behavior when product quality is only partially unraveled.

Contents

List of Tables

List of Figures

Acknowledgments

I am deeply grateful to my dissertation committee: Josh Lerner (Chair), Lauren Cohen, Paul Gompers, Oliver Hart and Matt Rhodes-Kropf. I am also grateful to Susan Athey, Andrei Shleifer, and Jeremy Stein for their invaluable guidance and encouragement. I also thank Philippe Aghion, Shai Bernstein, Josh Coval, Thom Covert, Ben Edelman, Robin Greenwood, Ben Iverson, Chris Malloy, seminar participants at the Harvard Finance Lunch and Harvard Organizational Economics Lunch for helpful comments and suggestions. Jinzhao Wang and Daniel Zhao provided superb research assistance. I am grateful for funding from the Harvard Business School Doctoral Fellowship and data from Microsoft Research New England.

To my father and mother, Yong Soon Lee and Kyung Hee Lee

Chapter 1

Value Access in Venture Capital[1]

1.1 Introduction

What makes venture capital successful? Determinants of venture capital performance is of particular interest in financial economics because private equity (including venture capital) is known to be a unique asset class in which market outperformance persists up to three consecutive funds ([26]). Unlike public equities that transact on exchanges, venture capital is a complex two-sided matching problem between entrepreneur and venture capitalist. A critical aspect of matching markets is the relative supply of (and demand for) each side of the matching transaction. While there is anecdotal evidence on the importance of access to ex-ante high-NPV entrepreneurs to explain much of cross-sectional variations in venture capitalists' performance, there is no good empirical evidence. I test the hypothesis that the well-connected venture capitalist may be successful by attaining access to great deals.

There are two significant empirical challenges. A measure for availability of access to deals which is a latent variable, is difficult to design. Second, endogenous managerial success must be disentangled properly, since access to great deals may be a direct function of superior talent. I address the first challenge by using social connections as proxy for availability of access. Social connections are instrumental in providing access to deals in an environment with limited investment opportunity and information asymmetry.[2] [3]

[1]Job-market paper

[2]I use access limitation to refer to cost of search, which can be prohibitively high, even if there are many more small businesses. In terms of raw supply, there were 5,294,970 businesses with less than 20 employees in 2008 (US Census Bureau).

[3]It may be argued that even in public equity markets, managers with different information advantage (insider information) are able to view a meaningful subset of opportunity set. However, I distinguish between access and evaluation of investments. If private

In particular, investors of public market equities all have access to every stock listed on a public index. Whereas in private markets, no such listings exist. Moreover, information asymmetry is a major hurdle in private markets due to non-mandatory disclosure of company statements, and especially in early-stage investments with non-existent track-records of business performance. Hence access to deals are often gained through syndications, or groups of co-investors, formed within trusted social contacts to diversify information signals.

To address the second obstacle, I employ HBS random section assignment to provide the necessary exogenous variation in social connections. The administrators of HBS use carefully surveyed profile of each student to assure a uniform distribution of various characteristics ([28]). I identify shares of section peers of an HBS executive who are venture capitalists and entrepreneurs as measures for availability of indirect and direct access to deals, respectively.

In the first set of results, I consider a limited sample of inaugural funds raised between 1980 and 2011 by US-based venture capital partnerships so that the effects are free from selection on observable past track-records. Also, if past track-record is a substitute for the benefits of social capital, its potential premium on various determinants of fund success will be presumably the strongest in the early stage of the firm ([6]).

In the baseline specification, I test to see if the share of peers is positively correlated with the number of syndication partners. If VC firms tend to syndicate with socially connected peers then the regression coefficient on the share should be positive and significant. Moreover, I restrict the definition of syndication partners in two ways following [22] for measures of *direct* and *indirect* access. First is a syndication partner who was invited into a round by the VC firm acting as a lead investor (outdegree); and second is a syndication partner who is the lead investor and invited the VC firm into a round (indegree). I show that a firm's outdegree is positively correlated with the share of section-peer entrepreneurs (EPs) and a firm's indegree is positively correlated with the share of section-peer venture capitalists.

To estimate economic implications, I compute the effect of knowing peers on the size of inaugural funds. The estimates translate to \$56.32M per company headed by a peer entrepreneur and \$33.79M per peer VC-firm unconditionally, or \$21.09M and \$7.43M, given the average size of HBS funds of \$100M. In order to test the idea that fund partners are able to raise larger funds by advertising their social-ties to limited partners, I run an OLS specification with a proxy for maturity of social contacts. I interact share of section/class-peer VCs/EPs with the age of HBS partners (difference between the average graduating class years of all partners in an HBS firm and the vintageyear of its inaugural fund). Results show that all

equity investors retain any information advantage beyond exclusive access, this would yield further heterogeneity in outcomes.

coefficient estimates on share of peers in VC and entrepreneurship are positive and significant, supporting the idea of raising larger funds by advertising established contacts.

Next, I examine investment performance to understand the effects of social connections on deal outcome. I show that the performance of section-peer funds on deals done prior to the inaugural fund's vintage year, has a positive effect on fund performance. Results show that one standard deviation higher section-peer performance leads to 6.3% higher likelihood of IPO to 20.79% from its baseline performance of 14.49% in the absence of peer activity, on deals exposed by HBS-firms through their inaugural funds. In economic terms, one standard deviation higher section-peer performance in the previous period leads to an estimated 22.16% increase in inaugural fund size, or $28.48M.

Lastly, I test three candidate hypotheses to better understand the mechanism underlying performance premium associated with either syndicating or financing a section peer. One possible explanation of the effect of social interactions on managerial success is ability screening. Suppose peers learn through the course of two years in HBS the underlying abilities of each section mate. By selection, each section will discriminate high and low talent managers in decisions to syndicate (VC-to-VC) or finance (VC-to-EP). Then a deal syndicated between non-section peers will be a match between low talent managers, possibly yielding the results in which likelihood to IPO is systematically higher by decreasing social distance. Additionally, if ability screening is the primary function of peer interaction during the MBA years, one should expect to observe section exclusivity to weaken and eventually disappear in the follow-on funds as market participants will be able to condition on observable past track-records. Regression results however suggest that section exclusivity and the associated marginal effect on likelihood to IPO remain robust to inclusion of all follow-on funds as of 2011.

Second explanation is that social-ties reduce cost of search and coordination such that the production function is enhanced among peers. Third hypothesis is that the most connected VCs receive privileged deal flow. Distinguishing between the two possibilities is difficult when using the share of portfolio companies that eventually go public as a performance measure. Production function improvement idea would suggest that conditional on accessing a deal, venture capitalists add value beyond providing access to capital through advice and other resources by working more harmoniously with peers. The access idea would suggest that best friends offer access to the best deals, and that investing into high-quality entrepreneurs drives the results of high IPO exit rates.

Using the average number of startups founded (unconditional on exit status) by c-level executives and founders of portfolio companies prior to the first investment round a la [17], I control for ex-ante deal quality. I regress deal quality on VC-to-EP financing matches and VC-to-VC syndications. Results show a systematically increasing positive and significant coefficient estimates with respect to social proximity in connections. This suggests that peers endow venture capitalists with access to better deals. Peer VCs may be able to sustain a stronger dynamic relationship that forms a proprietary deal flow. Indeed, I find evidence that leading investment rounds is associated with being invited to future rounds. [4]

Then I examine deal outcomes at the individual round-level, in order to distinguish between peer syndications formed simultaneously in a single round, and syndications formed across a sequence of rounds. Syndications of the latter form render an early investor who is presumably the leader providing access to the later investor. If the production function enhancement hypothesis is correct, then both forms of syndication should be positively and significantly correlated with deal success. Results however show that simultaneously formed peer syndications are not significantly correlated. Rather, estimates indicate that following an early section- or class-peer investor is positively and significantly correlated with deal success. Using this evidence, I conclude that deal access hypothesis is more likely true.

Lastly, in order to measure the persistence of the effects detected in inaugural funds, I consider an inclusive panel data set of all 3,799 funds raised between 1980 and 2007, and measure deal flow and performance across 10-year lifecycles. When considering follow-on funds, market participants are able to condition on observable track-record of managers, hence the previous exercise of using the raw number of peer firms will fail to identify the effect in question. To address this, I use data on peers working in the same sector as the fund's focus but in non-investment roles varied by residential proximity to tease out effects of social connections on deal flow. Results show that those targetting Computer Software & Internet receive positive and significantly related deal invitations, whenever the firm is surrounded by non-investing peers in the same state.

This research provides insight on deal flow, syndications and performance as functions of the underlying peer network in venture capital. Random section assignment of HBS MBA venture capitalists and entrepreneurs is used to identify and measure the economic value of social connections. Results suggest that the well-connected venture capitalist may be successful by attaining access to great deals. The effect is identified in key aspects of the inaugural fund, and appears to be persistent in the follow-on funds.

The paper is organized as follows: section II provides literature review. Section III elaborates on the econometric model used for measurement. Section IV describes the data set and provides summary

[4][27] show that the business product is more correlated with success than the quality of management. However, since the underlying idea associated with the business product is a transportable good, the original founder of the company may remain as the primary source of success.

statistics. Section V gives results, section VI provides extensions and section VII concludes. All tables and figures are available in the appendix.

1.2 Literature Review

[22] examine venture capitalists' underlying investment syndication network structure to show that there is strong positive correlation between performance and a variety of measures of the VC's network-connectedness. Authors find that VC networks, formed through past and on-going syndication partnerships, add value to investment success rather than provide privileged access to good deals. In my paper, I use networks formed through the years executives spent at Harvard Business School and suggest that HBS alumni benefit by accessing the best deals.

Several papers have attempted to use and test social-ties as a booster of performance. In each instance, authors have used same-degree holding as a link in either syndications or entrepreneur-financing matches. However, within-school heterogeneity and selective sampling of visible venture data pose significant hurdles to establishing proper identification. Moreover, the question of how social connections can achieve strong performance is not addressed in the analysis, whereas I posit that access available through social ties to ex-ante high-NPV entrepreneurs drives success.

[6] shows that general partners holding the same school degrees tend to syndicate more often and the associated IPO exit rates tend to be higher. Two major challenges in the paper are establishing identification given the same school degree measure, and addressing selection bias in surviving funds and startups observed in the sample. Measuring social-tie as a pair holding the same school degree does not rule out unobserved heterogeneity (across class and within class shocks) that may prevent establishment of a causal link.[5] Notably, managerial ability is difficult to control for the econometrician but most likely is the key confounding factor of the effect in question. Suppose some latent data regarding a manager is observable to other investment partners (managerial philosophy, past employment history, knowledge of the industry, etc) but unobservable to the econometrician and weakly correlated with any observable characteristics (school attended).[6] Then a selective subsampling of data - dropping Harvard/Stanford or top 5 schools - for robustness checks will be insufficient in ruling out ability-driven results. Author

[5] As part of a robustness check, author uses same school, same degree and similar age (within 4 +/- years for undergraduate and 6 +/- years for graduate degree ties).

[6] Although [18] reveal that holding a top business school degree or top college degree is positively correlated with likelihood to IPO, academic degree may remain as a noisy measure of ability.

presents evidence that school-tie premium diminishes after the first syndication relationship arguing against the talent hypothesis; but the change in effect may be driven by substituting either access to good deals or reduction in coordination cost with a proven track-record, encouraging former syndicating partners to become independent thereby improving the quality of unconnected syndications.[7] The value of establishing a strong track-record will be especially important for the sample of same-school, first-time syndications as such VC partnerships will constitute inaugural funds. Moreover, if same school matching is a result of ability-based screening, the measure of school connections will be more precise than dividing Harvard/Stanford and the rest, or top-5 and the rest, as alternative ability measures considered by the author in robustness checks.

Second, publicly available venture capital data is subject to sample selection. Both failed VC-backed companies and poorly performing funds will not be visible in the data. In contrast to the previous challenge of identification, if ability is actually correlated with the ranking of school attended by managers, lower-ranked schools will be over-represented by high-ability managers from such schools in the observed data, yielding an upward-biased measurement of school-tie effects. In my paper, I address both challenges of controlling for managerial ability and sample selection using balanced section assignment. Balanced section implies that talent is balanced across sections, hence the distribution of section graduates will be uncorrelated with ability and remain unchanged in the observable sample of surviving funds. Also, author focuses only on VC-to-VC syndication ties but in this paper, I also consider VC-to-EP financing ties.

[18] show VC-VC syndications with top-school degrees perform better on average. But syndications formed based on same ethnic backgrounds, same school (top or not), or worked for the same previous employer, which is a different measure of social connection (affinity links), tend to perform worse.

[36] considers both VC-to-VC syndications and VC-to-EP financing ties as well, similarly in attempt to measure the effect of social-ties on syndication propensity and performance. The measure of social-tie is limited to same-school degrees, and in this case, without distinguishing undergraduate or graduate degree or class year. In particular, the institution producing the most number of entrepreneurs and the third highest number of VCs is University of California (i.e. Berkeley, UCLA, UCSD and all other seven campuses are clustered as one unique institution). Hence in one extreme, a VC who received his undergraduate degree from UC-San Diego from class of 1980 investing into a startup run by an MBA-graduate of UC-Berkeley from class of 2003 will be presumed to be socially-tied in the same sense as a pair of class-peers from MIT.

[7] As an example, an inexperienced Stanford (Harvard) manager may begin by syndicating with another Stanford (Harvard) manager, successfully exit deals, then subsequently syndicate with a Harvard (Stanford) manager.

In addressing the concerns of sample selection bias, author follows a [20] two-stage selection model. Joint distribution of companies and funds across states is used as an instrument in the first stage following [7]. However in BDH, locational heterogeneity across different countries in Europe as a source of exogenous variation in company-investor match process is used. An adaptation of such method in the US across states will be problematic as funds are often endogenously agglomerated near San Francisco, California and Boston, Massachusetts.

In consideration of differences and commonalities in the previous works, one can notice a trade-off between establishing a well-defined measure for social-tie and the available breadth of data cross-section to support a meaningful estimate. While [36] provides an ambiguous measure of social-tie, author examines board-level connections among venture capitalists. [6] and I aggregate partners over the firm-level while providing improved measure of social-ties. Firm-level measures allow the possibility of falsely viewing a relationship link between Stanford partners as a link between HBS partners in a firm dominated by Stanford alumni. However, such upward bias in estimates do not exist between entrepreneur and VC connections since the number of CEO/founder in startups are limited to a few. Also, differential marginal effect on deal success when varying social proximity among VC-to-VC connections suggest that non-HBS link bias in estimates are unlikely driving the results.

[5] and [21] also study VC-to-EP financing process but ask different questions. [35] formulates a structural two-sided matching model of VCs and EPs but focuses on different questions.

[10] show similar results among mutual fund managers. Managers holding stocks of companies with CEOs or board members from the same school retain a performance premium, presumably due to information diffusion. My work resembles the result and methods employed in CFM but I apply the hypothesis to venture capital and also consider VC-to-VC syndications.

The proposed identification technique of using HBS random section assignment is not new but it is a novel application on answering a unique question, in the context of venture capital. [28] use a sub-sample of the same section data to show that graduates are more likely to pursue entrepreneurship if more students with previous entrepreneurial experience are present in the same section. [34] performs a similar exercise to show that executive compensation is influenced by peer effects. To address section-level common shocks, Shue employs cyclical reunions as a source of exogenous variation. I apply the data to identify and measure the effect of social connectedness on deal flow, fund size and performance in venture capital; and use ex-ante characteristics of section mates as instrumental variables to control for potential section-level common shocks.

1.3 Econometric Model

I now lead the discourse onto the particular empirical methodology that employs random section assignment to measure effects on inaugural funds raised by HBS executives for identification. A restricted sample of inaugural funds are used for two reasons: many factors that are influenced by observable measures of VC ability (number of syndication partners, access to quality of deals, sectoral targeting, fund size, etc) are normalized given the absence of observable fund track-records. Second, if past track-record is a substitute for the benefits of social capital, its potential premium on various determinants of fund success will be presumably the strongest in the early stage of the firm.

I specify two baseline models.

$$D_n = \alpha X_n + \gamma PeerTie + s_n + k_t + \epsilon_n \tag{1.1}$$

In the first version, each observation is at the unique deal-level. D_n is equal to 1 if the company eventually exits via IPO as of December 2011, and 0, otherwise. X_n includes deal-level vector of observavable characteristics (average number of co-investors per round, types of co-investors, average fund size, etc) and *PeerTie* is an indicator variable equal to 1 if there exists at least a pair of co-investors (or a co-investor and company executive pair) that are socially connected. The closest form of social connection is between section-peers, next is between class-peers, then between school-peers. I am interested in the coefficient estimate $\hat{\gamma}$ under various specifications.

In the second version, each observation is at the unique fund-level. It is defined as follows:

$$Y_{it} = \alpha_0 X_{it} + \beta_0 V_{it} + \gamma_0 s_{it} + u_i + k_t + \epsilon_{0it} \tag{1.2}$$

Y_{it} is the dependent variable of interest describing some dimension (number of unique syndicating partners, size of the inaugural fund, etc) of firm i at year t. X_{it} is the firm-level vector of observable characteristics including graduation year fixed effect, V_{it} include industry-wide observable components, s_{it} is the number of section-peer GP firms (normalized by section or class size) of firm i active in year t. u_i and k_t are firm and year fixed effects. I would reject the null hypothesis that section-peer firms do not influence outcome measures of firm i if γ is statistically different from zero. The extent to which peers actually influence the investment strategies of fund managers may go beyond syndications, deal access or ability to raise funds. Peers may influence an investment banker to pursue venture capital, or target a particular sector. As such, estimations of particular effects of social connections that operate through syndications and direct financing in venture capital represent an underestimation of the broader peer influence.

It should also be noted that there is a significant gap in number of years between the dependent variables and the explanatory factors that proxy social interactions from MBA years. Firm fixed effect u_i is included to control for interim-period idiosyncratic shocks. Section-level common shocks are addressed in the following.

Suppose the HBS executives in firm i from class of 1998 section C formulated similar projections regarding internet and software industries that induced several section-mates to pursue venture capital upon attending lectures by a famed professor. This would represent some latent shock z_{it} in the error term such that $E[\epsilon_{0it}s_{it}] \neq 0$. A section-level common shock during the MBA years can bias the estimate in question, or worse, yield spurious results in the absence of any causal peer factors.

I approach this problem by using ex-ante observable characteristics of section-peers as instruments and estimate γ using second-stage least squares (2SLS). Consider the following first-stage specification:

$$s_{it} = \alpha_1 X_{it} + \beta_1 V_{it} + \delta w_i + u_i + k_t + \epsilon_{1it} \tag{1.3}$$

w_i is a vector of two key instrumental variables for s_{it}. The first is mean share of U.S. News & World Report top-10 colleges[8] attended by the section-peers of all HBS executives of firm i.[9] The second is mean share of males among section-peers of all HBS executives of firm i. The first measure is strongly correlated with the number of section-peer firms as all schools in the top-10 with the exception of Columbia, University of Chicago and Caltech also represent the top suppliers of HBS venture capitalists. The second measure will also yield strong correlation as males constitute the majority of HBS VC executives (72%). No intermediate shocks during the MBA years have effected them, because both measure ex-ante characteristics of section-peers. Moreover, since HBS administrators actively balance the sections conditional on observable characteristics of an incoming class (i.e. undergraduate institution attended and gender) in attempts to ensure equitable experience across sections for each individual, such factors are unlikely to cause any particular shocks during the MBA years or to be systematically correlated with an individual characteristic.

In the second stage, I use the following specification:

[8]Top 10 national universities according to U.S. News and World Report as of September 2012 are: Harvard, Princeton, Yale, Columbia, Chicago, MIT, Stanford, Duke, Pennsylvania and Caltech.

[9]As a robustness check to rule out possible sub-section treatment effect (top school attendees form exclusive investment clubs within section), I tested a specification that includes covariates for whether or not any of the HBS executives in the firm attended a top institution or is male. Estimates remained unchanged.

$$Y_{it} = \alpha_2 X_{it} + \beta_2 V_{it} + \gamma_2 \hat{s}_{it} + u_i + k_t + \epsilon_{2it} \tag{1.4}$$

where $\hat{s}_{it}$ denotes the fitted values from the first-stage equation (1.3). If all required assumptions of the estimation steps are fully met, a statistically significant non-zero coefficient estimate of γ_2 will allow one to reject the null hypothesis that section-peers do not influence outcome measures of firms.

1.3.1 Firm-level Aggregation

When aggregating individuals over the firm, there are potentially two issues of concern. First, firm-level measures allow the possibility of falsely viewing a relationship link between Stanford partners as a link between HBS partners in a firm dominated by Stanford alumni.[10] To address this, I show differential marginal effect on deal success when varying social proximity among VC to VC connections. Observing a change in syndication coefficient estimate with respect to change in social proximity controlling for the number of HBS execs in the firm (hence also the firm's number of all partners), I am able to rule out spurious non-HBS connection as a driving effect. Second, if an HBS firm has more than one HBS alumnus from a different class, I need to be careful in assuring that the exclusion restriction is not violated when using instrumental variables for 2SLS estimations.

An HBS venture capital firm (firm with at least one HBS executive) on average has 3.60 alumni over the course of its life. Of the firms with at least two executives, 31% general partnerships include at least one pair from different classes. I examine the consequence of firm-level aggregation on the key operating assumption of balanced sections that ultimately provide the necessary exogeneous variation to correctly identify γ_2.

Suppose the first- and second-stage firm-level outcome equations can be disaggregated to the following pair of individual equations:

$$\begin{aligned} s_{jt} &= \tilde{\alpha}_1 X_{jt} + \tilde{\beta}_1 V_{jt} + \tilde{\delta} z_j + u_j + k_t + \epsilon_{1jt} && (1.5) \\ Y_{jt} &= \tilde{\alpha}_2 X_{jt} + \tilde{\beta}_2 V_{jt} + \tilde{\gamma} s_{jt} + u_j + k_t + \epsilon_{2jt} && (1.6) \end{aligned}$$

where j indexes each individual HBS executive in firm i and z_j denotes a section-level instrumental variable.

[10]However, such bias in estimates is unlikely to exist between entrepreneur and VC connections since the number of CEO/founder in startups is limited to a few.

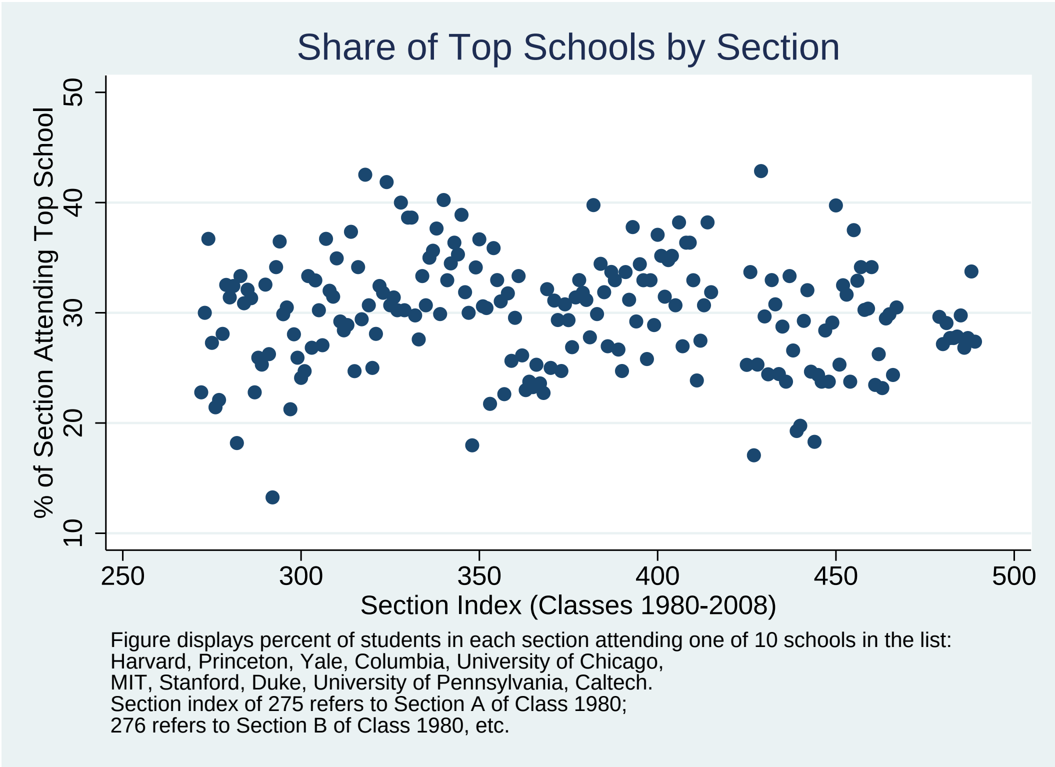
Then by aggregating over each individual j in firm i, I get the following second-stage firm-level outcome equation:

$$Y_{it} = \alpha_2 X_{it} + \beta_2 V_{it} + \tilde{\gamma}\Sigma s_{jt} + u_i + k_t + \Sigma\epsilon_{2jt} \tag{1.7}$$

where by construction, exogenous covariates aggregated over individuals equal the firm-level exogenous covariates.

We know that $E[z_j \epsilon_{2jt}] = 0$ by balanced section, however it remains to be shown that $E[z_j \epsilon_{2j't}] = 0$ for $j \neq j'$ in order to apply the firm-level w_i as proper instrumental variable in the aggregated equation. Intuitively, the correlation requirement stipulates that percentage of top undergraduate institutions attended by students from different sections (both within and across classes) is not correlated with any unobservable characteristic of an HBS executive. I need to rule out that some high (or low) percentage of top undergraduate institutions represented in an older class, 1: did not affect some unobservable section/class-specific shock included in the error $\epsilon_{2j't}$ and 2: did not cause a particular type of future students to sort themselves and apply to HBS.

Figure 1.1: *Percent of section attending top VC/EP-producing schools for classes 1970 - 2008*



I show in figure (1.1) a scatter plot of z_j for every section and class. Simple examination reveals that percentage of top schools represented in earlier classes (1950-1960) are slightly more disperse across section but otherwise there is no distinguishable time trend. An OLS regression of percentage of top schools on section-class index variable (low index for older classes) confirms this description with a coefficient effectively equal to 0. Additionally, if one superimposes IPO rates from 1975 to 2011 on the chart with percentage of top schools represented in each section-class, no noticeable correlation (0.0563) can be detected. This suggests that z_j of past section-class is unlikely to be meaningfully correlated with a future section experience or any sorting by future students. Moreover, HBS administrators actively intervene through the balancing process so that no section carries any college-specific reputation, i.e., 'Section I for the Ivy-leaguers', etc. Finally, I include graduating class-year dummies as controls in X_{it} as a measure for the types of HBS executives in a VC firm to partial-out any unobservable class-trends.

1.4 Data

1.4.1 HBS Section Network

In elaborating the unique institution of section assignment at HBS, [15] provides the following description and anecdotal evidence:

"The B-school [HBS] is the only school in the world I know of that practices the section system to such an extreme. A few other schools use it but for more limited periods. Its [section's] impact is greater than any person's or group's... it drives B-school students to learn, influencing them in countless ways... Yet here is a uniquely American invention of great potential that is affecting our fortunes in economic competition... He [a first-year student] was saying that much of what he learned at the B-school depended on the section, and therefore it behooved him to pay close attention to his relationships with the others." (*Inside Harvard Business School: Strategies and Lessons of America's Leading School of Business*, pages 15 - 17)

Philip Broughton, who was a member of section A of class of 2006 reveals his experience with learning about a section-mate:

"During my time at HBS, I often ran into Bob, my first-semester section neighbor. All my early fears about him had dissolved. Behind the glacial Air Force exterior was a warm, funny man with an elegant mind and a keen ambition to do the best for his family." (*Ahead of the Curve: Two Years at Harvard Business School*, page 209)

As a simple demonstration in showing the effect of section network visibly in the social lives of graduates, I consider a sample of Class of 1993 VCs that married another HBS alum. In the entire sample

Table 1.1: *HBS class of 1993 VC couples*

Firm	Name	Section	Position	College	Spouse Information		
					Class	Section	College
F1	ANON1	B	Senior Vice President	Wellesley	1990	I	Brown
F2	ANON2	C	Managing Director	Middlebury	1993	D	Brown
F3	ANON3	C	President	Oklahoma State	1993	D	Georgetown
F4	ANON4	D	Managing Partner	McGill	1993	B	Harvard
F5	ANON5	D	Managing Director	Ohio State	1993	D	UNC-CH
F6	ANON6	E	Managing Director	UPenn	1995	F	UPenn
F7	ANON7	H	Venture Partner	Dartmouth	1997	A	Dartmouth
F8	ANON8	H	General Partner	UPenn	1993	F	Stanford
F9	ANON9	I	Managing Director	Kansas	1993	I	Harvard

Note: The broader within-HBS coupling ratio is 6%. Among HBS VCs, the ratio is 10%; among class of 1993 VCs, 22.5%. Of the 9 pairs in class of 1993 venture capitalists, 6 are within the same class and 2 are within the same section.

of all graduates, approximately 6% are married to another HBS alum. Among the sample of 40 Class of 1993 VCs (Table 1.1), 9 (22.5%) executives are married to another HBS alumnus. Of the 9, 6 are married to a spouse who graduated from the same class. Of the 6, 2 are married to a spouse who graduated from the same section in the same class. In the absence of any section effect on the social interaction of graduates, one would expect a proportional figure between 0 and 1 executive married to a section-peer spouse, but the anecdotal evidence provides an observation that easily suggests section-class proximity as a meaningful measure of the underlying social interactions among HBS VCs.[11] In the marriage market, effect of face to face meetings and cohort interactions are compelling determinants on outcome, but it remains to be shown that similar effects can be both identified and measured with meaningful economic magnitudes when extended to venture capital fund performance.

1.4.2 HBS Alumni Database

Harvard Business School boasts a growing network of 44,906 MBA graduates. An alumni database maintained by HBS administrators includes names of students from the very early Class of 1915 to the most recent Class of 2012. In the Class of 2011, there were 941 students in total: 62 or 6.6% claimed to start a business, 25% pursued consulting, 39% pursued financial services (1% corporate finance, 7% investment banking, 11% investment management/hedge fund, 14% VC/private equity/LBO, and 6% other financial services).[12]

[11]Similar statistics can be shown for different classes, and also using more inclusive samples with non-VCs.

[12]http://www.hbs.edu/recruiting/mba/data-and-statistics/employment-statistics.html

Every HBS graduate was merged with executive names available in Thomson-Reuter's VentureXpert to be identified as either a venture capitalist or an entrepreneur. Each individual's last, first and middle name was used in that order of significance in the first stage of the matching algorithm to populate candidate executive name matches. Titles such as 'Lt. Col.', 'Dr.', 'Prof.' or 'Rear Admiral' and any suffixes were dropped in both sides, Thomson and HBS database, for uniformity. For names with common variations such as 'Dave' for 'David', a Levenshtein distance tolerance of 2 characters conditional on last name and first two letters of first name match were used. Names with more than three words ('Gonzalo Jose Temes Castrillon' or 'Jose Maria Arturo Tan Castillo') only the first and last words were used for matching. Relatively relaxed matching criteria for long names unlikely yielded much error in the data as two- and three-word names accounted for approximately 85% of all names.

In the second stage, a set of candidate firm matches were generated using firm names. Again, firm names were simplified using a common filtering process of excluding prefixes and suffixes appearing in the following list: pty, pte, plc, llp, llc, lc, lp, inc, l.p., ltd, the. A secondary list was generated matching on the first two words of a firm name, then individually examined to append verified pairs to the primary list of firm name matches (e.g. Bain Capital Ventures and Bain Capital are often used interchangeably).

In the third stage, by combining the lists generated in both first and second stages, pairs of firm-and-executive linkages were formed for each HBS alumnus. A row with only a firm name meant that some HBS executive was found to be associated with an identified venture capital firm or a start-up listed in Thomson, though his name is missing under Thomson; a row with only an executive name meant that an HBS executive was found to be associated with an identified venture capital firm or a start-up listed in Thomson, though his firm association is missing under HBS database; a complete pair meant that the firm-and-executive linkage was verified in both databases. The incomplete pairs can result if an executive discontinued his/her tenure and moved onto a different firm while HBS alumni database failed to keep such records; or if an executive left venture capital/entrepreneurship altogether and Thomson no longer retains historical executive affiliations; or simply the algorithm failed to properly match an individual.

Both partially and fully paired candidates from the third stage were then verified using major web sources (executive biography on the company homepage, CapitalIQ, Bloomberg and LinkedIn) and the Pratt's Guide to Venture Capital for complete identification (e.g. James D. Robinson III, class of 1961C and James D. Robinson IV, class of 1992E are father-and-son founder and general partner, respectively, of RRE Ventures LLC) and to populate employment tenure, paying close attention to employment exits and entries. The median tenure of venture capitalists and entrepreneurs are 11 and 13 years, respectively.

Table 1.2: *Distribution of peer venture capitalists and entrepreneurs by section and class*

	Venture Capitalists				Entrepreneurs			
	N	Mean	Median	%	*N*	Mean	Median	%
Observations (Executive x Time)	26,356				17,766			
Executives	1,742				1,033			
Executive Tenure		15.13	11			17.20	13	
		(12.85)				(14.09)		
Section Peers		4.16	3			2.33	1	
		(2.57)				(1.79)		
Class Peers		40.02	38			20.81	22	
		(16.27)				(8.36)		
# of Firms per Section		5.10	4	3.64%		3.33	2	2.30%
		(2.49)				(1.79)		
# of Firms per Class		38.78	37	3.12%		21.81	23	1.99%
		(16.70)				(8.36)		
# of Firms per Section (Same Focus)		1.80	0	1.39%				
		(2.50)						
# of Firms per Class (Same Focus)		13.39	8	1.12%				
		(14.38)						

Note: Peer data excludes self. Standard deviations are in parentheses.

Of the 44,906 alumni, 2,542 were identified to be venture capitalists (VC) and 1,033 to be entrepreneurs (EP) that started a VC/PE-backed company (Table 1.2).[13] When restricting the VC data sample to partners, managing directors, C-level executives and co-founders of funds, 1,742 remain. Titles held by entrepreneurs included CEO, founder/co-founder, CFO, COO and member of the board. As expected, a large portion of entrepreneurs were also venture capitalists, either through financing rounds and being a board member or through role transitions (successful entrepreneurs who became venture capitalists) - 733 of the entrepreneurs either simultaneously or in sequence were also venture capitalists. Although such numbers appear relatively large (5% VCs including non-partners) or small (3% entrepreneurs) given the initial career destinations reported by the Class of 2011, it is worth noting that many who initially enter private equity/LBO, investment banking/management or consulting often switch into venture capital; and not everyone who starts a company successfully receive VC/PE sponsorship.

1.4.3 Balanced Sections

Since 1949, HBS began dividing all entering MBA students into sections of approximately 90 students each: A, B, C, D, E, F, G. Sections H, I and J were added in the later years[14]. According to HBS sources,

"Students quickly discover that the section experience gets them fully engaged during their first year at HBS and beyond, helping to forge lasting friendships and invaluable contacts for life. In effect, the section becomes a safe and intimate haven where, under the encouragement of mutual support, students can apply newly acquired skills and leadership abilities. It's one of the most formative and defining experiences at HBS."[15]

Prior to section assignments, the administrators survey each student on his/her background information (pre-MBA career, undergraduate institution attended, hometown, etc) to uniformly distribute the class across sections conditional on all observable characteristics. The redistribution effort effectively reduces the variance of observable characteristics by ensuring that students with certain industry experience or undergraduate degree are not ex-post clustered in one particular section. As a result, the estimated errors of regressor coefficients will be larger than the ones derived from section assignments in the absence of administrative interventions.

[13] A few inactive/retired, possibly deceased alumni are included in the summary table. However the subsample of identified entrepreneurs and venture capitalists for whom the employment data was populated were verified to be active during their tenure.

[14] Section H, I and J were each added in 1970, 1971 and 1971, respectively.

[15] http://www.hbs.edu/mba/academics/sectionexperience.html

Table 1.3: *Distribution of HBS MBA graduates by section and class*

Sections	Classes
A (4,840)	After 2000 (11,692)
B (4,919)	1990 - 2000 (8,215)
C (4,862)	1980 - 1989 (7,642)
D (4,848)	1970 - 1979 (7,252)
E (4,820)	1960 - 1969 (5,391)
F (4,839)	1950 - 1959 (3,245)
G (4,549)	1940 - 1949 (1,184)
H (3,613)	1925 - 1939 (253)
I (3,498)	Before 1925 (32)
J (1,710)	
K (811)	
M (47)	
Missing (1,550)	
Total MBA	44,906

Table (1.3) displays the distribution of classes from pre-1925 to post-2000 revealing a steady growth of MBA students over the past 100 years. Distribution across sections A to G is well-balanced in raw number of students as expected; sections H, I and J have graduated fewer students due to their additions in the later years.[16]

1.4.4 Thomson-Reuter VentureXpert and Investment Rounds

Thomson-Reuter VentureXpert lists 5,139 US-based venture firms, or early-stage investors of private equity between 1960 and 2011.[17] These firms have collectively raised 10,667 funds and contain 26,029 executives (including non-investment professionals). Venture firms have been exposed to 134,870 investment rounds on 52,233 unique portfolio companies. Distribution of cross-sectional sectoral exposure on the deals is given in Table (1.4). More than 50% of the rounds are syndicated with at least one other co-investor, or an average number of 1.61 co-investors per round. I employ the following definitions in linking a subset of the data sample to HBS alumni network:

Definition 1.4.1 *HBS GP: VC firm with at least one HBS executive*

Definition 1.4.2 *Section-peer VC (EP, GP): Venture capital executive (entrepreneur, firm) from the same section and*

[16]Introduction of sections H, I and J in the later years do not pose a problem in the analysis because the oldest executive in the sample is from class of 1971, year in which all ten sections were present.

[17]Angel investors and small inactive funds (missing all exposed deal data) are excluded from the sample.

Table 1.4: *Distribution of sector exposure by portfolio company and investment round*

	Companies		Investment Rounds	
Sector	Frequency	Percent	Frequency	Percent
Biotechnology	2,987	5.72	10,480	7.77
Communications and Media	5,353	10.25	15,672	11.62
Computer Hardware	5,336	10.22	14,848	11.01
Computer Software	12,099	23.16	37,663	27.93
Consumer Related	4,340	8.31	8,105	6.01
Energy Related	372	0.71	1,152	0.85
Industrial Products	3,581	6.86	7,228	5.36
Internet Specific	6,158	11.79	12,190	9.04
Medical/Health	4,897	9.38	15,570	11.54
Other Products	4,817	9.22	6,450	4.78
Other Services and Manufacturing	2,293	4.39	5,512	4.09
Total	52,233	100.00	134,870	100.00

Note: Deals are from 1960 to 2011 invested by US-based early-stage funds.

graduating class

Definition 1.4.3 *Class-peer VC: Venture capital executive from the same graduating class*

Definition 1.4.4 *Close class-peer VC: Venture capital executive from the graduating class one year above or below*

Table (1.2) shows that each HBS VC executive has an average of 4.16 section-peer executives present at any given year in which he/she is active, 40.02 class-peer executive. The average tenure (total number of years active) as a venture capitalist among HBS executives is 15.13 years. The table also reveals that within a section, executives do not cluster into the same firm but moderately so within a class; on average, there are 5.10 and 38.78 VC firms per section and class, respectively.

Moreover, 12.7% of all VC firms is an HBS firm and these firms have raised proportionately more funds with 2,276. There are 2.78 HBS executives present per HBS firm per year; 9.88 HBS section-peer firms per year and 65.99 class-peer firms per year.

On the deal side, table (1.5) shows an interesting pattern: the share of portfolio companies eventually going public or being acquired is 10.78% when considering the entire sample, 14.49% when the deal is syndicated by at least one pair of HBS firms, 16.38% when syndicated by at least one pair of class-peer firms, and 18.36% when syndicated by at least one pair of section-peer firms.[18] Additionally, the raw number of deals syndicated between section-peer venture capital firms is 452, or approximately a quarter of 1,758

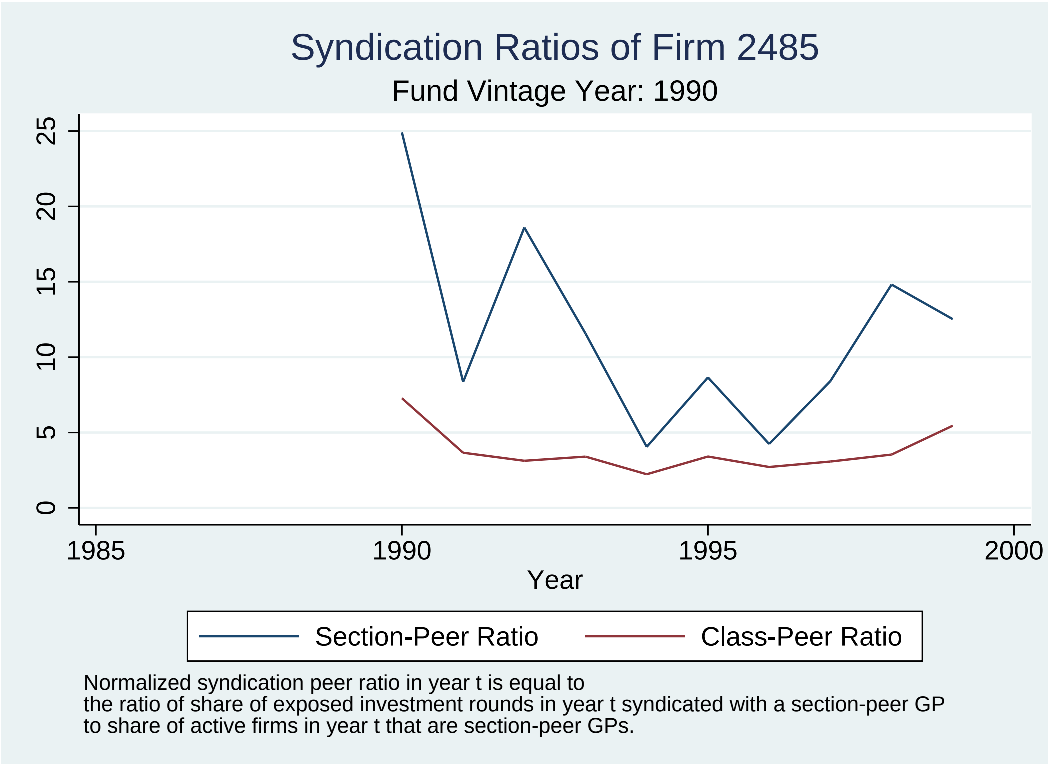
[18]Firm-level aggregation of individual connections can produce upward-biased estimate of a social-tie. However, change in syndication coefficient estimate with respect to change in social proximity controlling for the number of HBS execs in the firm (hence also the firm's number of partners), I am able to rule out non-HBS connection induced bias as a driving effect.

Table 1.5: *Deal success and investment syndicates by varying degrees of social proximity*

Social Proximity	All	VC-VC Syndication								VC-EP Financing		
HBS Alumni		√	√	√	√	√	√	√	√	√	√	√
College-peer			√	√	√							
SP's class-peer						√						
SP's section-peer							√					
Class-peer				√				√			√	
Section-peer					√				√			√
N	52,233	5423	1,483	825	261	354	58	1,758	452	1,475	156	26
Success	0.11	0.15	0.16	0.17	0.16	0.16	0.19	0.16	0.18	0.19	0.24	0.23
	(0.31)	(0.35)	(0.37)	(0.37)	(0.36)	(0.37)	(0.40)	(0.37)	(0.19)	(0.39)	(0.43)	(0.43)
Quality	2.71	4.16	4.31	4.41	4.54	4.50	3.98	4.24	4.34	3.94	3.91	3.91
	(3.23)	(3.77)	(3.74)	(3.89)	(4.24)	(3.75)	(3.71)	(3.83)	(4.11)	(3.64)	(2.75)	(2.75)
Avg. CoInvestors	1.80	2.93	3.10	3.20	3.17	3.06	2.96	3.15	3.18	2.44	2.21	2.21
	(1.00)	(1.13)	(1.12)	(1.17)	(1.12)	(1.01)	(0.97)	(1.20)	(1.24)	(1.01)	(0.92)	(0.92)
Avg. HBS Execs	1.00	3.56	4.87	5.99	6.37	6.32	7.64	4.78	5.37	3.05	5.56	5.56
	(2.34)	(3.52)	(4.08)	(4.56)	(4.54)	(5.09)	(7.13)	(3.92)	(4.11)	(3.16)	(4.39)	(4.39)
Rounds	2.58	5.41	5.63	5.80	5.70	5.66	6.05	5.66	5.83	5.08	4.89	4.89
	(2.51)	(3.24)	(3.17)	(3.17)	(3.19)	(2.95)	(3.45)	(3.25)	(3.47)	(3.02)	(3.43)	(3.43)
Avg. Valuation	55.60	70.39	78.11	85.61	71.85	88.72	77.88	81.29	73.01	77.69	77.24	77.24
	(85.20)	(94.97)	(101.61)	(111.86)	(96.34)	(115.98)	(79.66)	(105.98)	(93.95)	(120.66)	(77.40)	(77.40)

Note: VC-VC syndication among HBS-alumni requires at least one pair of executives from two different venture capital firms on the syndicate of an investment round. VC-EP financing among HBS-alumni requires at least one pair of executives from a venture capital firm and the portfolio company. SP refers to an HBS spouse. Success is defined as the share of portfolio companies exiting through IPO as of 2011. Quality measures firm age as of 2011 (mean of coinvestor fund sequence number). Average valuation is in millions of USD.

Figure 1.2: *Normalized section- and class-peer syndication ratios for firm 2485*



deals syndicated between class-peers. The proportion is even higher in venture capitalist-to-entrepreneur financing matches, in which section-peer matches account for all 100% of class-peer matches. This is clearly an over-representation given ten sections in each class (or seven sections for older classes), suggesting a higher propensity to remain socially exclusive in syndications and financing.

1.4.5 Deal Flow and Syndication Measures

Definition 1.4.5 *Normalized Section/Class-Peer Syndication Ratio (t) =*

$$\frac{\textit{Share of investment rounds in year t syndicated with a peer GP}}{\textit{Share of active firms in year t that are peer GPs}} \tag{1.8}$$

The normalized section-peer syndication ratio measures a firm's propensity to syndicate with a section-peer firm. It should equal 1, conditional on any possible sorting on ability (i.e. private information relevant to performance) in the absence of any preference given toward a socially proximate investing firm. In Figure (1.2), Firm 2485's investment activity through its inaugural fund (size: $77.8M) raised in 1990 focusing on Computer Software across 8 years reveal the significantly higher propensity to syndicate with a section-peer

Table 1.6: *Normalized syndication peer ratios*

	N	Mean	σ
	2,525	9.83	60.20
	2,525	5.47	14.13
Difference	2,525	4.35	50.44
t-statistic	4.34		
Degrees of Freedom	2,524		
p-value	0.0000		

Note: Ratio of share of exposed investment rounds syndicated with a section-peer GP to share of active firms that are section-peer GPs over 2,525 (year x firm) observations of HBS venture capital firms from 1980 to 2011.

firm than the propensity to syndicate with a class-peer firm. Since sections are randomly assigned to balance any private information relevant to ability, it is unlikely that the normalized section-peer syndication ratio is biased upward by ability measures in comparison to the normalized class-peer syndication ratio. In fact, this dominance in propensity is generalizable to all HBS firms as shown in Table (1.6).

Definition 1.4.6 *Leading firm j: Firm j is the leading firm on an investment round if it is either an identified leading investor in the round, or the highest cumulative equity investor in the portfolio company as of the round*[19]

Definition 1.4.7 *Degrees of VC j in year t: # of unique syndication partners of firm j in all investment rounds over a 5-year trailing window ending in year t*[20]

Definition 1.4.8 *Outdegrees of VC j in year t: # of unique syndication partners invited by firm j over a 5-year trailing window ending in year t (firm j is the leading firm)*

Definition 1.4.9 *Indegrees of VC j in year t: # of unique syndication leaders that invited firm j over a 5-year trailing window ending in year t*

Definition 1.4.10 *Active firm in year t: VC firm involved in at least one equity investment round of a portfolio company in year t*

1.4.6 Measures of Performance

I define performance of funds as the share of invested companies that exit via initial public offerings as of December 2011. This is a popular measure employed in previous works ([22, 29]). A more inclusive measure

[19] Thomson-reuter follows similar convention (highest cumulative equity investor in the portfolio company) to identify the lead investor in the round; its identification is usually helpful in deciding between two equally high cumulative equity investors.

[20] Results are robust to 1-, 3-, 7- year trailing windows, generally yielding stronger relationships with shorter trails

that also allows exits via mergers and acquisitions yield similar results in most specifications. Under ideal circumstances, cash flow data for each fund would allow a computation of capital-weighted measures such as IRR or TVPI, which can also be used to estimate economic implication of the effects. However, the usual trade-off faced by empirical researchers in venture capital has been between firm characteristic depths and cross-sectional breadths of information ([6, 18, 36]). While Thomson-Reuter VentureXpert is rich cross-sectionally with reasonably detailed fund-level information, firm valuation data in both pre- and post- rounds are missing for over 70% of investment rounds. These limitations yield noisy estimates of capital-weighted measures due to hidden values of exposed shares and are made worse by share dilutions from issuance of new shares, through a series of follow-on rounds. However, fund sizes are used to yield meaningful measures for economic implications.

Another potential point of objection is the selectivity of the deals shown in the database. It is certainly possible that many failed ventures are excluded and hence performance as measured by share of IPO exits may be biased upward. But in the context of comparing performance effects between section and class-peers, it is unlikely that equity values or likelihood to IPO would be systematically different given the balanced section assumptions. Also, if exit multiples are directly correlated with access (potentially due to improved production function or access to better deals) in addition to its correlation to IPO exit rates, then the performance measures limited to IPO exits would be underestimating the effects of social proximity in both syndication and financing networks to VC success.

1.5 Results

1.5.1 Inaugural Fund

Initial Deal Flow

The data sample consists of a panel of first-time funds raised by 3,538 firms between 1980 and 2011 that I follow for 10 years or up to December 2011, whichever is earlier. I estimate fixed-effects panel regression models under the assumption that disturbances are first-order autoregressive, to allow for persistence over time in a VC firm's network position. I use the Baltagi-Wu ([4]) algorithm to allow for unbalanced panels. Estimates in Table (1.7) reveal that section-peer VC firms contribute to firm deal flow in all three measures.

The average percentage of section-peers belonging to a firm with an active fund targetting the same

Table 1.7: *Initial deal flow from peer venture capital firms*

	(1) Degree	(2) Degree	(3) Indegree	(4) Indegree	(5) Outdegree	(6) Outdegree	(7) HBS Deals	(8) HBS Deals
% of section-peer VCs with same focus fund	185.0**		50.27**		97.45*		4.240*	
	(60.18)		(17.14)		(23.38)		(0.979)	
% of class-peer VCs with same focus fund		432.6*		100.6**		202.3*		6.636*
		(129.8)		(37.04)		(50.35)		(1.560)
# of HBS execs	2.607	2.728	1.158**	1.177**	0.315	0.353	0.000731	0.0220
	(1.739)	(1.740)	(0.497)	(0.498)	(0.674)	(0.674)	(0.0451)	(0.0470)
Log of total equity invested	0.169**	0.168**	0.00814	0.00794	0.0289	0.0284	-0.00545***	-0.00568***
	(0.0677)	(0.0678)	(0.0193)	(0.0193)	(0.0263)	(0.0264)	(0.00302)	(0.00301)
Log of active firms	7.539*	7.527*	1.903*	1.899*	1.725*	1.722*	0.0108	0.0114
	(0.261)	(0.261)	(0.0744)	(0.0744)	(0.101)	(0.101)	(0.00897)	(0.00899)
Log of new firms	0.0527	0.0531	0.0700	0.0702	0.0530	0.0537		
	(0.190)	(0.190)	(0.0542)	(0.0542)	(0.0737)	(0.0737)		
IPO count	0.623*	0.623*	0.180*	0.180*	0.174*	0.173*	-0.00167	-0.00133
	(0.0510)	(0.0510)	(0.0145)	(0.0145)	(0.0198)	(0.0198)	(0.00134)	(0.00134)
MA count	0.514*	0.514*	0.159*	0.159*	0.190*	0.190*	-0.00358*	-0.00381*
	(0.0276)	(0.0276)	(0.00786)	(0.00787)	(0.0107)	(0.0107)	(0.000938)	(0.000977)
Clustering by Firm	No	No	No	No	No	No	Yes	Yes
Firm FE?	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Sector FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Type FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class-year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE?	No	No	No	No	No	No	Yes	Yes
N	20451	20451	20451	20451	20451	20451	17649	17649
R^2_o	0.226	0.223	0.275	0.280	0.201	0.196		
ρ_{ar}	0.807	0.806	0.820	0.819	0.800	0.798		
Durbin-Watson	0.702	0.702	0.727	0.726	0.674	0.674		
Baltagi-Wu LBI	1.006	1.007	1.038	1.039	0.979	0.979		
Correlation	0.216	0.206	0.308	0.312	0.207	0.186		

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Note: The sample consists of a panel of first-time funds by 3,538 VC firms raised betwen 1980 and 2011 that are followed for 10 years or up to December 2011, whichever is earlier. I estimate fixed-effects panel regression models under the assumption that the disturbances are first-order autoregressive, to allow for persistence over time in a VC firm's network position. Percent of peer VCs with same fund focus measures the share of HBS executives' sections actively investing in the same sector. HBS deals refer to share of invested companies run by an HBS graduate. Degrees of VC j in year t measure the number of unique syndication partners of firm j in all investment rounds over a 5-year trailing window ending in year t. Indegrees of VC j in year t measure the number of unique syndication leaders that invited firm j over a 5-year trailing window ending in year t. Outdegrees of VC j in year t measure the number of unique syndication partners invited by firm j over a 5-year trailing window ending in year t. Active firm in year t is a VC firm involved in at least one equity investment round of a portfolio company in year t. All controls are lagged by one year. Baltagi-Wu (1999) algorithm is used to allow for unbalanced panels. Intercepts are not shown.

sector is 1.39% with a standard deviation of 1.61%.[21] The results show that an HBS-firm with one standard deviation higher percentage of section-peer firms with an active fund targetting the same sector syndicates with approximately 3 more firms; invited to syndicate by 1 more firm; and lead 2 more firms in an investment round. The difference in the number of unique firms associated with the three different measures of syndication network reflects the fact that degree measures any connection established between two firms, indegree measures exactly one leading firm per round, and outdegree measures all co-investors in a syndicate excluding the leading firm.[22]

Figure (1.2) provides an example of an inaugural fund raised in 1990 with a focus on Computer Software and a normalized section peer syndication ratio of at least 2 and mean of 11.61 across its 10-year life. A ratio greater than 1 indicates an excess propensity to syndicate with a peer-firm, beyond its porportional representation in the set of all available firms. Moreover, the section-peer syndication ratio remains greater than its class-peer syndication ratio consistently, each year through the entire life-cycle of the fund. This rules out the possibility that syndication patterns are driven by ability-based sorting. Table (1.6) generalizes the result across all HBS-firms with mean normalized section- and class-peer syndication ratios of 9.83 and 5.47, respectively. Using a two-sided unpaired t-test, I can reject the null hypothesis that two ratios are equal with 99.99% certainty.

Section-peer firms, ceteris paribus, appear to syndicate more often among themselves. There are many reasons why this may be true. Peer preference can be derived from an existence of utility gained from friendship that allows for a greater penalty imposition in sustaining the dynamics of an enduring professional relationship. In an environment of emphatic adverse selection and uncertainty of investment quality, cost of information acquisition may be lower given an underlying sustainability of collaboration. Socially-bound syndications may also be used to collude to attain better terms with the entrepreneur.

But what is the mechanism driving the results of outdegree? How is it that a firm is able to lead on more deals if it knows section-peer venture capital firms? One possibility is that the social network gives a firm the ability to create a rich syndicate for entrepreneurs, which is attractive for both broader and less costly access to resource. Indeed, Table (1.7) shows that the share of deals financing HBS- executive run companies are positively associated with the percent of section-peer venture capital firms with the same industrial focus.

When considering the deal side, section-peer entrepreneurs successful in attaining venture financing

[21] A fund is defined to be focusing on a particular sector if the majority of its capital is exposed to the sector.

[22] There are approximately 3 co-investors per round in HBS syndicated deals.

Table 1.8: *Initial deal flow from peer entrepreneurs using 2SLS*

	(1) Degree	(2) Degree	(3) Indegree	(4) Indegree	(5) Outdegree	(6) Outdegree
% of section-peer EPs	67.54**		8.255		24.34**	
	(31.79)		(5.972)		(8.297)	
% of class-peer EPs		59.77**		8.955**		24.44*
		(24.61)		(4.480)		(6.362)
Degree_1	0.623*	0.627*				
	(0.0276)	(0.0278)				
Indegree_1			0.871*	0.873*		
			(0.0107)	(0.0111)		
Outdegree_1					0.886*	0.888*
					(0.00833)	(0.00837)
# of HBS execs	-0.635	-0.403***	-0.148	-0.0214	-0.180	-0.129***
	(0.818)	(0.236)	(0.188)	(0.0457)	(0.270)	(0.0693)
Log of total equity invested	1.844*	1.866*	0.249*	0.256*	0.422*	0.433*
	(0.195)	(0.208)	(0.0613)	(0.0667)	(0.0354)	(0.0366)
Log of active firms	6.113*	6.028*	0.518*	0.495*	0.275*	0.263*
	(0.321)	(0.329)	(0.0660)	(0.0722)	(0.0578)	(0.0574)
Log of new firms	-18.76*	-19.39*	-4.874*	-5.218*	-4.740**	-4.919**
	(4.793)	(4.842)	(1.261)	(1.283)	(1.670)	(1.684)
IPO count	2.426*	2.393*	0.373*	0.367*	0.375*	0.371*
	(0.264)	(0.272)	(0.0451)	(0.0460)	(0.0572)	(0.0571)
MA count	1.891*	1.883*	0.332*	0.334*	0.300*	0.299*
	(0.135)	(0.142)	(0.0320)	(0.0358)	(0.0216)	(0.0220)
Clustering by Firm	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes	Yes	Yes
Type FE?	Yes	Yes	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes	Yes	Yes
Class-year FE?	Yes	Yes	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic	268.77	1271.37	261.61	1207.31	261.28	1207.24
N	24,249	24,249	21,082	21,082	21,082	21,082
R^2	0.845	0.843	0.944	0.943	0.919	0.918

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Table 1.9: *Funds and partnerships, 1960 - 2011*

	All Funds	HBS Funds
General Partnerships	5,139	654
Funds	10,667	2,276
Average Fund size	\$67.65M	\$99.75M
	(220.82M)	(240.99M)
Executives	26,029	1,742
Vintage Years	1960 - 2011	1960 - 2011
Avg. HBS-GP Execs/Year		2.78
		(4.31)
Avg. HBS-Section-Peer-GP/Year		9.88
		(14.13)
Avg. HBS-Class-Peer-GP/Year		65.99
		(66.43)

Note: Firm-level peer statistics are also summarized. Standard deviations are in parentheses. Source: Thomson-Reuter VentureXpert.

(not necessarily with successful exits) yield significant and positive effect in endowing a VC firm with an initial deal flow. I am able to rule out potential section-level shocks to identify and measure a causal effect. Distribution of top undergraduate institutions is used to instrument the share of section-peers that eventually become portfolio company executives. Table (1.8) shows the estimated effects are significant in both degree and out-degree. One standard deviation higher share of section-peers actively seeking financing appears to endow an HBS venture capital firm to lead 1 additional company and syndicate with 1 additional firm.

Fund Size

The conditional average of percentage of section-peers who are actively seeking venture-financing in the vintage year of an HBS-firm's inaugural fund, is 2.02% with a standard deviation of 1.73% (Table 1.9). Then, given the regression estimates in Table (1.10), an HBS-firm with one standard deviation higher percentage of active section-peer entrepreneurs raises 6.59%-19.81% (13.20%) larger inaugural fund. Similarly, an HBS-firm with one standard deviation higher percentage of section-peer VC firms raises 9.45%-26.63% (18.04%) larger inaugural fund.[23] When considering the corresponding two-stage least squares estimation in Table (1.11), I observe that estimations increase to 31.81%-61.43% (46.62%) and 19.71%-36.22% (27.97%), respectively.

In economic terms, the estimates translate to \$56.32M and \$33.79M (Table 1.9) or \$21.09M per peer

[23]Standard deviations are in parentheses.

Table 1.10: *Size of first fund using OLS*

	(1) Log of fund size	(2) Log of fund size	(3) Log of fund size	(4) Log of fund size
% of section-peer VCs	7.577**			
	(3.610)			
% of class-peer VCs		8.581**		
		(3.672)		
% of section-peer EPs			7.629***	
			(3.815)	
% of class-peer EPs				16.29*
				(4.086)
# of HBS execs	0.193	0.219	0.0980	0.234
	(0.292)	(0.279)	(0.220)	(0.239)
Log of total funds	0.622*	0.620*	0.623*	0.625*
	(0.0966)	(0.0972)	(0.0967)	(0.0969)
Log of new firms	2.288**	2.516**	2.075***	2.418**
	(0.968)	(0.983)	(1.042)	(0.993)
Clustering by Year	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes
Type FE?	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes
Class-year FE?	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes
N	2460	2460	2460	2460
R^2	0.453	0.452	0.452	0.452

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Table 1.11: *Size of first fund using 2SLS*

	(1) Log of fund size	(2) Log of fund size	(3) Log of fund size	(4) Log of fund size
% of section-peer VCs	11.75*			
	(3.474)			
% of class-peer VCs		11.62**		
		(3.894)		
% of section-peer EPs			26.95**	
			(8.557)	
% of class-peer EPs				22.21**
				(7.232)
# of HBS execs	0.314	0.306	0.416	0.329
	(0.266)	(0.275)	(0.337)	(0.281)
Log of total funds	0.620*	0.619*	0.619*	0.625*
	(0.0932)	(0.0939)	(0.0943)	(0.0937)
Log of new firms	0.634*	0.635*	0.648*	0.636*
	(0.0392)	(0.0413)	(0.0348)	(0.0319)
Clustering by Year	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes
Type FE?	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes
Class-year FE?	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes
F-statistic	223.05	1680.72	67.52	194.85
N	2460	2460	2460	2460
R^2	0.452	0.452	0.448	0.452

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Table 1.12: *Size of first fund with interactions for managerial experience using OLS*

	(1) Log of fund size	(2) Log of fund size	(3) Log of fund size	(4) Log of fund size
A * % of section-peer VCs	1.258**			
	(0.502)			
A * % of class-peer VCs		2.648**		
		(1.005)		
A * % of section-peer EPs			1.837**	
			(0.710)	
A * % of class-peer EPs				5.081**
				(1.531)
# of HBS execs	0.110**	0.128**	0.111**	0.119**
	(0.0435)	(0.0478)	(0.0424)	(0.0442)
Log of total funds	0.0741	0.0660	0.0891	0.176
	(0.104)	(0.107)	(0.110)	(0.120)
Log of new firms	0.334**	0.317**	0.331**	0.267**
	(0.114)	(0.117)	(0.112)	(0.103)
Clustering by Year	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes
Type FE?	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes
Class-year FE?	No	No	No	No
Year FE?	No	No	No	No
N	320	320	320	320
R^2	0.138	0.136	0.140	0.156

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Note: A is equal to the difference between the vintageyear of the fund and the average number of class-years of the HBS partners.

company and \$7.43M per peer VC-firm, given the average fund size of HBS-firms. In comparison, per peer company premium is close in magnitude to the effect of raising a second fund as estimated by Kaplan and Schoar (2005).

In order to test the idea that fund partners are able to raise larger funds by advertising their social-ties to limited partners, an OLS specification with a proxy for maturity of social contacts is used. If a class of 1995 graduate is raising his inaugural fund in 2000, his contacts (section and class mates) have unlikely had a sufficiently long runway to establish influential reputation. Thus I interact percent of section/class-peer VCs and section/class-peer EPs on the difference between the average graduating class years of all partners in an HBS firm and the vintageyear of its inaugural fund.

Table (1.12) shows indeed that all four coefficient estimates (section/class-peer VCs and section/class-peer EPs) are positive and significant, supporting the idea of raising larger funds using more established contacts.

Selection Bias

Of the 3,799 inaugural funds raised between 1980 and 2007 approximately third of the funds do not report fund size. Moreover, funds without size are also missing data on sector, type and stage. To ensure that estimates are not biased upward on a selective sub-sample of funds with disclosed data, I first consider an incomplete OLS specification without sector, type and stage fixed effects. The resulting coefficient estimates of percent of section-peer EPs and class-peer EPs increase slightly to 8.80 and 17.68 with greater than 95% and 99% probabilistic significance, respectively. The minor change in estimates suggest that the three missing fixed effects do not yield a substantially different results between the two sub-samples.

Second, I consider a corresponding tobit regression over all 3,799 inaugural funds, censoring the 1,339 funds with missing data. Estimates of percent of section-peer EPs and class-peer EPs then increase significantly to 32.76 and 89.43, retaining the same probabilistic significance, respectively.[24] Beta estimates in tobit models measure both the change in the observed fund size weighed by the probability of observing the value, and the change in the probability of observing the value weighed by the mean observed fund size. Thus the increase in estimates suggest that change in likelihood of observing fund size associated with higher percent of peer EPs dominate the diminished direct effect on size when considering the entire sample. This exercise shows that if selection bias in size disclosure yield significantly different results, the selective estimates are likely to be biased downward.

Performance

Using Baltagi-Wu (1999) algorithm for unbalanced panels, I estimate fixed-effects panel regression models under the assumption that the disturbances are first-order autoregressive, to allow for persistence over time in a VC firm's performance. The sample consists of a panel of first-time funds by 2,510 VC firms raised between 1980 and 2007 that are followed for 10 years or up to December 2011, whichever is earlier. Lagged firm performance does not yield any significant coefficient estimate due to its absorption in firm fixed effect. Peer performance refers to share of deals that exited via IPO invested by peer firms before firm inception, conditional on peers having invested in at least one deal.

Table (1.13) shows that if section-peer firms choose to invest in deals in the past and have performed sufficiently well, there is a significantly positive factor correlated with the firm's performance in its inaugural fund. The results cannot be used to ascertain a more rigorously identified causal effect as the preceding

[24]Results on percent peer venture capitalists are similar in both incomplete OLS specifications and tobit models.

Table 1.13: *Fund performance*

	(1) Performance	(2) Performance	(3) Performance	(4) Performance
Section-peer invested?	-0.0438			
	(0.0314)			
Section-peer performance	0.126*			
	(0.0329)			
Class-peer invested?		-0.0597**		
		(0.0237)		
Class-peer performance		0.107*		
		(0.0304)		
Close class-peer invested?			-0.0541**	
			(0.0206)	
Close class-peer performance			0.0925*	
			(0.0266)	
Random-peer invested?				-0.00916
				(0.00745)
Random-peer performance				0.0439*
				(0.0113)
# HBS execs	-0.0114	-0.0107	-0.0115	-0.0105
	(0.0130)	(0.0130)	(0.0130)	(0.0130)
Log of total funds	-0.0695*	-0.0679*	-0.0689*	-0.0697*
	(0.0163)	(0.0163)	(0.0163)	(0.0163)
Log of total equity invested	-0.0587*	-0.0588*	-0.0588*	-0.0594*
	(0.00542)	(0.00542)	(0.00541)	(0.00541)
Log of new firms	0.0392*	0.0397*	0.0395*	0.0390*
	(0.00405)	(0.00405)	(0.00405)	(0.00406)
Firm FE?	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes
Type FE?	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes
Class-year FE?	Yes	Yes	Yes	Yes
N	17677	17677	17677	17677
R^2_o	0.0126	0.0124	0.0119	0.0115
ρ_{ar}	0.404	0.404	0.404	0.405
Durbin-Watson	1.365	1.363	1.364	1.361
Baltagi-Wu LBI	1.872	1.871	1.872	1.869
Correlation	-0.401	-0.398	-0.407	-0.408

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

results on deal flow and fund size. But the results show a systemmatically weaker coefficient estimates from section-peer firms to a randomly chosen firm that was active in the past. Coefficient estimates range from 0.0439 for a randomly chosen firm to 0.126 for section-peer firms. This means that one standard deviation higher section-peer performance in the past leads to 6.3% higher likelihood of IPO to 20.79% from its baseline performance of 14.49%, on deals exposed by HBS-firms through their inaugural funds.

The dummy term for whether or not a peer firm was active in the past is negative, albeit insignificant, suggests a performance hurdle, conditional on being active, at which peer performance has a net-zero effect on the firm. Coefficient estimate on the dummy ranges from -0.009 to -0.044. If peer firms' deal flow improves following strong performance, it seems sensible that the quality of deals also degrades following failures. Table (1.14) shows evidence supporting deal quality changes, where quality of a deal is defined as the average number total funds raised by all co-investors on the deal as of 2011. In the following section for extensions, I also employ the average number of startups founded by c-level executives of portfolio companies as an alternative measure and show similar results.

1.5.2 Peer Syndication Effects

In previous results, I have shown that social connections lead to more deals, more syndication partners and to raising larger funds and improved performance. I now examine the underlying investment production process to distinguish between two competing hypotheses: do social connections reduce cost of search and coordination such that the production function is enhanced among peers, or do peers have access to the best deals?

Table (1.5) shows an interesting pattern: the share of portfolio companies eventually going public or being acquired is 10.78% when considering the entire sample, 14.49% when the deal is syndicated by at least one pair of HBS firms, 16.38% when syndicated by at least one pair of class-peer firms, and 18.36% when syndicated by at least one pair of section-peer firms.[25] Additionally, the raw number of deals syndicated between section-peer venture capital firms is 452, or approximately a quarter of 1,758 deals syndicated between class-peers. The proportion is even higher in venture capitalist-to-entrepreneur financing patterns, in which section-peer matches account for all 100% of class-peer matches. This is an over-representation given ten sections in each class, suggesting a higher propensity to remain socially exclusive in syndications and financing.

[25] Firm-level aggregation of individual connections can produce upward-biased estimate of a social-tie. However, change in syndication coefficient estimate with respect to change in social proximity controlling for the number of HBS execs in the firm (hence

Table 1.14: *Deal level quality OLS regression*

	(1) Quality	(2) Quality	(3) Quality	(4) Quality
Section-peer succeeded	0.417*			
	(0.0589)			
Section-peer failed		-0.190**		
		(0.0836)		
Class-peer succeeded			0.169	
			(0.176)	
Class-peer failed				-0.182
				(0.231)
Average fund size	0.00573*	0.00579*	0.00581*	0.00581*
	(0.000278)	(0.000279)	(0.000279)	(0.000279)
Average # of HBS execs	0.146*	0.151*	0.152*	0.152*
	(0.00836)	(0.00835)	(0.00832)	(0.00832)
Average # of co-investors	-0.00234	0.00142	0.00269	0.00264
	(0.0201)	(0.0201)	(0.0201)	(0.0201)
Corporate PE/Venture Fund	-0.409*	-0.388*	-0.381*	-0.381*
	(0.0562)	(0.0560)	(0.0561)	(0.0561)
Angel/Individuals	-0.263**	-0.249**	-0.250**	-0.252**
	(0.107)	(0.107)	(0.107)	(0.107)
Independent Private Partnership	1.040*	1.062*	1.070*	1.071*
	(0.0680)	(0.0682)	(0.0681)	(0.0681)
Investment Bank	-0.349*	-0.333*	-0.331*	-0.331*
	(0.0503)	(0.0502)	(0.0502)	(0.0502)
SBIC/Government	-0.403*	-0.398*	-0.398*	-0.398*
	(0.0523)	(0.0523)	(0.0522)	(0.0522)
Endowment/Pension Funds	0.0837	0.102	0.0992	0.101
	(0.108)	(0.108)	(0.108)	(0.108)
Others	-0.690*	-0.671*	-0.671*	-0.671*
	(0.0677)	(0.0679)	(0.0680)	(0.0680)
Year FE?	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes
N	19,070	19,070	19,070	19,070
R^2	0.163	0.162	0.161	0.161

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Table 1.15: *Deal level performance OLS regression when syndicated with peers*

	(1) Success	(2) Success	(3) Success	(4) Success
Section-peer syndicated	0.0588**			
	(0.0193)			
Class-peer syndicated		0.0393*		
		(0.00979)		
School-peer syndicated			0.0343*	
			(0.00615)	
College-peer syndicated				0.0122**
				(0.00475)
Average fund size	0.00000767	0.00000663	-0.00000136	0.00000649
	(0.0000147)	(0.0000147)	(0.0000148)	(0.0000147)
Average fund number	-0.0123**	-0.0123**	-0.0112**	-0.0123**
	(0.00461)	(0.00461)	(0.00461)	(0.00461)
Average # of HBS Execs	0.00437*	0.00381*	0.00324*	0.00427*
	(0.000818)	(0.000840)	(0.000858)	(0.000835)
Average # of co-investors	0.0240*	0.0228*	0.0199*	0.0240*
	(0.00236)	(0.00237)	(0.00245)	(0.00236)
Corporate PE/Venture Fund	-0.00438	-0.00482	-0.00344	-0.00454
	(0.00556)	(0.00556)	(0.00556)	(0.00556)
Angel/Individuals	-0.0380*	-0.0375*	-0.0369*	-0.0382*
	(0.00988)	(0.00990)	(0.00988)	(0.00991)
Independent Private Partnership	-0.0436*	-0.0434*	-0.0428*	-0.0438*
	(0.00491)	(0.00491)	(0.00491)	(0.00491)
Investment Bank	0.00274	0.00268	0.00338	0.00256
	(0.00584)	(0.00584)	(0.00584)	(0.00584)
SBIC/Government	-0.0537*	-0.0535*	-0.0527*	-0.0538*
	(0.00559)	(0.00559)	(0.00558)	(0.00559)
Endowment/Pension Funds	-0.0526**	-0.0525**	-0.0520**	-0.0528**
	(0.0190)	(0.0190)	(0.0190)	(0.0190)
Others	-0.0155*	-0.0153*	-0.0126**	-0.0156*
	(0.00411)	(0.00411)	(0.00414)	(0.00411)
Year FE?	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes
N	35,368	35,368	35,368	35,368
R^2	0.0586	0.0588	0.0592	0.0584

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Note: Errors are heteroskedasticity adjusted.

The generalization of social proximity and IPO likelihood pattern is shown using OLS estimation in (1.15) including controls for quality, number of coinvestors, HBS executives and rounds. Although this simple regression analysis offers some interesting observations, it is limited by the absence of affirmative causal statements regarding peer syndication and investment performance. First, it is unclear whether or not VC firms are syndicating on deals because they are friends (a purely social connection) or because social interaction allows ability screening and the syndicate is clustered by ability. Second, the direction of causality cannot be known: is the leading member on the best deal offering access to peer firms or is the investment production function improved because of peer-proximity?

If ability screening is the primary function of peer interaction during the MBA years, one should expect to observe section exclusivity to weaken and eventually disappear in the follow-on funds as market participants will be able to condition on observable past track-records. Regression results in Table (1.15) however suggest that section exclusivity and the associated marginal effect on likelihood to IPO remain robust to inclusion of all follow-on funds as of 2011.

Distinguishing between the two possibilities of improvement in production function through reduction in cost of search/coordination and access to high-quality deals is difficult when using the share of portfolio companies that eventually go public as performance measure. Production function improvement idea would suggest that conditional on accessing a deal, venture capitalists add value beyond providing access to capital through advice and other resources by working more harmoniously with peers. The access idea would suggest that best friends offer access to the best deals, and that investing into high-quality entrepreneurs drives the results of high IPO exit rates.

Using the average number of startups founded (unconditional on exit status) by c-level executives and founders of portfolio companies prior to the first investment round a la Gompers, Kovner, Lerner and Scharfstein (2005), I control for ex-ante deal quality. I regress deal quality on VC-to-EP financing matches and VC-to-VC syndications. Results in Table (1.16) show a systematically increasing positive and significant coefficient estimates with respect to social proximity in connections. This suggests that peers endow venture capitalists with access to better deals.

Then I also examine deal outcomes at the individual round-level, in order to distinguish between peer syndications formed simultaneously in a single round, and syndications formed across a sequence of rounds. Syndications of the latter form render an early investor who is presumably the leader providing access to the later investor. If the production function improvement hypothesis is correct, then either

also the firm's number of partners), I am able to rule out non-HBS connection induced bias as a driving effect.

Table 1.16: *Deal quality OLS regression when financing and syndicating with peers*

	(1) ln(ep quality)	(2) ln(ep quality)	(3) ln(ep quality)	(4) ln(ep quality)	(5) ln(ep quality)
VC-EP Section-peer	0.0390***				
	(0.0202)				
VC-EP School-peer		0.0208*			
		(0.00312)			
Section-peer syndicated			0.0261		
			(0.0172)		
Class-peer syndicated				0.0266**	
				(0.00986)	
School-peer syndicated					0.0470*
					(0.0106)
Average fund size	0.000160*	0.000151*	0.000160*	0.000159*	0.000148*
	(0.0000314)	(0.0000302)	(0.0000315)	(0.0000312)	(0.0000293)
Average fund number	-0.0110**	-0.0101**	-0.0110**	-0.0111**	-0.00964**
	(0.00451)	(0.00445)	(0.00451)	(0.00447)	(0.00434)
Avg. # of HBS execs	0.00289**	0.00215***	0.00269**	0.00218***	0.000751
	(0.00120)	(0.00117)	(0.00118)	(0.00117)	(0.00108)
Avg. # of co-investors	0.0211*	0.0197*	0.0207*	0.0197*	0.0152*
	(0.00341)	(0.00335)	(0.00337)	(0.00329)	(0.00318)
Corporate PE/VC	0.0131	0.0135***	0.0133	0.0132	0.0146***
	(0.00811)	(0.00803)	(0.00809)	(0.00805)	(0.00809)
Angel/Individuals	-0.00445	-0.00356	-0.00351	-0.00275	-0.00158
	(0.0209)	(0.0204)	(0.0206)	(0.0204)	(0.0199)
Ind. Private Partnership	0.0297*	0.0296*	0.0299*	0.0304*	0.0311*
	(0.00545)	(0.00538)	(0.00548)	(0.00551)	(0.00546)
Investment Bank	-0.000421	-0.000151	-0.000206	0.0000351	0.00106
	(0.00414)	(0.00417)	(0.00414)	(0.00412)	(0.00416)
SBIC/Government	-0.00719	-0.00713	-0.00700	-0.00660	-0.00565
	(0.00489)	(0.00489)	(0.00492)	(0.00497)	(0.00493)
Endowment/Pension	-0.0258**	-0.0244***	-0.0257**	-0.0254**	-0.0237***
	(0.0127)	(0.0123)	(0.0127)	(0.0127)	(0.0126)
Others	0.000464	0.00165	0.000539	0.000819	0.00413
	(0.00415)	(0.00412)	(0.00416)	(0.00416)	(0.00418)
Year FE?	Yes	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes	Yes
N	35397	35397	35397	35397	35397
R^2	0.0973	0.100	0.0973	0.0975	0.0990

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Note: Specifications in columns (1)-(2) are for VC-EP financing. Specifications in columns (3)-(5) are for VC-VC syndications. Errors are heteroskedasticity adjusted.

Table 1.17: *Deal level performance OLS regression when financing or syndicating with peers*

	(1) Success	(2) Success	(3) Success	(4) Success	(5) Success	(6) Success	(7) Success	(8) Success	(9) Success
Section-lead	0.0231								
	(0.0183)								
Section-follow		0.0420**							
		(0.0212)							
Section-sim			0.0251						
			(0.0163)						
Class-lead				0.0391*					
				(0.0108)					
Class-follow					0.0199***				
					(0.0102)				
Class-sim						0.00569			
						(0.00744)			
School-lead							0.0173**		
							(0.00552)		
School-follow								-0.000111	
								(0.00539)	
School-sim									0.00542
									(0.00458)
Log of EP Quality	0.0171**	0.0169**	0.0172**	0.0164**	0.0169**	0.0172**	0.0163**	0.0172**	0.0172**
	(0.00566)	(0.00566)	(0.00566)	(0.00566)	(0.00566)	(0.00566)	(0.00569)	(0.00566)	(0.00566)
Average fund size	-0.0000191***	-0.0000189	-0.0000189	-0.0000201***	-0.0000192***	-0.0000190	-0.0000199***	-0.0000190	-0.0000194***
	(0.0000116)	(0.0000116)	(0.0000116)	(0.0000116)	(0.0000116)	(0.0000116)	(0.0000116)	(0.0000116)	(0.0000116)
Average fund number	-0.00663	-0.00663	-0.00670	-0.00664	-0.00661	-0.00670	-0.00652	-0.00668	-0.00659
	(0.00420)	(0.00420)	(0.00420)	(0.00420)	(0.00420)	(0.00420)	(0.00420)	(0.00420)	(0.00420)
Avg. # of HBS execs	0.00326*	0.00307*	0.00320*	0.00322*	0.00300*	0.00323*	0.00332*	0.00329*	0.00319*
	(0.000646)	(0.000641)	(0.000650)	(0.000646)	(0.000646)	(0.000660)	(0.000648)	(0.000645)	(0.000670)
# of Coinvestors	0.00712*	0.00690*	0.00706*	0.00705*	0.00677*	0.00708*	0.00714*	0.00715*	0.00702*
	(0.00141)	(0.00140)	(0.00141)	(0.00141)	(0.00137)	(0.00141)	(0.00141)	(0.00137)	(0.00142)
Investor-type FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	118071	118071	118071	118071	118071	118071	118071	118071	118071
R^2	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143

Standard errors in parentheses
*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Note: Section/class/school-lead is equal to 1 if a subsequent round includes a section/class/school-peer firm. Section/class/school-follow is equal to 1 if a previous round includes a section/class/school-peer firm. Section/class/school-sim is equal to 1 if the current round includes a section/class/school-peer firm. Errors are clustered by company.

form of syndication should be both positively and significantly correlated with deal success. Results in Table (1.17) show that simultaneously formed peer syndications are not significantly correlated with deal success. Rather, estimates indicate that following an early section- or class-peer investor is positively and significantly correlated with success. Using this evidence, I conclude that deal access hypothesis is more likely true. Further details on the analyses are provided in Section VI.

1.5.3 Across Funds

Deal Flow

In the previous section, it was shown that both peer venture capital firms and entrepreneurs have a significant and positive effect on providing the firm with excellent deals in its inaugural fund. Before a firm's performance is revealed to the market, it is plausible that socially proximate peers would endow a firm with deals given lower search cost and higher propensity to syndicate. Indeed, this effect is identified using 2SLS and presumably most amplified at the firm's inception. But as the firm ages, does the underlying social network continue to be operative? In the subsequent funds, market players can select on the firm's past track-record and the previous regression exercise on first funds can confound the measures with ability. I approach this problem by examining the number of peers who are not directly involved in venture capital but potentially relevant to providing deal flow. I subdivide the data into venture firms that target[26] one of six (computer software & internet, computer hardware, energy, telecommunication, biotechnology, retail) major sectors and examine number of peers working in the corresponding sector in non-investment/financial roles. Often times, these individuals are part of the company operations or management or work as consultants.

Columns (1) - (6) of Table (1.18) show estimations with significant coefficients in Computer Software & Internet. Across all three measures of deal flow, the effect of number of section-peers working in the industry dominates corresponding effects from the number of class-peers. These estimations however fail to rule out the possibility of section-level common shocks during the peers' MBA years. For example, a particular professor of information technology could have indoctrinated a section of students a certain type of management philosophy that subsequently lead the section to pursue the industry, albeit in different roles (operations vs. investment). To address this concern, I consider the number of section peers residing in the same state as the executives of the venture capital firm. Then I interact this term with the number of

[26] Fund must deploy more than 50% of its capital to the target sector

Table 1.18: *Deal flow in computer software & internet*

	(1) Degree	(2) Outdegree	(3) Indegree	(4) Degree	(5) Outdegree	(6) Indegree	(7) Degree	(8) Outdegree	(9) Indegree
Outdegree_1			0.422*			0.421*			0.413*
			(0.0469)			(0.0470)			(0.0445)
Section-peers in same sector (A)	0.625**	0.239**	0.124***				0.373***	0.144***	0.0634
	(0.271)	(0.0857)	(0.0673)				(0.221)	(0.0773)	(0.0577)
Class-peers in same sector				0.0828**	0.0329**	0.0155**			
				(0.0354)	(0.0141)	(0.00736)			
Section-peers in same state (B)							0.317**	0.100***	0.0552***
							(0.145)	(0.0591)	(0.0330)
A*B							0.0434***	0.0160	0.0102**
							(0.0236)	(0.0132)	(0.00339)
# of HBS execs	-1.137**	-0.317***	-0.304**	-1.510**	-0.488***	-0.355**	-1.862**	-0.554**	-0.442**
	(0.569)	(0.170)	(0.146)	(0.680)	(0.254)	(0.145)	(0.740)	(0.219)	(0.185)
Log of total equity invested	2.399*	0.933*	0.270***	2.378*	0.925*	0.267***	2.338*	0.914*	0.267***
	(0.339)	(0.0718)	(0.140)	(0.340)	(0.0720)	(0.140)	(0.348)	(0.0711)	(0.141)
IPO count	0.0106	-0.166**	0.0212	0.0234	-0.160**	0.0232	0.0118	-0.165**	0.0196
	(0.214)	(0.0633)	(0.0636)	(0.213)	(0.0627)	(0.0637)	(0.209)	(0.0619)	(0.0623)
MA count	2.126*	0.679*	0.403*	2.126*	0.678*	0.404*	2.073*	0.660*	0.397*
	(0.247)	(0.0447)	(0.0552)	(0.248)	(0.0442)	(0.0553)	(0.233)	(0.0397)	(0.0539)
Clustering by Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Type FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	33684	33684	33683	33684	33684	33683	33684	33684	33683
R^2	0.510	0.456	0.629	0.511	0.456	0.629	0.519	0.464	0.633

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Note: Peers in Computer Software & Internet work in non-investment roles. Section-peers in same state reside in the same state as the HBS executives of venture capital firms. A*B is the interaction term between the number of section-peers working in Computer Software & Internet in non-investment roles and the number of section-peers in same state of residence.

section peers working in the same sector as a way to provide a source of variation that is likely correlated with the level of social interaction (availability to meet at conferences, trade shows, networking events, reunions, etc). Indeed, both degree and indegree regressions in columns (7) and (9) reveal significantly correlated interaction terms. This provides some evidence for the fact that even after the inaugural fund, venture capitalists continue to receive deal contacts from their social network.

Performance on Fund Size

It has been shown that there exists a prediction of performance using lagged peer-performance when considering inaugural funds. Does this pattern continue and remain persistent across second, third and subsequent funds? What are the economic implications of positive peer-performance? I address these questions by considering the wider panel of all funds raised by 2,955 VC firms betwen 1980 and 2007 followed for 10 years or up to December 2011, whichever is earlier. I estimate fixed-effects OLS regression models with specifications that vary in social proximity: section-peer, class-peer, close class-peer (one class above and below the graduating year) and a randomly chosen peer firm with an active contemporaneous fund. Table (1.19) shows the resulting estimate, revealing positive and significant estimates on peer performance on fund size in diminishing strengths as I vary the social proximity of peers, with vanished significance for random peer firm performance. A 95% confidence interval estimation on section-peer firms exclude the coefficient point estimate of a randomly chosen firm, suggesting a statistically different and stronger effect among connected peers. Coefficient estimates range from 0.0450 on performance of randomly chosen firms to 0.774 on performance of section-peer firms. One standard deviation (28.63%) higher section-peer performance in the previous period leads to an estimated 22.16% increase in fund size, or $28.48M.

1.6 Extensions

Is it possible to distinguish between production enhancement and deal access hypotheses? Distinguishing between the two possibilities is difficult when using the share of portfolio companies that eventually go public as performance measure. Unlike the investment process in public markets, venture capital managers not only choose to invest into startups, they are also influential in achieving successful exits. Moreover, a measure of deal quality defined as the aggregate age of co-investors in the syndicate will be endogenous to performance, as managers' ability to raise follow-on funds depends on investment outcome.

Table 1.19: *Fund performance on fund size*

	(1) Log of fund size	(2) Log of fund size	(3) Log of fund size	(4) Log of fund size
Section-peer invested?	-0.202			
	(0.136)			
Section-peer performance	0.774**			
	(0.238)			
Class-peer invested?		-0.202		
		(0.128)		
Class-peer performance		0.636*		
		(0.193)		
Close class-peer invested?			-0.214***	
			(0.111)	
Close class-peer performance			0.613*	
			(0.153)	
Random-peer invested?				0.0434
				(0.0367)
Random-peer performance				0.0450
				(0.0576)
# HBS execs	-0.0881**	-0.0864**	-0.0838**	-0.0837**
	(0.0282)	(0.0276)	(0.0277)	(0.0261)
Log of total funds	-0.258*	-0.254*	-0.254*	-0.253*
	(0.0606)	(0.0608)	(0.0607)	(0.0607)
Log of total equity invested	0.540*	0.540*	0.540*	0.540*
	(0.0234)	(0.0235)	(0.0234)	(0.0234)
Clustering by Firm	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes
Type FE?	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes
Class-year FE?	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes
N	37755	37755	37755	37755
R^2	0.480	0.480	0.480	0.480

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

An alternative measure for deal quality following Gompers, Kovner, Lerner and Scharfstein (2005) is the average number of startups founded (unconditional on exit status) by c-level executives and founders of portfolio companies prior to the first investment round. GKLS find that serial entrepreneurs, regardless of the outcome of their previous ventures, will be more likely to succeed than a first-time entrepreneur. This is because serial entrepreneurs appear to learn also from failures.

Using this measure, the analysis is run at the deal level to examine VC-to-EP financing matches and VC-to-VC syndications on the ex-ante quality of deals. Results are available in Table (1.16). Columns (1) and (2) show that there is nearly 100% increase in the marginal effect of VC-to-EP financing matches among HBS peers on the log of deal quality when the matches are restricted to section-peers.[27] A similar pattern does not emerge among peer syndications (columns 3 to 5), however coefficients for class-peer and school-peer syndications are positive and significant. Together, the results suggest that social-ties endow venture capitalists with access to better deals.

1.6.1 Round leaders and followers

One approach to further understand the investment process between peers is discriminating syndication types. I use individual deal-round data that consists of 118,071 deal-rounds for 35,368 portfolio companies financed by US-based venture funds between 1960 and 2011. Each indicator variable for section/class/school-lead is equal to 1 if a subsequent round for the same company includes a corresponding peer firm in its investment syndicate. Each indicator variable for section/class/school-follow is equal to 1 if a previous round includes a corresponding peer firm. Lastly, each indicator variable for contemporaneous syndications is equal to 1 if the current round includes a corresponding peer firm. The idea is to distinguish between being an early investor versus being a late investor; then examine the effects separately across various social proximities.

Table (1.17) shows the result of regressing investment success (equal to 1 if exit via IPO, 0 otherwise) on each early, contemporaneous and late indicator variables and other controls. In addition to investor-type, year and sector fixed effects, round stage fixed effect is also included. This control allows me to rule out the possible interpretation that a positive and significant effect of being a follower investor on success is driven by gaining access to selectively matured deals. Results reveal that leading a round as a class-peer or school-peer firm is positively and significantly related to success. Following into a round as a section-peer

[27] In this particular dataset, all class-peer matches occurred between section-peers.

or class-peer firm is positively and significantly related to success. In particular, following a section-peer leads to twice the marginal effect on success than following a class-peer. Moreover, investing simultaneously on the rounds for section/class/school-peers is not significantly related to success. Together, the data suggest that peers are not likely improving the investment production process. Rather, most closely-knit peers benefit from gaining access to deals identified by their early-investor contacts.

1.7 Conclusion

The mission of this paper has been identifying and measuring effects of deal access in venture capital. I employed the random section assignment at Harvard Business School to show that being socially connected in venture capital leads to more deal flow, larger asset under management and better performance in the inaugural funds of HBS executive-run venture capital firms.

The first set of results are derived from inaugural funds raised between 1980 and 2011 by US-based venture capital partnerships so that the effects are free from selection on observable past track-record of funds. I examined regression coefficients on the number of peers (normalized by section or class size) working as either venture capitalists or entrepreneurs seeking financing in the vintage years as the key estimate of interest. Since this is an ex-post peer outcome (decision to pursue VC or entrepreneurship is endogenous), I addressed this problem by using ex-ante characteristics of the incoming section-peers as instruments and showed that estimates are robust to such controls. The results showed that one standard deviation higher share of section-peers actively seeking financing appear to endow an HBS venture capital firm to lead 1 additional company and syndicate with 1 additional firm. In economic terms, the estimates translate to \$56.32M and \$33.79M (Table 1.9) uncondtionally per peer company and VC firm, respectively, or \$21.09M and \$7.43M, given the average fund size of HBS-firms. In comparison, per peer company premium is close in magnitude to the effect of raising a second fund as estimated by Kaplan and Schoar (2005).

I also showed that the lagged outperformance of section-peer funds (funds lead by a section-peer general partner) on deals prior to the inaugural fund's vintage year, has a positive effect on fund performance. The results showed that one standard deviation higher section-peer performance in the past leads to 6.3% higher likelihood of IPO to 20.79% from its baseline performance of 14.49% in the absence of peer activity. In economic terms, one standard deviation (28.63%) higher section-peer performance in the previous period leads to an estimated 22.16% increase in inaugural fund size, or \$28.48M.

In order to provide insights into the mechanism through which social-ties operate, I tested three candidate hypotheses. One possible explanation of the effect of social interactions on managerial success is ability screening. Regression results however suggested that section exclusivity and the associated marginal effect on likelihood to IPO remains robust to inclusion of all follow-on funds as of 2011, arguing against ability screening.

Second explanation is that social-ties reduce cost of search and coordination such that the production function is enhanced among peers. Third hypothesis is that most connected VCs receive privileged deal flow. Distinguishing between the two possibilities is difficult when using the share of portfolio companies that eventually go public as performance measure. Production function improvement idea would suggest that conditional on accessing a deal, venture capitalists add value beyond providing access to capital through advice and other resources by working more harmoniously with peers. The access idea would suggest that best friends offer access to the best deals, and that investing into high-quality entrepreneurs drives the results of high IPO exit rates.

Using the average number of startups founded (unconditional on exit status) by c-level executives and founders of portfolio companies prior to the first investment round, I control for ex-ante deal quality. I regress deal quality on VC-to-EP financing matches and VC-to-VC syndications. Results show a systematically increasing positive and significant coefficient estimates with respect to social proximity in connections. This suggests that peers endow venture capitalists with access to better deals. Peer VCs may be able to sustain a stronger dynamic relationship that forms a proprietary deal flow. Indeed, I find evidence that leading investment rounds is associated with being invited to future rounds.

Then I took another step by examining deal outcomes at the individual round-level, in order to distinguish between peer syndications formed simultaneously in a single round, and syndications formed across a sequence of rounds. Syndications of the latter form render an early investor who is presumably the leader providing access to the later investor. If the production function improvement hypothesis is correct, then either form of syndication should be both positively and significantly correlated with deal success. Results however show that simultaneously formed peer syndications are not significantly correlated with deal success. Rather, estimates indicate that following an early section- or class-peer investor is positively and significantly correlated with success. Using this evidence, I conclude that deal access hypothesis is more likely true.

The second set of results are derived from an inclusive panel data set of 3,799 funds raised between 1980 and 2007, and measure deal flow and performance across 10-year lifecycles. I used data on peers working in the same sector as the fund's focus but in non-investment roles varied by residential proximity to tease out effects of social connectedness on deal flow.

A large obstacle in venture capital research is limitation in data due to non-mandatory discloure of information in private markets. This limitation poses interesting challenges to both the researchers and practitioners. In this paper, I focused on the variation in investment opportunity sets particular to venture capital, as a function of social-ties. I showed that being socially connected leads to better deal flow, ability to raise larger funds and investment performance. Results suggest that the well-connected venture capitalist may be successful by attaining access to great deals. The effect is identified in the inaugural fund and appears to be persistent in the follow-on funds.

However, the evidence and analysis presented are suggestive and indirect as rejections of deals or syndication invitations are unobserved. Future research could benefit from more exhaustive data sources or an application of structural model estimations.

Chapter 2

Competing Ad Auctions[1]

2.1 Introduction

Online ad auctions sell advertising placements on search engines and elsewhere—providing key funding for a variety of online resources. Advertisers sign up with one or more ad platforms, specify their advertising preferences (including conditions in which they want their ads to be shown, and how much they are willing to pay), and receive clicks from interested users.

In this paper, we explore competition among ad platforms that offer search engine advertising services. Our motivations are several. For one, we want to understand why some advertisers choose to use only certain ad platforms but not others. After all, if an ad network cannot attract a broad selection of advertisers, it will be unable to present ads related to users' requests, and it will also garner substantially lower revenue. Second, we want to explore competition across auction platforms. Auction competition is a relatively unexplored topic in the literature. The primary difficulty in considering auction competition is that classic competition models (e.g. Bertrand and Cournot) view competition through the setting of market price, either directly or via production quantities. But ad auctions set prices through importantly different mechanisms: strategic interaction among advertisers yields price discovery, while an ad platform can influence outcomes through, e.g., setting of reserve prices.

Competition in online ad auctions also differs from models of auction competition examined in the literature. [31] explores optimal mechanisms for a seller employing a second-price auction with a reserve

[1]Co-authored with Ben Edelman, Harvard Business School and Itai Ashlagi, Massachusetts Institute of Technology, Sloan School of Management

price in a general context. In contrast, we limit our auction competition to the search engine industry. This restriction puts useful structure on competition: in order to attract users, an ad platform seeks to reduce cost of search to users. But reducing the cost of search decreases the number of clicks users perform, thereby reducing payments from advertisers. Separately, [16] consider competition among auctions occuring at the level of strategic entry by auctioneers in two segmented markets with differing buyer-to-seller ratios. There, an auctioneer chooses between entering in a market saturated with buyers by expecting to compete with many sellers (competition effect) and selling to great magnitudes (scale effect). Applying this framework to online ad auctions implies that search engines compete by specializing on exclusive sets of keywords with correspondingly segmented sets of buyers. However, the leading ad platforms do not strategically limit keywords or even focus on distinctive sets of keywords. Rather, leading ad platforms sell clicks from web searchers searching for all manner of subjects, and ad platforms sell to a common pool of advertisers and would-be advertisers.

We model competing ad auctions in two stages. In the first stage, advertisers choose to enter one or more ad auctions after considering each search engine's user base (capacity) and click-through rates (technology) in light of an exogenous cost of joining each ad auction. Using the symmetric Bayes-Nash equilibrium, we derive advertisers' entry strategies that lead to the VCG outcome of the ad auctions. In the second stage, participating advertisers submit bids in the ad auction(s) they chose to enter, and ad platforms strategically compete on increasing capacity while preserving an optimal level of technology to maximize revenue. This follows a model of Cournot-style competition among ad platforms.

Closest to our paper is [24] which studies competition across ad auctions with auction capacities endogenized by consumer choice. However, they abstract away from the mechanisms that allocate and price advertising positions. In contrast, our approach is grounded in the unusual mechanisms used in selling online advertising.

We proceed in five parts. In Section 2.2, we develop model fundamentals and notation. In Section 2.3, we model ad auctions with participation costs—offering an initial explanation of why not all advertisers use all ad platforms, and testing that explanation with available data. In Section 2.4, we analyze competing ad auctions in light of user preferences over search engines. In Section 2.5 we develop a concept of "joining" two ad auctions, and we identify conditions where joins increase or reduce advertiser welfare. In Section 2.6, we conclude.

2.2 Ad Auctions: GSP and VCG

In an *ad auction*, there are advertisers $\mathcal{N} = \{1, \ldots, N\}$ and K slots for sale. Each advertiser can receive at most one slot. The positions are sold for a single period of time. Each slot k has an expected *click-through rate* $\alpha_k > 0$. The auction has a known capacity $C > 0$. Thus, if an advertiser wins slot k, the advertiser will receive $C\alpha_k$ clicks in expectation. We assume that $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_K$. Define $\alpha_k = 0$ for every $k > K$.

The *value* per click for advertiser j is $v_j \in [0, 1]$. Advertisers are risk neutral, and the payoff to advertiser j for winning slot k is $C\alpha_k v_j$ minus its payments to the ad platform. We also assume that for each advertiser j, v_j is drawn from a commonly known distribution F, and the value of each advertiser is private information.

Each advertiser j is required to submit a bid b_j. We denote by $b_{(j)}$ the j^{th} highest bid. Similarly we denote by $g(j)$ the identity of the j^{th} highest advertiser. In case of ties, the order among those advertisers is determined randomly.

We consider payments and outcomes under two distinct auction mechanisms. Modern ad platforms generally use the *Generalized Second Price* (GSP) structure, wherein an advertiser $g(j)$ receiving position k pays a total of $C\alpha_j b_{(j+1)}$. In the *Vickrey-Clarke-Groves* (VCG) ad auction, each advertiser pays its impact on all others' social welfare, assuming bids equal values. Hence, under VCG, advertiser $g(j)$ receiving position k pays a total of $p^{VCG,k} = \alpha_k C \sum_{j=\min(K+1,n)}^{\min(k+1,n)} (\alpha_{j-1} - \alpha_j) b_{(j)}$. Both GSP and VCG ad auctions allocate the first position to the highest-bidding advertiser, the second position to the second-highest bidder, and so forth.

Bidding truthfully is a dominant strategy for every advertiser in the VCG ad auction. (See [38], [9], [19].) [14] and [37] show that there exists an equilibrium in the GSP ad auction under complete information such that the GSP outcome coincides with the outcome of the VCG auction in which each advertiser bids truthfully. Furthermore, this equilibrium yields the lowest revenue for the seller, so it is in some sense best for the advertisers. Finally, [8] shows that a reasonable myopic GSP bidding strategy converges to this equilibrium.

2.3 Ad Auctions with Participation Costs

Consider a GSP ad auction, and suppose there is a cost $Z > 0$ for each advertiser to participate in the auction; Z can be interpreted as an advertiser's transaction cost in submitting its campaign into the ad platform. Elements of Z include creating an account, setting advertising parameters, monitoring effectiveness, adjusting bids, and paying bills. Z is not transferred to the auctioneer. Therefore, a high participation cost, relative to a platform's capacity, will cause an advertiser to forego use of that platform, even if the advertiser would otherwise find it profitable to use that platform. For example, if the following inequality holds

$$\frac{Z}{C} > \alpha_1 v, \tag{2.1}$$

then an advertiser with value v will never enter the auction since the advertiser would realize negative utility even if he managed to receive the first slot with zero payment.

In this section we show that there exists a unique threshold function that determines whether an advertiser will participate in a given ad auction. In particular, consider the following two-stage game: In the first stage, each advertiser decides whether to participate in the auction, as a function of the advertiser's value v. The value of each advertiser is private information (whether or not the advertiser enters the auction). In the second stage, all advertisers that decide to participate in the auction submit a bid. The advertisers who enter face the same conditions as in [14] and [37], and the outcome coincides with the VCG outcome of the ad auction with these advertisers.[2]

A strategy for an advertiser j is a function of $s : [0,1] \to \{0,1\}$ where $s_j(v_j) = 1$ means that advertiser j enters the auction given that its value v_j, and $s_j(v_j) = 0$ means j does not enter. We employ symmetric Bayes-Nash equilibrium as our benchmark equilibrium concept when further analyzing advertisers' strategies.

Denote by $U(v,n)$ the expected utility (before considering entry cost Z) of an advertiser with valuation v, who decides to participate in the auction given that exactly n other advertisers also participate. For simplicity, we assume that if an advertiser's expected utility is 0 (including costs), the advertiser prefers to enter the auction. Thus if all advertisers use the strategy s^*, then an advertiser with valuation v will choose to enter the auction if and only if

$$E_n[U(v,n)|s^*] \geq Z. \tag{2.2}$$

After advertisers decide whether to enter, a set of advertisers B bid in the auction, and their equilibrium bids coincide with the GSP equilibrium in [14] and [37].

Theorem 2.3.1 *Suppose $U(1,N) > Z$. Then there exists a unique strategy s^* which forms a symmetric Bayes-Nash equilibrium in the two-stage game. In particular there exists a threshold $v^* > 0$ such that for each $v \in [0,1]$, advertiser j*

$$s_j^*(v) = 1 \tag{2.3}$$

if $v \geq v^*$ and

$$s_j^*(v) = 0 \tag{2.4}$$

[2]In [1] the authors do not model participating costs, but rather let advertisers choose a single ad auction to participate in.

otherwise.

To prove theorem 2.3.1, the following lemma will be useful. The lemma provides two monotonicity properties: the expected utility for an advertiser who enters the auction is non-increasing in the number of participants, and is increasing in the advertiser's value.

Lemma 2.3.2 *For any $n < N$, and any v:*

1. $U(v, n) > U(v, n+1)$.
2. $\partial U(v, n)/\partial v > 0$.

Proof 2.3.3 *Let $Q(v, n)$ and $P(v, n)$ denote the expected number of clicks and expected payment for an advertiser with value v that decides to enter auction l given that there are $n-1$ other advertisers in the auction. Thus the expected utility of such an advertiser is*

$$U(v, n) = vQ(v, n) - P(v, n). \tag{2.5}$$

Note that

$$Q(v, n) = C \sum_{1 \le k \le n} \binom{n-1}{k} \alpha_k \tilde{F}(v)^{n-k} [1 - \tilde{F}(v)]^{k-1}. \tag{2.6}$$

where $\tilde{F}(v)$ is the probability that a bidder that entered the auction has a valuation lower than v.[3]

Let X_n be a random variable distributed over $\{\alpha_1, ..., \alpha_N\}$ such that $P(X_n = \alpha_k) = \binom{n-1}{k} p^{n-k} (1-p)^{k-1}$ where $p = \tilde{F}(v)$, and let $Y_n \sim Bin(n-1, p)$, i.e. $P(Y_n = k) = \binom{n-1}{k} p^{n-k} (1-p)^{k-1}$. Note that $P(X_n = \alpha_k) = P(Y_n = k)$. For every $m > 1$, define

$$\mu(m, n) = (\alpha_m - \alpha_{m-1}) P(Y_n \ge m-1),$$

and let $\mu(1, n) = \alpha_1 P(Y_n \ge 0)$. Observe that $Q_l(v, n) = CE[X_n] = C \sum_{m \ge 1} \mu(m, n)$ and that $\mu(1, n) = \alpha_1$. By our assumption on click-through rates, $\alpha_{k+1} - \alpha_k \le 0$. Moreover, $P(Y_n \ge k)$ is decreasing in n because more trials with the same probability of success will lead to more successes. Therefore, for all $m > 1$, every $\mu(m, n)$ is non-increasing in n, implying that $Q(v, n) > Q(v, n+1)$. Finally, by [30],

$$U(v, n) = \int_0^v Q(x, n) dx, \tag{2.7}$$

implying that $U(v, n) > U(v, n+1)$.

[3] We make no assumption on the entry strategy profile, except requiring that it induce a measurable distribution $\tilde{F}$.

The second part follows because the derivative of $U(v, n)$ with respect to v is equal to $Q(v, n)$ by (2.7) and because $\partial Q(v, n)/\partial v > 0$.

Proof of Theorem 2.3.1:

Fix the entry strategies of all other advertisers except j. By Lemma 2.3.2, j is better off entering given some fixed value v^*, as long as $v_j \geq v^*$. Thus, every symmetric Bayes-Nash equilibrium is characterized by a threshold value: there exists a $v^* \in [0, 1]$ such that every advertiser will enter the auction if and only if its value is at least v^*.

Fix a symmetric strategy profile that is characterized by a threshold s^*. To show existence of the threshold value, define

$$G(v) = E_n[U(v, n)|s^*] - Z.$$

By Lemma 2.3.2, $G(v)$ is continuous and strictly increasing in v and by assumption. Furthermore, $G_l(0) = -E_n[P(0, n)|s^*] - Z < 0$, and by assumption $G_l(1) = E_n[U(1, n)|s^*] > U(1, N) > Z > 0$. Therefore there exist v^* such that $G(v^*) = Z$. Since $G_l(v)$ is strictly increasing in v, v^* is unique. □

The preceding proof indicates that the cut-off value v^* is a function of the platform's primitives, $Z, C, \alpha_1, \alpha_2, \ldots, \alpha_k$ which uniquely define the auction. Comparative statics follow directly:

$$\frac{\partial v^*}{\partial C} < 0 \tag{2.8}$$

$$\frac{\partial v^*}{\partial Z} > 0 \tag{2.9}$$

$$\frac{\partial v^*}{\partial \alpha_k} < 0 \quad \forall k. \tag{2.10}$$

2.3.1 Advertiser Size and Multi-homing

Consider now a set of $L > 0$ GSP auctions, indexed by $l = 1, \ldots, L$. Auctions may differ in both click-through rates and capacity. For simplicity, let participation cost $Z > 0$ remain fixed, although our results can easily be extended to consider varying participation costs. We add a subscript l for relevant parameters of each auction l.

Observe that the symmetric equilibrium in each auction form a symmetric equilibrium in the *extended game* in which each advertiser chooses which auction(s) to enter. Furthermore, the argument in the previous section implies the uniqueness of this equilibrium. In particular, advertisers' entry decisions are independent.

Corollary 2.3.4 *Let $v_1^*, \ldots, v_L^*$ be the cut-off values corresponding to the unique symmetric Bayes-Nash equilibria*

Table 2.1: *Advertiser size and multi-homing status (source 1)*

	Normalized Impression Count	Normalized Advertiser Count
Google only	17.34	1000.00
Yahoo only	26.01	338.65
Microsoft only	80.92	84.21
Google and Yahoo	663.58	209.26
Google and Microsoft	78.03	24.61
Yahoo and Microsoft	*	*
Google, Yahoo, and Microsoft	1000.00	65.76

Table 2.2: *Advertiser size and multi-homing status (source 2)*

	Proportion Ranked	Average Rank	Average Reach	Average Page-Views
Google only	0.696	17,428,974	4.19	0.22
Yahoo only	0.735	14,784,742	5.20	0.40
Microsoft only	0.705	15,838,598	4.40	0.41
Goole and Yahoo	0.862	8,305,741	17.11	1.18
Google and Microsoft	0.888	7,234,154	4.70	0.26
Yahoo and Microsoft	0.871	8,325,335	3.07	0.15
Google, Yahoo, and Microsoft	0.940	3,803,684	62.94	5.64

as in Theorem 2.3.1 and let B_l be advertisers entering auction l in this equilibrium. For each $v_l^*, v_{l'}^* \in V^*$ and $B_l, B_{l'} \in B$, if $v_l^* \geq v_{l'}^*$ then $B_l \subseteq B_{l'}$.

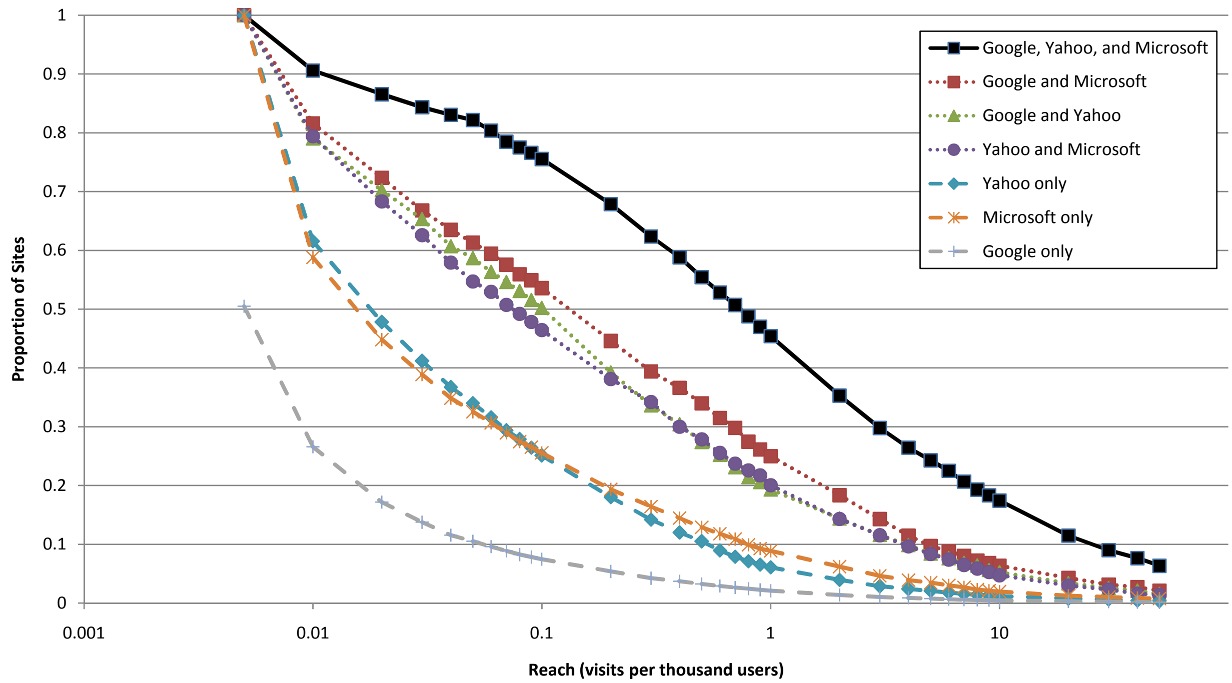
Proof 2.3.5 *Pick any v_l^* and $v_{l'}^*$ in V^* such that $v_l^* \geq v_{l'}^*$. Then any advertiser with value $v_j \geq v_l^*$ will enter auction l, hence $j \in B_l$. Moreover, by assumption, $v_j \geq v_{l'}^*$ hence $j \in B_{l'}$.*

Corollary 2.3.4 offers a testable implication: Large advertisers multi-home because they can spread participation cost Z across a large volume of ad purchases, whereas small advertisers find the participation costs too large to justify signing up with smaller platforms.

We test these claims with data from multiple services that track and preserve advertising at multiple ad platforms. We report normalized advertiser sizes based on independent observations by two different ad monitoring services.[4] Based on data from a first data collection service, Table 2.1 compares the size (impression count, normalized with maximum value set to 1,000) of advertisers that use one, two, or all three of the ad platforms we examine. Based on data from a second service, Table **??** measures the size of multi-homing and non-multihoming advertisers by other metrics: Proportion of sites achieving an Alexa

[4]By agreement with our data sources, we do not report their names.

Figure 2.1: *Advertiser reach and multi-homing status (source 2)*



ranking (data avialable only for approximately the 25 million most popular sites on the web), average rank (bottom-coding unranked sites at a rank of 40 million), as well as average "reach" (number of users visiting the site) and average page views. [5] Figure 2.1 plots the distribution of reach by multi-homing and non-multihoming advertisers.

By each metric, these tables are consistent with Corollary 2.3.4 and our model of advertisers' participation costs. In particular, the advertisers that purchase ads from all three platforms are strikingly larger than the

[5] Reach and page-views are reported per thousand users.

Table 2.3: *Advertiser presence rank*

	Presence Rank
Microsoft only	1
Yahoo only	2
Yahoo and Microsoft	3
Google Only	4
Google and Microsoft	5
Google and Yahoo	6
Google, Yahoo, and Microsoft	7

Table 2.4: *Ordered probit regressions of advertiser presence rank on reach, rank and page views*

	Reach	Rank	Page Views
Presence Rank	0.0832 ***	-1.01e-08 ***	0.8360 ***
	(0.0208)	(4.29e-10)	(0.2551)
Pseudo R^2	20.64	0.0079	0.0002
Model p-value	<0.001	<0.001	0.001
Observations	26286	26286	26286

$* - P < 0.10$

$** - P < 0.05$

$*** - P < 0.01$

advertisers that purchase ads only from one or two platforms: They buy more ad impressions (Table 2.1) and are more likely to be ranked by Alexa, achieve a lower average rank (i.e. greater traffic), larger reach, and more page-views (Table 2.2). Meanwhile, the advertisers who choose to use only Google are the smallest by far—further confirming that small advertisers tend not to multi-home. Finally, Figure 2.1 confirms that triply-multihoming advertisers have pointwise larger reach than double-multihoming advertisers which in turn have pointwise larger reach than single-homing advertisers. We also verify this relationship through a simple ordered probit regression. First, we rank combinations of search engines to form a presence rank—reflecting, for example, that Google is larger than Yahoo and Microsoft together. See Table 2.3. Then, we run an ordered probit of each advertiser's presence rank on its web site rank, reach, and page-views. See Table 2.4. The coefficients on reach, rank, and page-views are all positive and significantly different from zero.

2.4 Competing Auctions

In this section, we explore the micro-foundation of ad auction competition. Our approach is grounded in a fundamental tradeoff facing each search engine: On one hand, a search engine must provide high-quality results to satisfy users' requirements as quickly as possible. On the other hand, a search engine will reap greater revenues if it designs its listings to cause users to click more advertisements. Google early recognized this tradeoff: In their seminal 1998 paper, Google co-founders Brin and Page flagged an "inherent" incentive for a search engine to reduce the quality of its algorithmic results or otherwise encourage users to click on advertisements ([33]). The tradeoff remains timely: For example, when launching its new "Instant" search service, Google touted a savings of 2 to 5 seconds per search. Such time-savings could benefit users, but some advertisers found that Instant Search reduced clicks on their advertisements ([12]). Meanwhile, ([13]) points out that for some search terms, advertisements can be distracting or affirmatively harmful—providing a further tradeoff between satisfying users and increasing revenue.

We formalize search engines' response to the quality/revenue tradeoff via a *search cost* parameter, s. A greater search cost implies more advertisement clicks and more revenue to the search engine. Conversely, a search engine with a lower search gives users the "right" links more quickly, yielding fewer advertisement clicks but increasing user satisfaction and attracting, all else equal, more users.

Following Hotelling's model of consumers distributed on a circle ([23]), we build a model of consumer choice over search engines. Through this model, we endogenize a search engine's user base (capacity), and we incorporate a search cost parameter to capture the tradeoff between increasing consumer welfare versus increasing search engine revenue.

2.4.1 Consumer Choice

Suppose consumers are uniformly distributed around a circle with unit circumference. Suppose two search engines, A and B, occupy diametrically opposite locations on the circle. (The result can be extended to three or more engines.) A consumer who accesses search engine $l \in \{A, B\}$ receives value δ_l, less the search engine's cost of search, s_l. To access search engine l, the consumer incurs a cost td (where d reflects the consumer's distance from l and t gives a coefficient on distance). Thus, a consumer chooses engine A if the consumer is located at distance d such that:

$$\delta_A - s_A - td \geq \delta_B - s_B - t(1/2 - d). \tag{2.11}$$

Given a uniform distribution of consumers, the fraction of consumers choosing search engine A is $\frac{1}{t}(\Delta\delta - \Delta s + \frac{1}{2}t)$ where $\Delta\delta = \delta_A - \delta_B$. Similarly, the fraction of consumers choosing search engine B is $\frac{1}{t}(\Delta s - \Delta\delta + \frac{1}{2}t)$. As a result, A's capacity is $C_A = \frac{C}{t}(\Delta\delta - \Delta s + \frac{1}{2}t)$ where C is the total number of consumers in the market. Thus, search engine A's user base decreases as search cost increases, but increases in the value derived from use.

2.4.2 Ad Platform Selection of Search Cost

At the start of the second stage, advertisers have committed to the bidding process, and the number of entrants, $n^* = |B_l|$, and the pool of advertisers B_l (with valuations $\{v_1, ..., v_{n^*}\}$) are fixed. Then the revenue to search engine l is:

$$C_l(s_l)\Sigma_{j=1}^{n^*}\Sigma_{k=j}^{n^*}(\alpha_{k,l} - \alpha_{k+1,l})v_{(k+1)}$$

which is equivalent to:

$$C_l(s_l)\Sigma_{k=1}^{n^*}(\alpha_{k,l}-\alpha_{k+1,l})kv_{(k+1)}.$$

Since search cost is proportional to the expected number of clicks required to reach the desired information, any weighted aggregate click-through rate is strictly increasing in s_A: $\frac{\partial \Sigma_{k\in K} w_k \alpha_{kl}}{\partial s_A} > 0$ for any $\Sigma w_k = 1$. Moreover, assume that this function is linear, i.e. $\frac{\partial^2 \Sigma_{k\in K} w_k \alpha_{kl}}{\partial s_A^2} = 0$.

For tractability, suppose the search engine's technology is exponential (subsection B.1). Then $\alpha_{k,l} = \beta_l^k$. Define $\overline{V} = \sum_{k=1}^{n^*} kv_{(k+1)}$ and $w_k = kv_{(k+1)}/\overline{V}$.

Then we can rewrite the revenue as:

$$C_l(s_l)(1-\beta_l)\overline{V}\Sigma_{k=1}^{n^*} w_k \beta_l^k.$$

By the linearity assumption, we can parameterize the sum as $\Sigma_{k=1}^{n^*} w_k \beta_l^k = a_l s_l + b_l$, yielding the simplified expression:

$$C_l(s_l)(1-\beta_l)\overline{V}(a_l s_l + b_l).$$

This revenue function has a unique internal maximum attained by the following first order conditions:

$$s_l = \frac{1}{1+a_l}(\Delta\delta + \frac{1}{2}t - b_l + s_{l'}) \quad (2.12)$$

$$s_{l'} = \frac{1}{1+a_{l'}}(-\Delta\delta + \frac{1}{2}t - b_{l'} + s_l). \quad (2.13)$$

2.4.3 Comparative Statics

We now consider two special cases: two search engines have pointwise equal technologies ($a_l = a_{l'} = a, b_l = b_{l'} = b$), and one search engine pointwise dominates the other ($a_l > a_{l'}, b_l > b_{l'}$).

Pointwise Equality

Consider a search engine that is identical to its competitor ($a_l = a_{l'} = a, b_l = b_{l'} = b$) except that it enjoys a single advantage: a higher value delivered to consumers during search ($\delta_l > \delta_{l'}$). Such a search engine can leverage that strength by raising search cost while retaining higher capacity. The following theorem formalizes that advantage:

Theorem 2.4.1 *In the two-stage game with two auction platforms l and l' with pointwise identical baseline technologies, if $\delta_l \geq \delta_{l'}$, then in equilibrium: $s_l \geq s_{l'}, C_l \geq C_{l'}, \Sigma_{k\in K} w_k \alpha_{kl} \geq \Sigma_{k\in K} w_k \alpha_{kl'}, v_l^* \leq v_{l'}^*$ and $B_l \supseteq B_{l'}$.*

Proof 2.4.2 *The first order conditions (2.12) and (2.13) require* $s_l = \frac{\Delta\delta}{2+a} + \frac{\frac{1}{2}t-b}{a}$ *and* $s_{l'} = \frac{-\Delta\delta}{2+a} + \frac{\frac{1}{2}t-b}{a}$*. The resulting capacities are:*

$$C_l^* = C\left(\frac{1}{2} + \frac{a\Delta\delta}{t(2+a)}\right) \tag{2.14}$$

$$C_{l'}^* = C\left(\frac{1}{2} - \frac{a\Delta\delta}{t(2+a)}\right). \tag{2.15}$$

Thus, capacity, technology and search cost are all monotonic in δ*; if* $\delta_j \geq \delta_{l'}$ *then* $C_l \geq C_{l'}$, $\Sigma_{k\in K} w_k \alpha_{kl} \geq \Sigma_{k\in K} w_k \alpha_{kl'}$ *and* $s_l \geq s_{l'}$.

The threshold values and advertiser set relation follow from Corollary 2.3.4.

Pointwise Dominance

Consider a search engine l that enjoys technology superior to its competitor l': $a_l > a_{l'} \geq 0, b_l > b_{l'} \geq 0$. In that case, search engine l receives greater market share even though its value to consumers matches its competitor. In particular, l and l' pick search costs s such that

$$\Delta s = \frac{1}{a_{l'} + a_l + a_{l'}a_l}\left[(a_{l'} + a_l)\Delta\delta + (a_{l'} - a_l)\frac{t}{2} - a_{l'}b_l + a_l b_{l'}\right]. \tag{2.16}$$

Suppose $\Delta\delta = 0$. Then $\Delta s < 0$ but $\Delta C > 0$ since $a_l > a_{l'} > a_{l'}\left(\frac{t-2b_l}{t-2b_{l'}}\right)$. That is, l sets lower search cost and enjoys greater market share.

Alternatively, suppose a search engine l enjoys pointwise dominant technology (as in subsection 2.4.3), but gives consumers less value than its competitor l'. (That is, $\delta_l < \delta_{l'}$.) Then l will receive lower market share ($C_l < C_{l'}$) if its technology is not sufficiently greater than l''s. Formally, $\Delta C < 0$ if and only if $2a_{l'}a_l\Delta\delta + a_l < a_{l'}\left(\frac{t-2b_l}{t-2b_{l'}}\right)$.

2.5 Joining Auctions

In this section, we establish a concept of "joining" auctions such that their available positions are pooled, and all advertisers participating in one auction automatically participate in the other. What happens to advertiser welfare and ad platform revenue if two ad auctions are joined? These questions take on special relevance in light of the 2009 partnership between Microsoft and Yahoo, as well as a 2008 proposed partnership between Google and Yahoo (ultimately aborted after antitrust regulators raised concerns).

We begin with several definitions.

Definition 2.5.1 *A set of click-through rates has the property of* diminishing differences *if* $\alpha_k - \alpha_{k+1} \geq \alpha_{k+1} - \alpha_{k+2}$ *for each* $k \leq K$.

Definition 2.5.2 *Auction* l *and auction* l' join *to form auction* $\tilde{l}$ *if auction* $\tilde{l}$ *has capacity* $C_{\tilde{l}}$ *such that*

$$\max\{C_l, C_{l'}\} \leq C_{\tilde{l}} \leq C_l + C_{l'} \tag{2.17}$$

and if auction $\tilde{l}$ *has click-through rate*

$$\alpha_{k\tilde{l}} = max\{\alpha_{kl}, \alpha_{kl'}\} \tag{2.18}$$

where α_{kl} *(*$\alpha_{kl'}$*) is the click-through rate of slot* k *in auction* l *(*l'*).*

By taking the capacities of each auction as the size of the set of consumers choosing the engine, inequality (2.17) bounds $C_{\tilde{l}}$. Because some consumers use multiple search engines, we allow for overlap of capacity between two ad platforms—meaning a joined ad platform might have less capacity than the sum of capacities of its contributors. Furthermore, we assume that no consumer of either engine is lost upon join.

When auctions join, what click-through rates result? We envision ad platforms choosing the best components of each contributor, which implies click-through rates given by the stronger of the joining platforms. Hence the approach in (2.18).

2.5.1 Joining Auctions to Make All Advertisers Better Off

In the following theorem, we provide a condition in which a joined auction offers a sufficient improvement in capacity and technology to make every advertiser weakly better off ex post.

Theorem 2.5.3 *Suppose auctions* l *and* l' *join to form auction* $\tilde{l}$*. Let* $C_l \geq C_{l'}$ *and* $\alpha_{kl'} \geq \alpha_{kl}$ *for each* $k \leq K$*. Then every advertiser is weakly better off ex post if:*

$$C_{\tilde{l}} \geq \frac{C_{l'}\alpha_{kl'} + C_l\alpha_{kl}}{(N-k)\alpha_{k+1,l'} - (N-k-1)\alpha_{kl}} \quad \forall k \leq K. \tag{2.19}$$

Proof 2.5.4 *Suppose auction* l *and* i' *join to form* $\tilde{l}$ *and, without loss of generality, let* $n_l^* \geq n_{l'}^*$.

Consider an advertiser who does not receive any position in either auction before the join. Such an advertiser is ranked in position $k > n_l^*$*. Because that advertiser already achieves utility of 0, the joined auction* $\tilde{l}$ *cannot make it worse off.*

Consider an advertiser who receives a position in each of the auctions before the join. We introduce new notation to characterize ex post utility: denote by $u_l(v,k)$ the utility of a advertiser with value v who wins position k in auction l. Under VCG, we have:

$$u_l(v,k) = C\alpha_{kl}v - C \sum_{n^* > j \geq k} (\alpha_{jl} - \alpha_{j+1,l})v_{(j+1)} \tag{2.20}$$

where $\alpha_{kl} = 0$ for $k \geq K$.

Prior to the join, an advertiser with valuation v who receives positions in both auctions must receive position $k \leq n^*_{l'}$ in auction l', yielding utility $u_l(v,k) + u_{l'}(v,k)$.

Using assumption (2.19) we get:

$$\begin{aligned} &C_{\tilde{l}}\alpha_{k\tilde{l}}v - C_l\alpha_{kl}v - C_{l'}\alpha_{kl'}v \geq C_{\tilde{l}}(N-k)(\alpha_{k\tilde{l}} - \alpha_{k+1,\tilde{l}})v \geq \\ &\qquad C_{\tilde{l}} \sum_{n^*_{\tilde{l}} > j \geq k} (\alpha_{j\tilde{l}} - \alpha_{j+1,\tilde{l}})v_{(j+1)} \geq C_{\tilde{l}} \sum_{n^*_{\tilde{l}} > j \geq k} (\alpha_{j\tilde{l}} - \alpha_{j+1,\tilde{l}})v_{(j+1)} - S, \end{aligned} \tag{2.21}$$

where $S = C_{l'} \sum(\alpha_{jl'} - \alpha_{j+1,l'})v_{(j+1)} + C_l \sum(\alpha_{jl} - \alpha_{j+1,l})v_{(j+1)}$.

The last inequality (2.21) implies:

$$\begin{aligned} &C_{\tilde{l}}\alpha_{k\tilde{l}}v - C_{\tilde{l}} \sum(\alpha_{j\tilde{l}} - \alpha_{j+1,\tilde{l}})v_{(j+1)} \geq \\ &\qquad C_l\alpha_{kl}v - C_l \sum(\alpha_{jl} - \alpha_{j+1,l})v_{(j+1)} + C_{l'}\alpha_{kl'}v - C_l \sum(\alpha_{jl'} - \alpha_{j+1,l'})v_{(j+1)}, \end{aligned}$$

which is equivalent to:

$$u_{\tilde{l}}(v,k) \geq u_l(v,k) + u_{l'}(v,k). \tag{2.22}$$

Therefore an advertiser receiving a position in both auctions is better off after the join.

Consider an advertiser with valuations v who receives position k in auction l with $n^*_l \geq k > n^*_{l'}$. Such an advertiser will have ex post utility of $u_l(v,k)$. Since we have already shown (2.22), such an advertiser, who wins a single position, is also better off after the join.□

The conditions in Theorem 2.5.3 stipulate intuitive requirements for advertisers to gain from a joined ad auction: the resulting click-through rate and auction capacity must be sufficiently improved relative to the offerings of the ad auctions when separate. First, the auction with fewer advertisers must add value to the join through a point-wise larger click-through rate. Second, there must be minimal overlap between the two auctions, so that $C_{\tilde{l}}$ is sufficiently larger than both C_l and $C_{l'}$. It is necessary for $C_{\tilde{l}}$ to be sufficiently large so that advertisers in auction l (who already face higher prices due to more advertisers in l) gain sufficiently from joining the two auctions. An example with exponential click-through rates is given in the appendix.

2.5.2 Joining Auctions that Make Some Advertisers Worse Off

In this section we show that it a joined auction can negatively affect overall advertiser welfare. Consider the following definition.

Definition 2.5.5 *Auction l is* uniformly stronger *than auction l' if and only if $C_l \geq C_{l'}$, and $\alpha_{kl} \geq \alpha_{kl'}$ for each k. We denote this by $A_l \geq_{us} A_{l'}$.*

If one of the auctions is uniformly stronger than the other and if the resulting capacity remains equal to its original capacity (i.e. the uniformly weaker auction represents a subset of consumers of the stronger auction), then joining the auctions will make some advertisers weakly worse off. The following theorem identifies sufficient conditions that make advertisers worse off:

Theorem 2.5.6 *Suppose $A_l \geq_{us} A_{l'}$, and auctions l and l' join in the manner of Definition 2.5.2. If $C_{\tilde{l}} = C_l$, then any advertiser that wins in both auctions is worse off.*

Proof 2.5.7 *The assumption implies that $\alpha_{k\tilde{l}} = \alpha_{kl} \geq \alpha_{kl'}$ for each $k \leq K$ and $C_{\tilde{l}} = C_l$. Then the joined auction will be identical to auction l, and thus $u_{\tilde{l}}(v,k) = u_l(v,k)$ for any v and k. Prior to the join, if an advertiser wins position $k \leq n^*_{l'}$ (and thus the advertiser receives a placement in both auctions), then its total pre-merger utility of $u_l(v,k) + u_{l'}(v,k)$ is greater than its post-merger utility $u_{\tilde{l}}(v,k)$.$\square$*

2.5.3 Joint Auctions with Endogenous Capacity and Technology

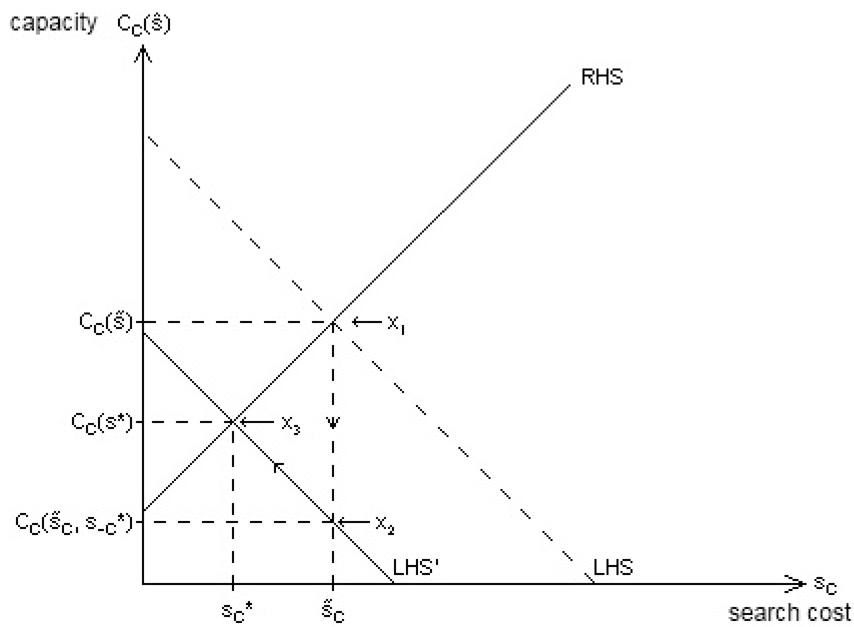
We now return to the foundations of consumer choice to more closely model the change in capacities when joining auctions. Suppose there are three engines $l \in \{A, B, C\}$ positioned on the unit circle with uniformly distributed consumers. Individually, each engine solves the following optimization:

$$\max_{s_l} C_l(\hat{s})(1-\beta_l)\overline{V}(a_l s_l + b_l) \quad \forall l \tag{2.23}$$

where $\hat{s}$ represents a vector of three search costs. Denote the solution triple as $(\tilde{s}_A, \tilde{s}_B, \tilde{s}_C)$.

Without loss of generality, let engines A and B join; the new platform's technology will be the pointwise maximum of A and B, and the new platform will retain the sum of the prior capacities. Moreover, consider a case in which one engine, A, has a technology that pointwise dominates that of engine B ($a_A > a_B$ and

Figure 2.2: *Joint auctions with endogenous capacity and technology*



$b_A > b_B$) and $C_A < C_B$. The optimization program under a joined platform will be:

$$\max_{s_A, s_B} \left(C_A(\hat{s}) + C_B(\hat{s})\right)(1 - \beta_A)\overline{V}(a_A s_A + b_A). \tag{2.24}$$

When A and B join, the joined platform replaces the technology of B with A. Then the marginal benefit of increasing search cost on engine B is the increase in C_A (increase in technology no longer applies since technology B has been replaced) while the marginal cost is a reduction in C_B. By noting that $C - C_C(\hat{s}) = C_A(\hat{s}) + C_B(\hat{s})$ and $\frac{\partial C_l}{\partial s_{l'}} < 0 \quad \forall l \neq l'$, we can conclude that total capacity will only decrease in response to increasing s_B. Then the solution to the joint maximization becomes $s_B^* = 0 \leq \tilde{s}_B$.

On the other hand, the marginal cost of increasing s_A is exactly offset by the increase in C_B while the marginal benefit to technology remains. Then the joined platform will increase s_A such that $s_A^* \geq \tilde{s}_A$.

What will be the effect on the resulting individual and aggregate capacity and technology? In order to predict engine C's best response to the join between A and B, we examine the first order condition of (2.23) for C:

$$C_C(s_c, s_{-c}) = -\frac{1}{a_C}\frac{\partial C_C}{\partial s_c}(a_C s_C + b_C). \tag{2.25}$$

Note that since consumers' utilities are linear in the cost of search, the partial derivative on the RHS is a constant. Point x_1 in Figure 2.2 illustrates the equilibrium market share of C prior to join. After the join, A and B adjust their search costs from $\tilde{s}_{-c}$ to s^*_{-c}, effectively reducing C's market share. Thus, C would find it undesirable to retain the former level of $\tilde{s}_C$ would be too high (point x_2), and C best-responds by reducing its search cost to s^*_C. At location x_3, C's market share has decreased while the combined market share of A and B has increased. Therefore

$$C_A(s^*) + C_B(s^*) \geq C_A(\tilde{s}) + C_B(\tilde{s}) \geq \max\{C_A(\tilde{s}), C_B(\tilde{s})\}.$$

Thus, both A and B enjoy increased capacity. Moreover, A's superior technology replaces B's technology, so more advertisers will enter B. Similarly, by retaining the same level of technology while increasing capacity, A attracts more advertisers. Because A and B each attract more advertisers, the expected payment of advertisers on A and B will increase.

2.6 Conclusion

Increasing search engine concentration makes it important to understand how advertisement auctions compete. If each buyer were restricted to entering exactly one auction, then auction platforms could compete on reserve price or on the cost of entry. But in fact advertisers can multi-home, albeit with additional costs for each ad platform they join. With multi-homing possible, an advertiser enters an ad auction if the expected benefit to entry outweighs the associated cost, independent of outcomes in other auctions. Hence traditional considerations of competition on reserve price or entry cost do not apply. We therefore turn to the supply side of ad auctions: search engines compete on capacity (each user picks only one search engine) and compete by setting search cost to balance advertisement clicks (auction real estate) against user utility.

Using the model's first stage equilibrium entry strategies, we find empirical evidence for the result that multi-homing advertisers are associated with greater traffic and larger reach. In the second stage, comparative statics around consumer choice parameters offer a stylized view of search engine market conditions: Google has greater market share and (some suggest) superior technology, while Yahoo retains higher market share than Microsoft despite (some say) inferior technology.

Our approach informs understanding of the advertiser welfare implications of joining auction platforms. Joining ad platforms can attract substantial regulatory attention: In November 2008, the Department of Justice planned to file antitrust charges to stop the proposed Google-Yahoo transaction. Then, in 2009, the Department of Justice approved the proposed partnership between Microsoft and Yahoo. At first glance it might seem paradoxical for Microsoft or the Department of Justice to claim that the Google-Yahoo transaction is undesirable, for advertisers and for the economy as a whole, while the Microsoft-Yahoo transaction offers net benefits. But our analysis suggests that that conclusion is entirely possible. In particular, by creating a joined ad platform of larger size than Microsoft or Yahoo alone, the transaction lets advertisers spread participation costs over a larger purchase—making it worth the while of small to midsize advertisers to sign up with the joined Microsoft-Yahoo platform.

We view our contribution as threefold. First, whereas standard models of online advertising take advertisers' participation as exogenous, we explicitly model an advertiser's decision to use or ignore a given ad platform, and we provide empirical support for our model. Second, we offer a model of search cost to demonstrate competition across platforms, and we provide empirical support for our model. Third, we analyze the prospect of joining auctions and characterize when such joins do and do not increase welfare.

Chapter 3

Consumer Search Pattern in Internet Advertising[1]

3.1 Introduction

Internet advertising is the fastest and the only growing form of advertising in the industry. Search advertising generated $25 billion in 2008 from approximately 100 billion searches performed mostly on just a few dominant search engines. Of the total spending on internet advertising, search advertising accounted for 47%, display ads 34%, and classifieds 19%. The market is controlled by an oligopolistic regime consisting of three platforms with the following shares: Google has 60%, Yahoo 18.1%, and Microsoft 11.8%. Given the interdisciplinary nature of many relevant topics of research in this field, much literature has been published in computer science, operations research, economics and marketing.

A wealth of theoretical research has produced useful results regarding optimality of ad auction designs for ex-post efficiency ([3, 14]) and a general framework for analysis of position auctions ([2, 37]). But largely due to limited availability of data, specifically on any one of the three search platforms, there is no developed empirical literature on consumer search pattern in search advertising using structural analysis. However, there are recent studies in different contexts: [25] show that the release of a new album increases sales of old albums, and the increase is substantial and permanent, especially if the new release is a hit. Authors pursue this explanation by developing and estimating a model of market demand based on a

[1] I thank Microsoft Research, New England for the generous support on data access and provision. I write on my own behalf, not at Microsoft's request or for Microsoft's benefit. Do not circulate without permission.

binary consumer learning model. In [11], the author uses observed search intensities from the online book industry to estimate search cost distributions that allow for asymmetric consumer sampling.

By viewing the search platform as an intermediary in a two-sided market, the platform is a match-service between advertisers and consumers. On one side of the market, consumers type in a keyword, then click and review K ads to complete a task while facing a per-period search cost. The dynamics of optimal search process will be governed by the fact that quality of each ad, as measured by the likelihood a match occurs, is private information to advertisers and the engine. Additionally, consumers receive signal via a short text provided by each ad and draws a bernoulli variable each period until a match occurs (task is completed in a binary sampling). On the other side of the market, N advertisers acquire slot $k < K$ for a given keyword(s) a la Generalized Second Price ([14]) auction, design appropriate ad text and link to represent their ads and pay per click.

Using the platform framework, I develop and compute a dynamic model of search in internet advertising. Micro-level browsing data from Microsoft's Bing.com (formerly known as Live.com) is used for structural estimations. The model predicts that users do not click on any ad with weak signals due to accumulating search cost and monotonicity of the value function. Rational search reveals a cascading pattern: the user clicks on a sufficiently high, highest-signal ad first, then moves on to the ad with the next highest conditionally expected probability of match once his assessment of likelihood of match on the current ad degrades over time. The user exits when maximum assessment of likelihood of match over all ads is below a threshold value. This essay provides a novel approach to understanding rational herding behavior when product quality is only partially unraveled.

There are two guiding assumptions on developing the consumer search model: consumers are epsilon-small and individual actions are negligible (assuming away collusive behavior) but individuals are rational and search optimally. Hence I do not impose a game-theoretic environment but employ dynamic programming. Then using such model, I use data from Microsoft which include over 20 million micro-search records on quasi-unique consumer identifier, time/date in which search occurred, keyword used to search, ads exposed, ads clicked, time spent on the advertiser's target site and conversion for estimation.

Despite the inordinate magnitude of breadth of data, I face several econometric hurdles. First, I must find an appropriate set of keywords with sufficient variation in ad configurations over a set of nearly-homogeneous advertisers. Given such keywords, I must require two further levels of restrictions: the associated search session must have generated sufficient number of clicks from users and conditional on a click, dwell times must be available. Due to user privacy protection agreement, Microsoft is limited to infer dwell time on an ad by taking the difference between the time when user clicks out and the time when user returns to any site belong to Microsoft Network. Hence in the estimation procedure, biases in dwell time measurement must be appropriately adjusted after a careful analysis of possible qualitative changes in dwell times. Moreover, conversion - a proxy for completion of task - is an inferred measurement which rely on consumer having viewed a secure web-page (page links that begin with 'https://') and a sufficiently high dwell time is recorded. Measuring consumer search cost and value of task completion will be limited to inferring data from aggregate statistics - median household income and average internet purchase amounts.

The paper is organized as follows: section 2 provides the modeling framework and discussion for ads under random ordering; section 3 provides description of the data set in greater detail, and proposes a measure for quality of ads to use as an empirical building block; section 4 gives detail on estimation technique; and section 5 concludes.

3.2 Consumer Search Pattern: Discrete Time and Random Ordering of Ads

Due to lack of variations in consumer-level data, I use a model which suppresses consumer heterogeneity by normalizing individual-level parameters that simplifies the estimation process. This allows for non-parametric primitive distribution estimations but lacks a robust micro-foundation. [2]

Consumers optimize their search to complete a task at a cost c per period and receive a payoff of 1 upon completion or 0 upon failure. Denote the set of quality of each advertiser as $Q = \{q_1, q_2, ..., q_K\}$. Each q_k is private information of each advertiser and drawn i.i.d. from F_q. This distribution is common knowledge. Assume ads are randomly ordered and thus their positions are not correlated with q_k. After entering a query, he views K sponsored links and receives some signal s_k from each ad-text associated with every ad. In each period t, user decides to either exit search, or continue to review ad k. During the review, user draws a bernoulli variable with latent probability $p_k(q_k)$ (likelihood of match) drawn i.i.d. from F_p; if it is true then user converts and terminates session, else he goes on to period $t+1$. User employs Bayesian updating to infer underlying probability of match in each period.

Upon observation of the signal vector $\mathbf{s}$, the consumer forms a set of prior beliefs regarding the

[2]Segal and Jeziorski 2009 ([32]) uses random effects utility to generalize consumer heterogeneity, but the model depends heavily on parametric assumptions.

probability of match, $\mu_k^0 = E[p|p \geq s_k] \forall k \leq K$.

Denote $x_{(i)}$ as the i^{th} *reverse-order statistic* so that $x_{(i)} > x_{(i+1)} > ... > x_{(K)}$. Also define $l(k)$ as a map from index k to $l(k)^{th}$ *reverse-order statistic* and $m(i) = k$ be the map from the i^{th} *reverse-order statistic* to ad k.

Let $v(\mathbf{s}, H^t, t)$ be the expected payoff in a search session with signal $\mathbf{s}$, history of review result H^t at period t.

In period $t \geq 1$, if consumer chooses to stay on the destination page, consumer expects to receive $g(s_k, h_k^t) - c + \delta(1 - g(s_k, h_k^t))v(\mathbf{s}, H^{t+1}, t+1)$ where δ is his discount factor for future payoffs and $g(s_k, h_k^t)$ is his updated belief of probability of match given history on ad k at t. It is important to note that because ads are randomly ordered, his experience with ad k does not affect his belief of probability of match on ads $j \neq k$. Upon completion of task in period t, his payoff is simply $1 - c$ and the session terminates. Thus the consumer search bellman equation becomes:

$$v(\mathbf{s}, h^t, t) = max\{0, g(s_k, h_k^t) - c + \delta(1 - g(s_k, h_k^t))v(\mathbf{s}, h^{t+1}, t+1)\} \tag{3.1}$$

In order to fully characterize equation 3.1, one must first understand how $g(s_k, h_k^t)$ evolves over time. Intuitively, given a history of outcomes at the end of period $t \geq 0$, h^t,

$$g(s_k, h_k^t) = E[p|p \geq s_k, h^t] \tag{3.2}$$

3.2.1 Properties of the Value Function

The following lemma presents a generalized version of updating belief of probability of match on ad k:

Lemma 3.2.1 *Suppose ads are randomly ordered. In period $t \geq 1$, if consumer is viewing ad k, belief regarding the probability of match on ad k is updated using the following rule:*

$$g(s_k, h_k^t) = \frac{E[p(1-p)^t|s_k]}{E[(1-p)^t|s_k]} \quad \forall t \geq 1 \tag{3.3}$$

$$g(s_j, h_j^{t+1}) = g(s_j, h_j^t) \quad \forall j \neq k \tag{3.4}$$

Proof 3.2.2 *To show (3.4), I use the assumption of random ordering of ads: new information revealed under ad k does not yield any information regarding the quality of of ad $j \neq k$.*

To show (3.3), I solve for the expression in (3.2) using the model framework: in period $t > 1$, I know that user has drawn an independent sequence of t failures. Hence I let $h^t = \{z_k^1 = ... = z_k^t\}$. Consider the density $\tilde{f}(p) = f_{p|s_k}(p|s_k)$. By Bayes' rule, I get:

$$\tilde{f}(p|h^t) = \frac{f_h(h^t|p)\tilde{f}(p)}{f_z(z_k^1 = ... = z_k^t = 0)} \tag{3.5}$$

$$= \frac{(1-p)^t \tilde{f}(p)}{\int (1-p)^t \tilde{f}(p) dp} \tag{3.6}$$

Note that expression (3.3) is not monotonically non-increasing over t because $p(1-p)^t$ is both concave and convex over the different sub-intervals of $[0,1]$ for different t. In fact, if p is distributed $Beta(\alpha, \beta)$, $g(s_k, h_k^t) \geq g(s_k, h_k^{t+1})$ if and only if $\alpha \leq 1$, hence the distribution must not possess a mode. This parametric restriction will play an important role in the estimation procedures described in the latter sections.

Lemma 3.2.3 *Suppose the value function $v(x) = max\{0, x - c + \delta(1-x)v(y(x))\}$ is continuously differentiable on $[0,1]$ and $c, \delta < 1$ are constants. Also let $y'(x) < 0$ on $[0,1]$. Then $v(x)$ is monotonically non-decreasing in $x \in [0,1]$.*

Proof 3.2.4 *Suppose $v(x) = 0$. Then $v(x)$ is certainly monotonically non-decreasing in x. Now suppose $v(x) = x - c_1 + \delta(1-x)v(y(x))$.*

Compute the derivative of the value function:

$$v'(x) = 1 - \delta v(x) + \delta(1-x)v'(y(x))y'(x) \tag{3.7}$$

Define set $S = \{x | v'(x) \leq 0 \forall x \in [0,1]\}$. If $S = \emptyset$ then I am done. Suppose $S \neq \emptyset$. Let $b = \inf S$.

Claim 3.2.5 *b is an element of S.*

Suppose $b \notin S$. Then by definition of b and continuity of v', for any $\epsilon > 0$, $v'(b+\epsilon) \leq 0$. But by the intermediate value theorem (IVT), there exists a value $\tilde{b} \in (b, b+\epsilon)$ such that $v'(b) > v'(\tilde{b}) > v'(b+\epsilon)$. If $v'(b+\epsilon) = 0$ then I have a contradiction. If $v'(b+\epsilon) < 0$ then I can set $\tilde{b}$ such that $v'(\tilde{b}) = 0$ which imply every point in S is bounded above by $\tilde{b}$ while $\tilde{b} > b$ hence I have another contradiction. Hence $b \in S$.

Claim 3.2.6 $v'(b) = 0$

Since $b \in S$, $v'(b) \leq 0$. Suppose $v'(b) < 0$. Then by taking the IVT, there exists $0 < \tilde{b} < b$ such that $v'(b) = 0$. But this means that b cannot be the infimum of S hence I have a contradiction.

Claim 3.2.7 *For any* $\tilde{x} < b$, $v'(\tilde{x}) > 0$

Since b *is the minimum of set* S, *any* $\tilde{x} < b$ *is not an element of* S, *hence* $v'(\tilde{x}) > 0$.

Claim 3.2.8 *For any* $x \in [0,1]$, $y(x) < x$

By 3.3 under Lemma 3.2.1, the claim is obviously true.

Now I return to the expression for v' *and evaluate it at* b:

$$v'(b) = 1 - \delta v(b) + \delta(1-b)v'(y(b))y'(b) = 0 \tag{3.8}$$

Using the upperbound for v, *we know* $1 - \delta v(b) > 1 - \delta(1-c) > 0$. *Also, because I found* y *is non-decreasing on the unit interval,* $y'(b) \geq 0$. *This means that* $v'(y(b))$ *must be negative. Since I have shown that* $y(b) < b$ *and* $b = \inf S$, v' *evaluated at* $y(b)$ *must be positive. Hence I have a contradiction.*

Note that Lemma 3.2.3 can be proven for *any* $y(x)$ that is monotonically non-decreasing in x. This allows one to fully analyze the characteristics of consumer's expected payoff function, v with a generalized belief-updating method that 1. reduces the quality of match upon failures (monotonically non-increasing in t) and 2. sets higher probability of match upon update when the prior is also higher (monotonically non-decreasing in x).

Proposition 3.2.9 *In optimal search pattern, consumers click from highest-quality (as measured by likelihood of match) ad to next highest-quality ad. Thus consumers expect to receive* $v(\mathbf{s}/s_{(2)}, H^t, t)$ *after the* i^{th} *click during a search session.*

Proof 3.2.10 *Since a consumer of any type is an expected utility maximizer, the very first click he makes must yield* $v(s_{(1)}, H^0, 0)$ *(since the value function is monotonically non-decreasing in signal by lemma 3.2.3) and his experience with ad* k *does not affect his belief of quality of ads* $j \neq k$. *Then conditional on having explored advertiser associated with the highest signal,* $s_{(1)}$, *when choosing another ad to view, he must click on the ad with the next highest-quality ad, following the bellman equation in 3.1.*

Lemmas 3.2.1 and 3.2.3 and equation 3.1 describe a decision-making process in which consumers choose to spend periods in reviewing advertiser k then eventually switch at some threshold maximum period $\tilde{t}_k$. The value function is initially set to $v(s_{(1)}, H^0, 0) = \max_{k,t} v(s_k, H^t, t)$ but begins to decrease after each period of unsuccessful match as updated probability of success also decreases. Then at $\tilde{t}_k$, the value function is set to $v(s_{(2)}, H^{t_k}, t_k)$ and the entire process repeats itself. It is clear that the implicit 'time limit' placed on reviewing each ad is set endogenously by the set of signals, $\mathbf{s}$. Formally, I present this result in the following proposition:

Proposition 3.2.11 *For each ad $k \leq K$, consumer never spends more than t_k periods to review the ad before switching. Moreover, each t_k is some function of $g(s_k, h_k^{t_k})$ and $g(s_{(l(k)+1)}, h_{l(k)+1}^{t_k})$.*

Proof 3.2.12 *At period t_k, the period immediately before switching, consumer's value function must be such that:*

$$\begin{aligned} g(s_{(l(k)+1)}, h_{m(l(k)+1)}^{t_k}) - c + \delta(1 - g(s_{(l(k)+1)}, h_{m(l(k)+1)}^{t_k}))v(\mathbf{s}, H^{t_k+1}, t_k+1) &\leq \\ g(s_k, h_k^{t_k}) - c + \delta(1 - g(s_k, h_k^{t_k}))v(\mathbf{s}, H^{t_k+1}, t_k+1) &= \\ v(\mathbf{s}, H^{t_k}, t_k) \end{aligned}$$

and

$$\begin{aligned} v(\mathbf{s}, H^{t_k+1}, t_k+1) &\geq \\ g(s_k, h_k^{t_k+1}) - c + \delta(1 - g(s_k, h_k^{t_k+1}))v(\mathbf{s}, H^{t_k+2}, t_k+2) \end{aligned}$$

Now prior to switching, if consumer is viewing ad k, following 3.4 I have $g(s_{l(k)+1}, h_{m(l(k)+1)}^{t_k}) = g(s_{l(k)+1}, h_{m(l(k)+1)}^{t_k+1})$. Since v is monotonic in s_k:

$$g(s_k, h_k^{t_k+1}) < g(s_{l(k)+1}, h_{m(l(k)+1)}^{t_k}) \leq g(s_k, h_k^{t_k}) \tag{3.9}$$

It was shown in 3.2.1 that g is monotonically non-decreasing in s_k and non-increasing in t. Hence it remains to show that for some $\tilde{t}$,

$$g(s_k, h_k^{\tilde{t}+1}) < g(s_{l(k)+1}, h_{m(l(k)+1)}^{\tilde{t}}) \tag{3.10}$$

Since $g(s,t)$ approaches 0 as t goes to infinity for any s, I know at some sufficiently large $\tilde{t}$, inequality 3.10 is true.

Using the lemmas and propositions developed, I can exploit the structure of of consumer expected payoff to formalize an estimation strategy.

Table 3.1: *Summary statistics*

Position	% of Clicks	Avg. Dwell Time (In Seconds)	Avg. Quality
ML-1	11.74%	152.6456	0.71%
ML-2	11.68%	258.9321	4.34%
ML-3	11.04%	178.4019	1.89%
ML-4	10.68%	197.004	0.19%
SB-1	10.98%	190.9281	0.00%
SB-2	10.97%	558.0493	0.05%
SB-3	10.97%	150.4906	0.00%

3.3 Empirical Analysis

3.3.1 Data Description

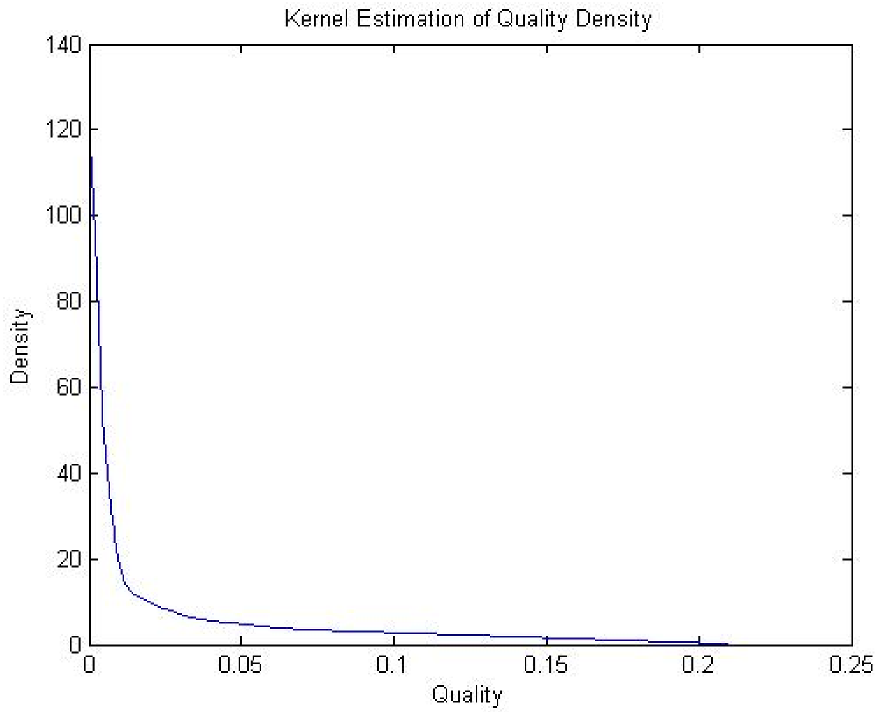
Using internal Microsoft search data, I extracted a sample of approximately 40,000 clicks resulting from a set of keywords in March 2010. There are 500 unique ads (text and link) shown to approximately 10,000 consumers. Each observation identifies the search session, ad, position, time of click, time spent reviewing the landing page (dwell time) and conversion. A search session is defined as the period in which a user begins by 1) entering a search query, 2) views K slots of sponsored links, 3) clicks on one of the ads, 4) reviews the ad destination page, 5) end the session or return to the same search result page and repeat the process as in 2). Variation in consumer characteristics is rarely available due to specialized profiling by the engine. Amount of transaction, if occurred, is not available. Dwell time is available only when the user has returned to Bing.com; it is available in approximately 20% of clicks (total: 40,000). A basic summary statistics is available in Table 3.1.

3.3.2 Observable Primitives: Conversion and Dwell Time

In the empirical analysis, I am interested in the effects of configuration of ad placements on the quality of user experience. In order to measure quality of ad, I use the sample average of *IsConverted* which takes on the value 1 whenever a pixel located on the objective page within the advertiser's destination website is reached (such pixel is placed by the advertiser). Else, it is set to 0. The non-parametric kernel density estimate of quality is shown in Figure 3.1.

Dwell times were separated into 14 partitions, representing intervals of 30 seconds beginning from 0. Dwell time is the difference in time between the moment a user lands on the destination page and the moment he returns to Microsoft search (not necessarily the original search result page which prompted the

Figure 3.1: *Estimated density of quality*



ad). Hence a fraction of clicks do not record dwell times and such data points are recorded as missing. To evaluate the relative quality of such missing dwell times, I use a simple regression on each dwell time interval while controlling for both position and advertiser fixed effects.

Formally:

$$IsConverted_i = \beta_3 \cdot \sum_{j \in J} dt_{ij} + \beta_2 \cdot \sum_{k \in K} position_{ik} + \beta_1 \cdot \sum_{a \in A} \mathbf{1}\{adid_{ia}\} + \beta_0 + \epsilon_i \tag{3.11}$$

Explanatory variables on the RHS of 3.11 are indicators for dwell times that belong to one of 14 partitions, set of indicator variables for each position and ad. Result of the regression is shown in Table 3.2. I find that regression coefficient for missing dwell times is within the range of coefficients for dwell times in the intervals 240-270 and 270 to 300 seconds. This is consistent with the intuition that a missing dwell time will tend to be drawn from the right tail of the distribution and signals a relatively satiated user.

Of the 400 ads, I take a subsample of 179 that have some recorded dwell times to run a kernel estimation of ad-specific dwell times with unsuccessful conversion. I restrict the sample to dwell times with failed conversion because if a conversion occurred during a review session, true dwell times would be either

Table 3.2: *Position, dwell time and ad fixed effects on conversion*

Variable	Coef.	Std. Err.	t	$P > \|t\|$
ML-2	0.019836	0.003323	5.97	0.000
ML-3	-0.1193	0.004612	-25.87	0.000
ML-4	-0.10454	0.004904	-21.32	0.000
SB-1	-0.10302	0.005059	-20.37	0.000
SB-2	-0.10249	0.00517	-19.82	0.000
SB-3	-0.10296	0.005205	-19.78	0.000
30-60	0.002948	0.004004	0.74	0.462
60-90	0.000728	0.004337	0.17	0.867
90-120	-0.0008	0.004411	-0.18	0.856
120-150	0.013577	0.004731	2.87	0.004
150-180	0.00967	0.005045	1.92	0.055
180-210	0.011034	0.005306	2.08	0.038
210-240	0.033486	0.00582	5.75	0.000
240-270	0.021701	0.006559	3.31	0.001
270-300	0.022854	0.006989	3.27	0.001
300-330	0.025148	0.008362	3.01	0.003
330-360	0.036638	0.008051	4.55	0.000
> 360	0.057653	0.003231	17.85	0.000
Missing	0.021868	0.002525	8.66	0.000
cons	0.062412	0.003759	16.6	0.000

missing or will be longer than the recorded value. Estimated distributions for two ads are shown in the appendix.

Note that the true distribution will be more concave: estimated distributions tend to overestimate the frequency of low-dwell times and underestimate the frequency of high-dwell times, due to selective subsampling of data. However, the following section reveals that such bias is irrelevant to the general shape (modality) of the estimated density of quality, as required by Lemma 3.2.1.

3.4 Model Estimation

Step 1: Estimate F_p

In this section I recover the distribution of a latent variable, $p_k(q_k)$ non-parametrically using the assumption on the ad review process in each period.[3]

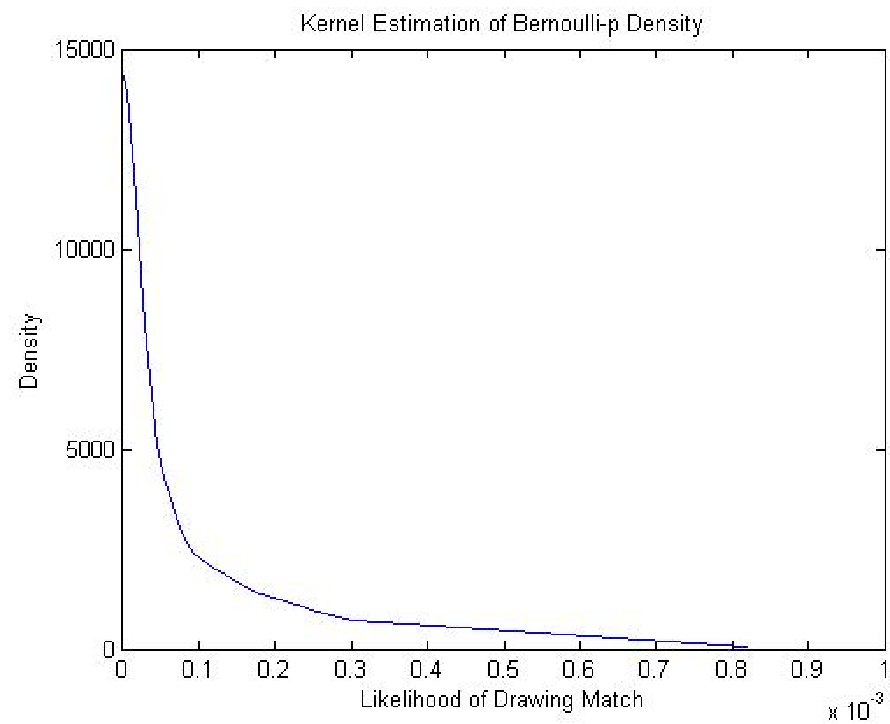
The following equality is true:

[3] An alternative estimation technique is described in the appendix.

$$
\begin{aligned}
q_k &= Prob(IsConverted_k = 1) \\
&= 1 - E_{t_k}[(1 - p_k(q_k))_k^t] \\
&= 1 - \sum_{t=0}^{T}(1 - p_k(q_k))^t Prob(t_k = t)
\end{aligned}
$$

If I let $x = 1 - p_k(q_k)$, I can generate a sequence of p_k that correspond to each observable q_k. A unique (q_k, p_k) with p_k in the unit interval is guaranteed to exist by continuity of the T-degree polynomial and Intermediate Value Theorem; its uniqueness can be shown by differentiating the expression with respect to x which is always positive in $[0, 1]$. A kernel estimation of F_p is shown in Figure 3.2.

Figure 3.2: *Estimated distribution of p*



As noted in the dynamics of conditional expected probability of match, I can confirm that the density actually fits very close to a fitted Beta distribution with $\alpha \leq 1$ and hence yields the required monotonicity properties.

Step 2: Estimate Signals, Θ of G

Unlike the procedure used to estimate F_p, parameters of the distribution of signals to consumers, Θ is difficult primarily due to lack of any direct information available to the econometrician. Therefore, I apply the structural model presented in lemma 3.2.1 on belief updating and proposition 3.2.11 to exploit terminal dwell times faced by consumers with prior belief of match, $\mu_k^{\tilde{t}_k}$ in a particular search session to derive moment conditions and use GMM for estimation.

At time $t = 0$, consumer's belief of match on viewing ad k is equal to $E[p_k|s_k]$. Then after sufficient number of periods has passed (in this case, $\lceil 1/E[p_k|s_k] \rceil$) producing consecutive failures, consumer's belief deteriorates to $\int_0^{E[p_k|s_k]} y dF_k(y|s_k)$. I can recursively define the value of each step in equation 3.1 as:

$$h(s_k, n) = \int_0^{h(s_k, n-1)} y dF_k(y|s_k) \quad \forall n > 0 \tag{3.12}$$

$$h(s_k, 0) = \int_0^1 y dF_k(y|s_k) \tag{3.13}$$

Then I have $\mu_k^t = h(s_k, n) \quad t \in [\lceil h(s_k, n-1)^{-1} \rceil, \lceil h(s_k, n)^{-1} \rceil) \quad \forall n \geq 1$.

Now consider a set of search sessions I_F (controlling for the set of unique ads shown) in which ad in position k was clicked first and subsequently ad in position k' was clicked second. Then by proposition 3.2.11, the dwell time $\tilde{t}_k$ on ad k is such that:

$$\mu_k^{\tilde{t}_k+1} < \mu_{k'}^{\tilde{t}_k} \leq \mu_k^{\tilde{t}_k} \tag{3.14}$$

Equivalently, dwell time will be mapped by some unique $\tilde{n}$ such that:

$$h(s_k, \tilde{n}+1) < h(s_{k'}, 0) \leq h(s_k, \tilde{n}) \tag{3.15}$$

Yielding, $\tilde{t_k} = \sup[\lceil h(s_k, n-1)^{-1} \rceil, \lceil h(s_k, n)^{-1} \rceil) = \lceil \frac{1}{h(s_k, \tilde{n})} \rceil$.

Then by using the observed dwell times, $t_{ik} \quad \forall i \in I_F$, I set:

$$\sum_{i \in I_F} \frac{1}{|I_F|} t_{ik} = E_{s_k, s'_k} \left[\frac{1}{h(s_k, \tilde{n}(s_k, s_{k'}))} |\phi_k \right] \quad \forall k \tag{3.16}$$

Note that given the computed parameters of quality distribution from the previous step, a simulation based on initial guess of Θ will allow one to use the moment for estimation.

Step 3: Estimate c and δ

After using dwell time for paramter estimation, I now turn to the value function as in equation 3.1 and the observed click-through rate of each ad k to derive a vector of CTR moments to compute search cost and the discounting factor.

In the dataset, I restrict my attention to search sessions in which ad k is clicked exactly once and it is the first click. Then I set:

$$Pr(v(\mu_k^0) = z(s, c, \delta) \geq c \cap s_k = \max_{j \in K} s_j) = CTR_k \quad \forall k \tag{3.17}$$

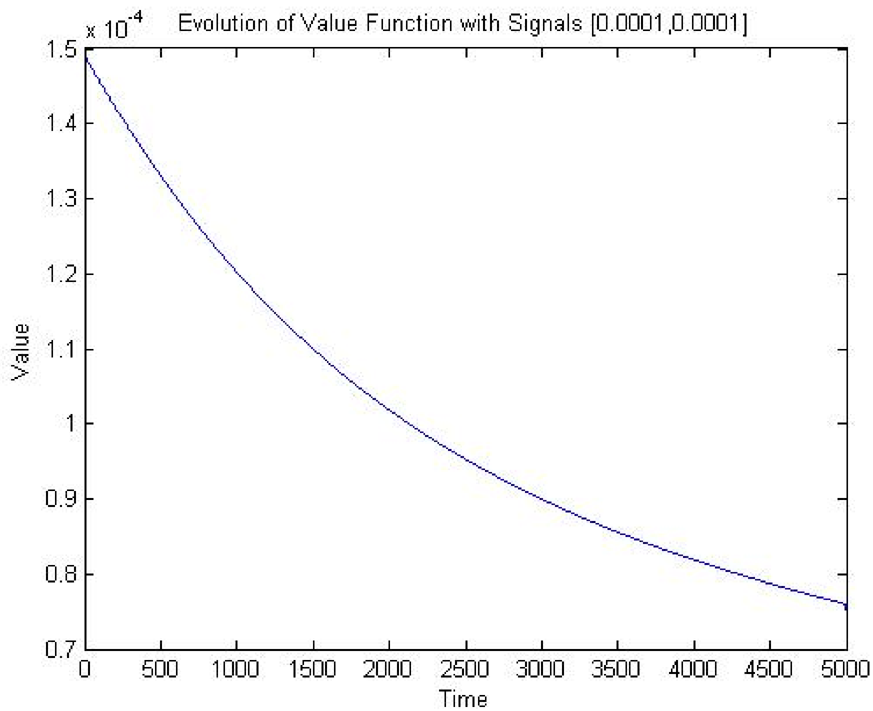
where $z(\cdot)$ is a function of vector of signals over each ad, search cost and the discounting factor. This function can be estimated using the parameters computed in the previous steps. CTR_k is the particular position-adjusted click-through rate computed using clicks in the chosen set of search sessions. Note that I have (over) identification of the parameters as long as $k > 1$.

Step 4: Computing the Value Function

After I have estimated a distribution of signals, I can draw a signal vector $\mathbf{s}$ for K ads from $\mathbf{G}$. Proposition 3.2.11 shows that there is some terminal duration period for a given vector of signals, hence the dynamic program is non-stationary and finite-horizon. I also know that if the terminal duration is T, because the reservation utility of each consumer is normalized to 0, $v(\mathbf{s}, H^T, T) = 0$. Attempting to solve for $v(\mathbf{s}, H^0, 0)$ using backward induction is yet an arduous process because for large T, the specifying every terminal node over a game of all possible histories, H^t in each $t < T$ is intractably large. In fact, given the estimated distribution of terminal dwell times in our data set, T is often approximately 1 hour, or 3600 seconds.

The proposed method of computing the value function is to first initialize a vector of value function from $t = 1$ to $t = T$ periods. Instead of using backward induction, I compute *forward* using the initial guess and iterate each time over a convergence criterion. Figure 3.3 shows that for a vector of signals $(1e-4, 1e-4)$, the value function is monotonically decreasing over time, following the convex shape of $g(1e-4, t)$ as expected. Figure 3.4 is a plot of initial values $(t = 1)$ over a plane of signal vectors. It reveals a Leontief cross-section reflecting the consumer's click-sequence that is driven by the maximum value over the vector of signals. What is important to note is that the value function is strictly monotonic in signal at time t. This means that a low-quality advertiser with a convincing sub-text (high signal) yields expectedly high value to rational consumers.This equilibrium is sutained by an accumulating search cost incurred

Figure 3.3: *Estimated path of signal over time*



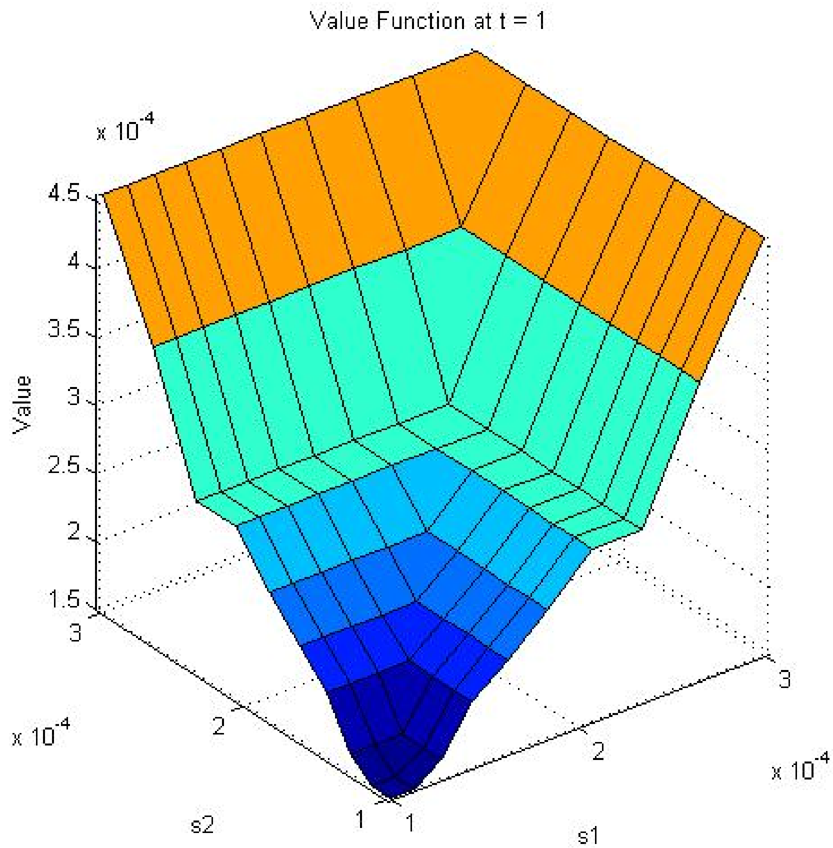
by consumers, preventing them from completely unraveling the underlying quality of ads. Additionally, the search engine assigns higher scores to ads with more clicks, places the low-quality ad at a higher position, further strengthening misguided signals. This results in a rational herding among consumers, or a tendency to concentrate on ads based on signals (beliefs). Summary statistics in Table 3.1 reflects this: clicks are distributed monotonically across positions, indicating consumer preference for signals; yet, the actual underlying quality of ads as proxied by dwell times are not distributed in any systematic order.

3.5 Conclusion

Using a platform framework, I developed and computed a dynamic model of search in internet advertising. Micro-level browsing data from Microsoft's Bing.com (formerly known as Live.com) was used for structural estimations. Given the limitations in data, I suppressed consumer heterogeneity by normalizing individual-level parameters that simplified the estimation process. This allowed for non-parametric primitive distribution estimations but lacked a robust micro-foundation.

The model predicts that users do not click on any ad with weak signals due to accumulating search

Figure 3.4: *Estimated value function over vector of signals*



cost and monotonicity of the value function. Rational search reveals a cascading pattern: the user clicks on a sufficiently high, highest-signal ad first, then moves on to the ad with the next highest conditionally expected probability of match once his assessment of likelihood of match on the current ad degrades over time. The user exits when maximum assessment of likelihood of match over all ads is below a threshold value.

Results imply that a low-quality advertiser with a convincing sub-text (high signal) yields expectedly high value to rational consumers.This equilibrium is sutained by an accumulating search cost incurred by consumers, preventing them from completely unraveling the underlying quality of ads. Additionally, the search engine assigns higher scores to ads with more clicks, places the low-quality ad at a higher position, further strengthening misguided signals. This results in a rational herding among consumers, or a tendency to concentrate on ads based on signals (beliefs).

References

[1] I. Ashlagi, D. Monderer, and M. Tennenholtz. Simulataneous Ad Auctions. Working paper. Harvard University, 2008.

[2] Susan Athey and Glenn Ellison. Position auction with consumer search. Working Paper, 2008.

[3] M. Schwarz B. Edelman. Strategic bidder behavior in sponsored search auctions. *Decision Support Systems*, 43(1):192–198, 2007.

[4] Ping W. Baltagi, B. Unequally spaced panel data regressions with ar(1) disturbances. *Econometric Theory*, 15(1):814–823, 1999.

[5] O. Bengtsson and D. Hsu. How do venture capital partners match with startup founders? Working paper. Wharton School, 2010.

[6] V. Bhagwat. Manager networks and investment syndication: Evidence from venture capital. Working paper. Northwestern University, 2011.

[7] Da Rin M. Bottazi, L and T. Hellman. What is the role of legal systems in financial intermediation? theory and evidence. *Journal of Financial Intermediation*, 18(1):559–598, 2009.

[8] M. Cary, A. Das, B. Edelman, I. Giotis, K. Heimerl, A. Karlin, C. Mathieu, , and M. Schwarz. Greedy Bidding Strategies for Keyword Auctions. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, 2007.

[9] E. Clarke. Multipart pricing of public goods. *Public Choice*, 18:19–33, 1971.

[10] Frazzini A. Cohen, L. and C. Malloy. The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy*, 116(5):951–979, 2008.

[11] Babur De Los Santos. Consumer search on the internet. NET Institute Working Paper 08-15, 2009.

[12] B. Edelman. Towards a Bill of Rights for Online Advertising. *Journal of Advertising Research*, 2009.

[13] B. Edelman. The Pathologies of Online Display Advertising Marketplaces. *ACM Sigecom Exchanges*, 2010.

[14] B. Edelman, M. Ostrovsky, and M. Schwarz. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97(1):242–259, 2007.

[15] D. Ewing. *Inside Harvard Business School: Strategies and Lessons of America's Leading School of Business*. Times Books, 1990.

[16] M. Mobius G. Ellison, D. Fudenberg. Competing Auctions. *Journal of the European Economic Association*, 2(1):30–66, 2004.

[17] Kovner A. Lerner J. Gompers, P. and D. Scharfstein. Performance persistence in entrepreneurship, journal of financial economics. *Journal of Financial Economics*, 62(1):731–764, 2007.

[18] Mukharlyamov V. Gompers, P. and Y. Xuan. Cost of friendship. Working paper. Harvard Business School, 2012.

[19] T. Groves. Incentives in teams. *Econometrica*, 41:617–631, 1973.

[20] J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1):153–161, 1979.

[21] Hegde and J. Tumlinson. Can birds of a feather fly together? evidence for the economic payoffs of ethnic homophily. Working paper, 2011.

[22] Ljungqvist A. Hochberg, Y. and L. Yang. Whom you know matters: Venture capital networks and investment performance. *Journal of Finance*, 62(1):251–301, 2007.

[23] H. Hotelling. Stability in competition. *The Collected Economics Articles of Harold Hotelling*, pages 50–63, 1990.

[24] D. Chiu J. Chiu. Mathematical modeling of competition in sponsored search market. *Proceedings of the 2010 Workshop on Economics of Networks, Systems and Computation*, 10, 2010.

[25] A. Sorensen K. Hendricks. Information and the skewness of music sales. *Journal of Political Economy*, 117(2):324–369, 2009.

[26] S. Kaplan and A. Schoar. Private equity performance: Returns, persistence, and capital flows. *Journal of Finance*, IX(4):1791–1823, 2005.

[27] Sensoy B. Stromberg P. Kaplan, S. Should investors bet on the jockey or the horse? evidence from the evolution of firms from early business plans to public companies. *Journal of Finance*, 64(1):75–115, 2009.

[28] J. Lerner and U. Malmendier. With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. Working paper. Harvard Business School, 2011.

[29] Richardson M. Ljungqvist, A. and D. Wolfenzon. The investment behavior of private equity fund managers. Working paper. New York University, 2005.

[30] R. B. Myerson. Optimal Auction Design. *Mathematics of Operations Research (6)*, 21:58–73, 1981.

[31] M. Pai. Competition in Mechanism. *ACM Sigecom Exchanges*, 9(7), 2010.

[32] Jeziorski Przemyslaw and Ilya Segal. What makes them click: Empirical analysis of consumer demand for internet search advertising. NET Institute Working Paper, 2009.

[33] L. Page S. Brin. The anatomy of a large-scale hypertextual Web search engine. *Computer networks and ISDN systems*, 1998.

[34] K. Shue. Executive networks and firm policies: Evidence from the random assignment of mba peers. Working paper. University of Chicago Booth School of Business, 2011.

[35] M. Sorensen. How smart is smart money? a two-sided matching model of venture capital. *Journal of Finance*, 62(6):2725–2762, 2007.

[36] T. Sunesson. Alma mater matters: the value of school ties in the venture capital industry. Working paper. Stockholm School of Economics, 2009.

[37] H. Varian. Position auctions. *International Journal of Industrial Organization*, 25:1163–1178, 2007.

[38] W. Vickrey. Counterspeculations, auctions, and competitive sealed tenders. *Journal of Finance*, 16:15–27, 1961.

Appendix A

Appendix to Chapter 1

A.1 Supplementary Tables and Figures

Figure A.1: *Share of HBS class involved in venture capital as financiers or entrepreneurs*

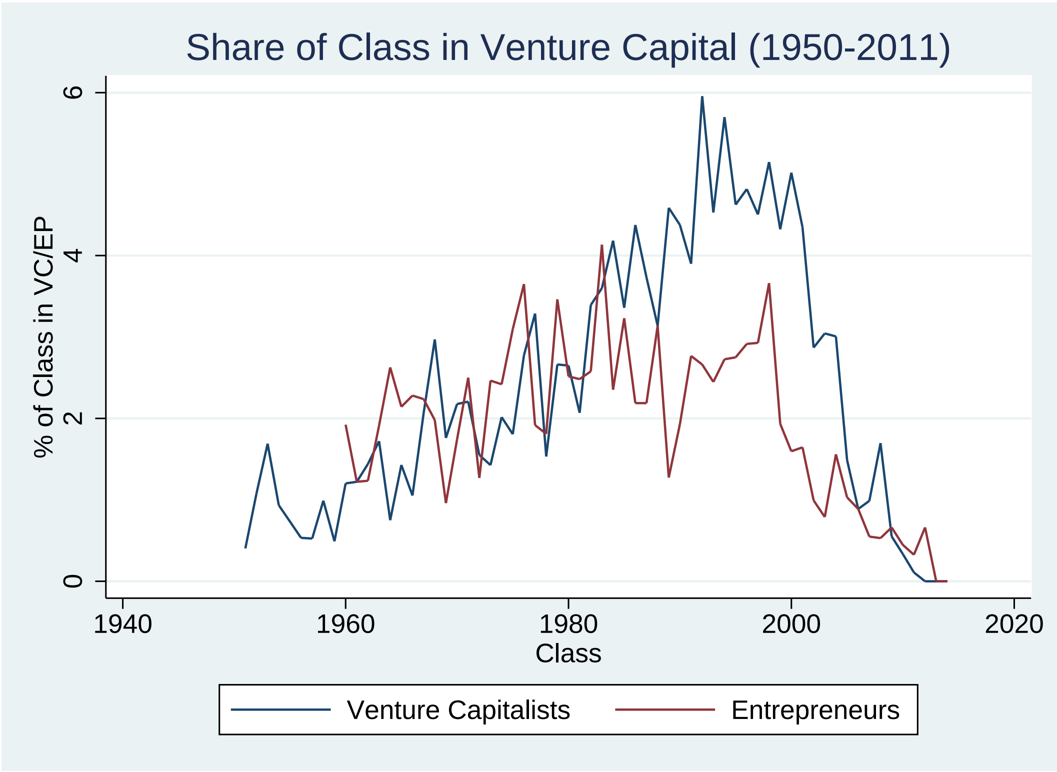


Figure A.2: *Range of shares of HBS section attending top colleges*

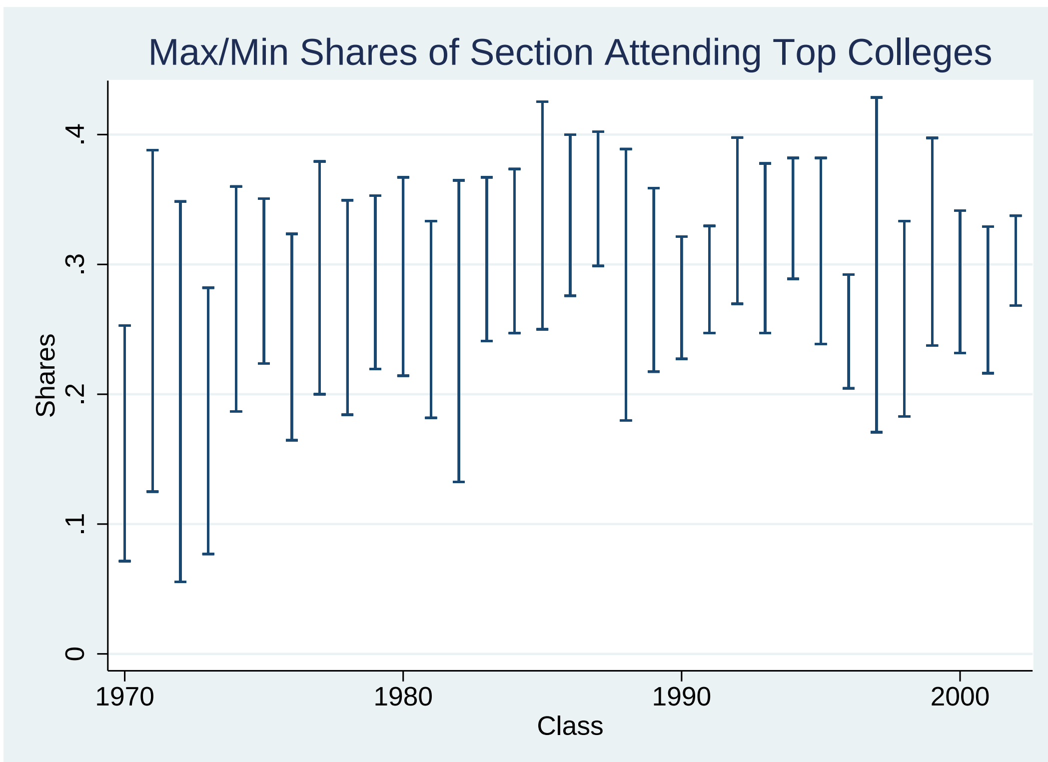


Figure A.3: *Share of investment rounds of portfolio companies that have gone public as of 2011*

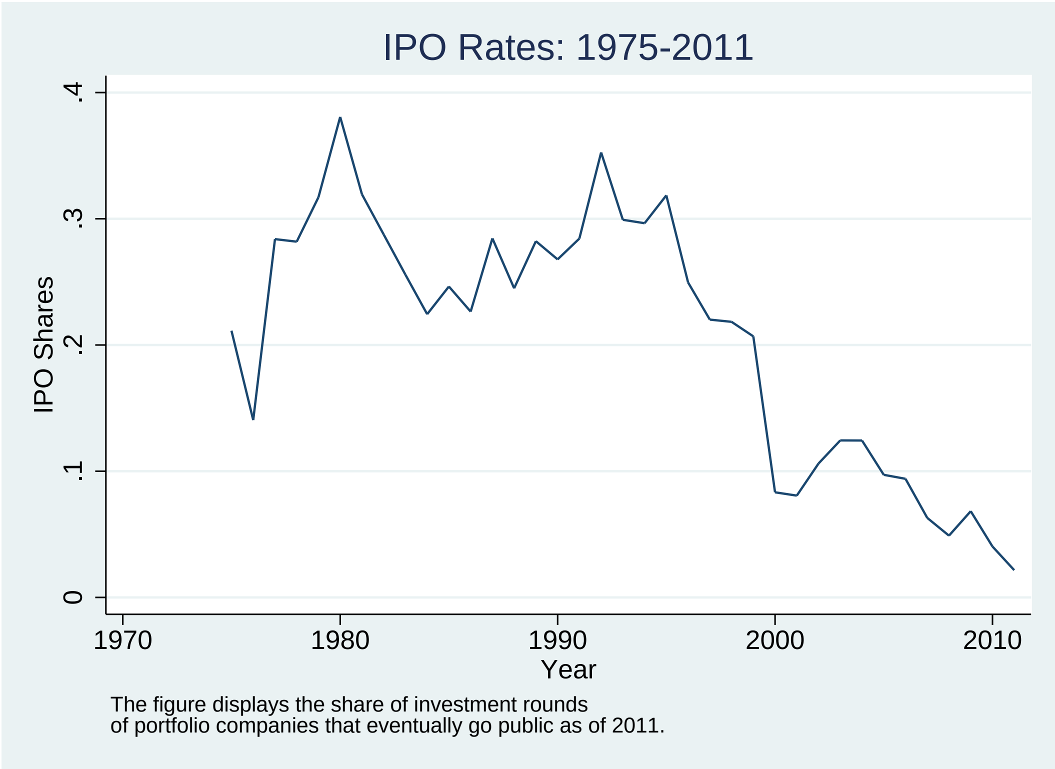


Figure A.4: *Total number of investment rounds, 1975 - 2011*

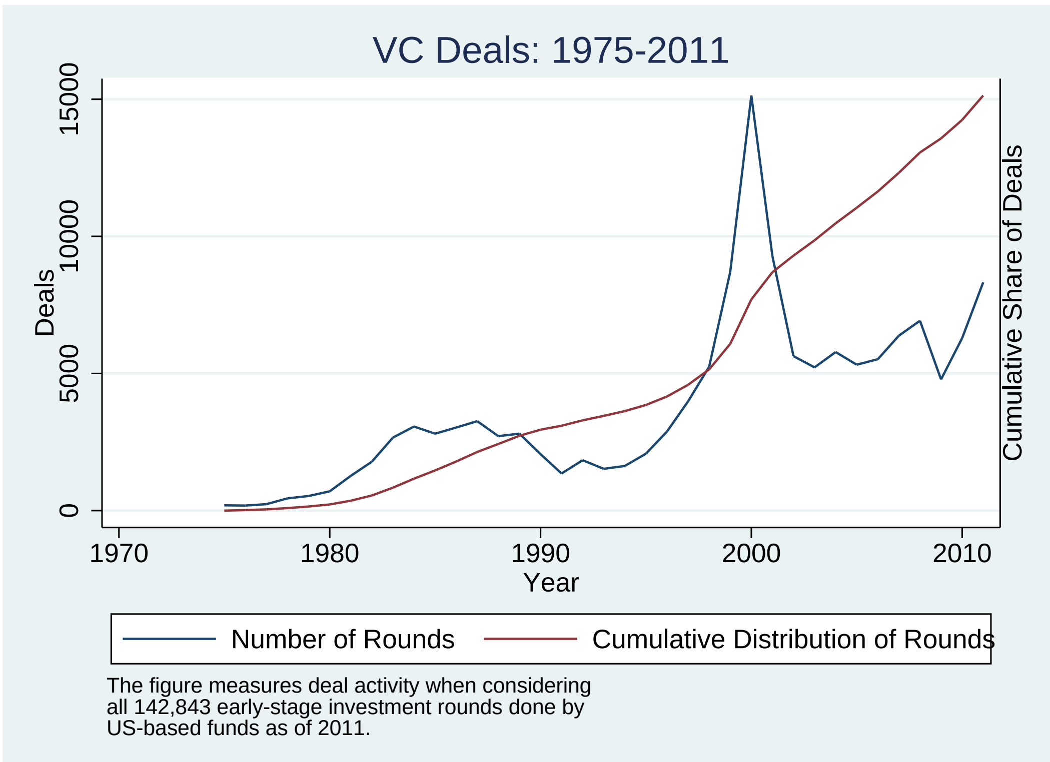


Figure A.5: *Correlation between IPO rates and percent of section attending top VC/EP-producing schools, 1970 - 2008*

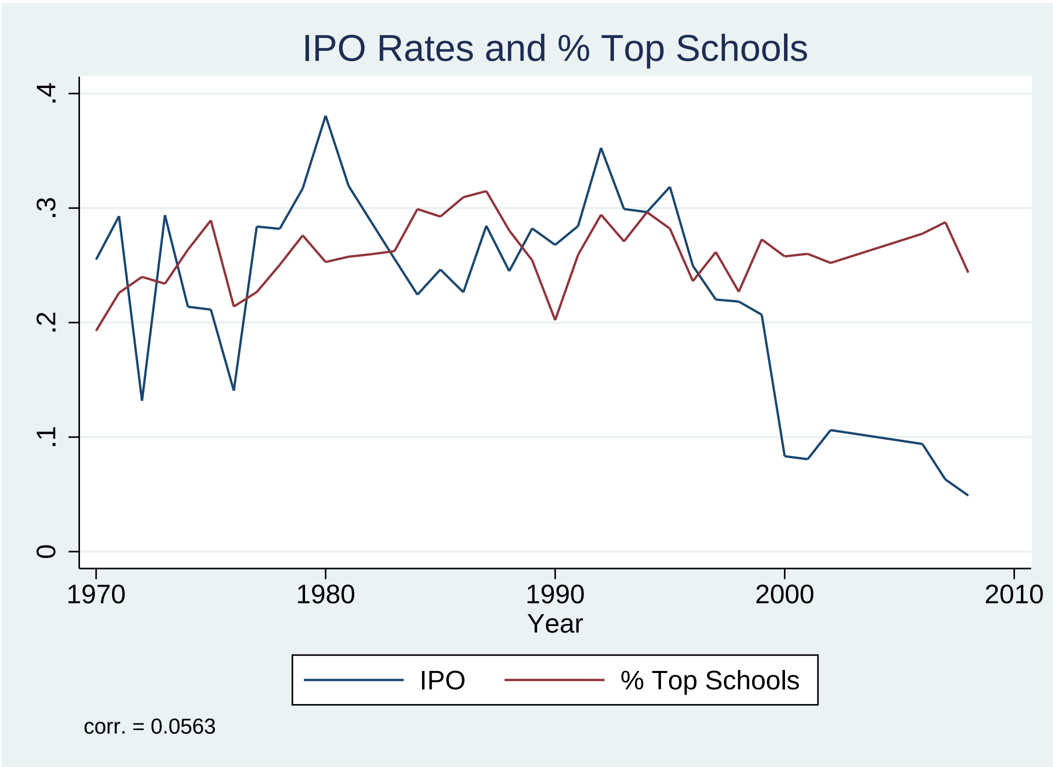


Table A.1: *Initial deal flow from peer entrepreneurs*

	(1) Degree	(2) Degree	(3) Indegree	(4) Indegree	(5) Outdegree	(6) Outdegree
% of section-peer EPs	29.41***		1.238		12.65**	
	(15.76)		(2.905)		(4.319)	
% of class-peer EPs		32.69		1.424		17.09**
		(23.90)		(4.762)		(6.979)
Degree_1	0.623*	0.623*				
	(0.0276)	(0.0276)				
Indegree_1			0.871*	0.871*		
			(0.0107)	(0.0107)		
Outdegree_1					0.886*	0.886*
					(0.00837)	(0.00835)
# of HBS execs	-1.123	-1.070	-0.235	-0.233	-0.325	-0.263
	(0.724)	(0.766)	(0.161)	(0.176)	(0.250)	(0.256)
Log of total equity invested	1.859*	1.863*	0.252*	0.252*	0.426*	0.427*
	(0.196)	(0.195)	(0.0611)	(0.0611)	(0.0357)	(0.0356)
Log of active firms	6.098*	6.098*	0.515*	0.515*	0.270*	0.271*
	(0.322)	(0.322)	(0.0665)	(0.0663)	(0.0578)	(0.0577)
IPO count	2.421*	2.420*	0.372*	0.372*	0.372*	0.372*
	(0.265)	(0.266)	(0.0451)	(0.0451)	(0.0573)	(0.0573)
MA count	1.889*	1.889*	0.332*	0.332*	0.299*	0.299*
	(0.135)	(0.135)	(0.0321)	(0.0321)	(0.0218)	(0.0218)
Clustering by Firm	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE?	Yes	Yes	Yes	Yes	Yes	Yes
Type FE?	Yes	Yes	Yes	Yes	Yes	Yes
Stage FE?	Yes	Yes	Yes	Yes	Yes	Yes
Class-year FE?	Yes	Yes	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes
N	24249	24249	21082	21082	21082	21082
R^2	0.845	0.845	0.944	0.944	0.919	0.919

Standard errors in parentheses

*** $p < 0.10$, ** $p < 0.05$, * $p < 0.001$

Note: The sample consists of a panel of first-time funds by 3,538 VC firms raised betwen 1980 and 2011 that are followed for 10 years or up to December 2011, whichever is earlier. I estimate fixed-effects OLS models with network measure lags to allow for persistence over time in a VC firm's network position. Percent of peer EPs measures the share of HBS executives' sections actively seeking investments. Degrees of VC j in year t measure the number of unique syndication partners of firm j in all investment rounds over a 5-year trailing window ending in year t. Indegrees of VC j in year t measure the number of unique syndication leaders that invited firm j over a 5-year trailing window ending in year t. Outdegrees of VC j in year t measure the number of unique syndication partners invited by firm j over a 5-year trailing window ending in year t. Active firm in year t is a VC firm involved in at least one equity investment round of a portfolio company in year t. All controls, with the exception of year dummies, are lagged by one year. Intercepts are not shown. Errors are clustered by firm/fund.

Appendix B

Appendix to Chapter 2

B.1 Example with Exponential Click-through Rates

Suppose click-through rates $\alpha_{kl} = \beta_l^k$ for each $k \leq K$ where $\beta_l < 1$. Note that these exponential click-through rates obey both monotonic ordering of α_{kl} and the *diminishing differences* property. Following the framework in Theorem 2.5.3, we set $\beta_{\tilde{l}} = \beta_{l'} > \beta_l$. Condition (2.19) then becomes:

$$C_{\tilde{l}} \geq \frac{C_{l'}\beta_{l'}^k + C_l\beta_l^k}{(N-k)\beta_{l'}^{k+1} - (N-k-1)\beta_{l'}^k} \quad \forall k \leq K.$$

Simplifying and rearranging:

$$C_{\tilde{l}} - C_{l'} - C_l\frac{\beta_l^k}{\beta_{l'}^k} \geq C_{\tilde{l}}(1-\beta_{l'})(N-k). \tag{B.1}$$

Because $\beta_l < \beta_{l'}$, we know:

$$C_{\tilde{l}} - C_{l'} - C_l\frac{\beta_l}{\beta_{l'}} \leq C_{\tilde{l}} - C_{l'} - C_l\frac{\beta_l^k}{\beta_{l'}^k}$$

and

$$C_{\tilde{l}}(1-\beta_{l'})(N-1) \geq C_{\tilde{l}}(1-\beta_{l'})(N-k).$$

Thus the generalized restrictions in equation (B.1) become:

$$C_{\tilde{l}} - C_{l'} - C_l\frac{\beta_l}{\beta_{l'}} \geq C_{\tilde{l}}(1-\beta_{l'})(N-1) \tag{B.2}$$

because this condition implies the required restrictions for each $k \leq K$.

By the subadditive property of joined auction $\tilde{l}$ and equation (B.2):

$$C_{l'} + C_l \geq C_{\tilde{l}} \geq C_{\tilde{l}}(1-\beta_{l'})(N-1) + C_{l'} + C_l \frac{\beta_l}{\beta_{l'}}$$

which requires:

$$C_{l'} + C_l \geq C_{\tilde{l}}(1-\beta_{l'})(N-1) + C_{l'} + C_l \frac{\beta_l}{\beta_{l'}}.$$

Rearranging:

$$\frac{C_l}{C_{\tilde{l}}} \geq (N-1)\frac{1-\beta_{l'}}{1-\frac{\beta_l}{\beta_{l'}}} \tag{B.3}$$

Since $C_l \leq C_{\tilde{l}}$ and $N \geq 2$ (so that the RHS of (B.3) is not trivially 0), inequality (B.3) requires:

$$1 - \frac{\beta_l}{\beta_{l'}} > 1 - \beta_{l'}.$$

Rearranging yields the requirement:

$$\beta_{l'} > \sqrt{\beta_l}. \tag{B.4}$$

Condition (B.4) requires that the auction with superior technology have $\beta_{l'}$ sufficiently larger than β_l. Furthermore, if N is very large, then $\beta_{l'}$ must approach 1 in order to satisfy the requirement in (B.3). Note that these are necessary conditions, but not sufficient conditions.

Appendix C

Appendix to Chapter 3

C.1 Alternative Estimation of F

F_k is the distribution of the true, underlying likelihood of match occurring between the consumer and the product offered via ad k. I use the fraction of clicks that resulted in $IsConverted_k = 1$ to proxy for the likelihood of match. Consider the following latent variable model:

$$Y_k^* = X_k' \beta_k + \epsilon_k \tag{C.1}$$

where $\epsilon_k \sim N(0,1)$, X_k is a vector of appropriate regressors (i.e. position dummies) and $IsConverted_k$ is the indicator variable for $Y_k^* > 0$. Then given a sequence of observations, $\{y_{ik}, x_{ik}\}_{i=1}^N$, I estimate the parameters using maximum-likelihood:

$$\hat{\beta}_k = \arg\max_{\beta_k} \sum_{i=1}^{N} \left[y_{ik} ln\Phi(x_{ik}'\beta_k) + (1 - y_{ik}) ln(1 - \Phi(x_{ik}'\beta_k)) \right] \quad \forall k \tag{C.2}$$

But since I have a large number of impressions with small variation in regressors, we can also compare estimated parameters under MLE to Berkson's minimum chi-square method. Suppose among N observations $\{y_{ik}, x_{ik}\}_{i=1}^N$ there are only $M << N$ distinct vectors of the regressors. Group all x_{ik} into a distinct set $\{x_{1k}, ..., x_{Mk}\}$. Let r_m be the number of observations with $x_{ik} = x_{mk}$ and $y_{ik} = 1$ and let n_m be the total number of observations with $x_{ik} = x_{mk}$. I will denote:

$$\hat{p}_m = r_m / n_m \tag{C.3}$$

$$\hat{\sigma}_m^2 = \frac{1}{n_m} \frac{\hat{p}_m(1-\hat{p}_m)}{\phi^2(\Phi^{-1}(\hat{p}_m))} \tag{C.4}$$

Then the generalized least squares estimator with weights $\hat{\sigma}_m^{-2}$ regressing $\Phi^{-1}(\hat{p}_m)$ on x_{mk} is:

$$\hat{\beta}_k = \left(\sum_{m=1}^{M} \hat{\sigma}_m^{-2} x_{mk} x'_{mk} \right)^{-1} \sum_{m=1}^{M} \hat{\sigma}_m^{-2} x_{mk} \Phi^{-1}(\hat{p}_m) \tag{C.5}$$

C.2 Supplementary Tables and Figures

Figure C.1: *Estimated dwell time distribution of ad (1)*

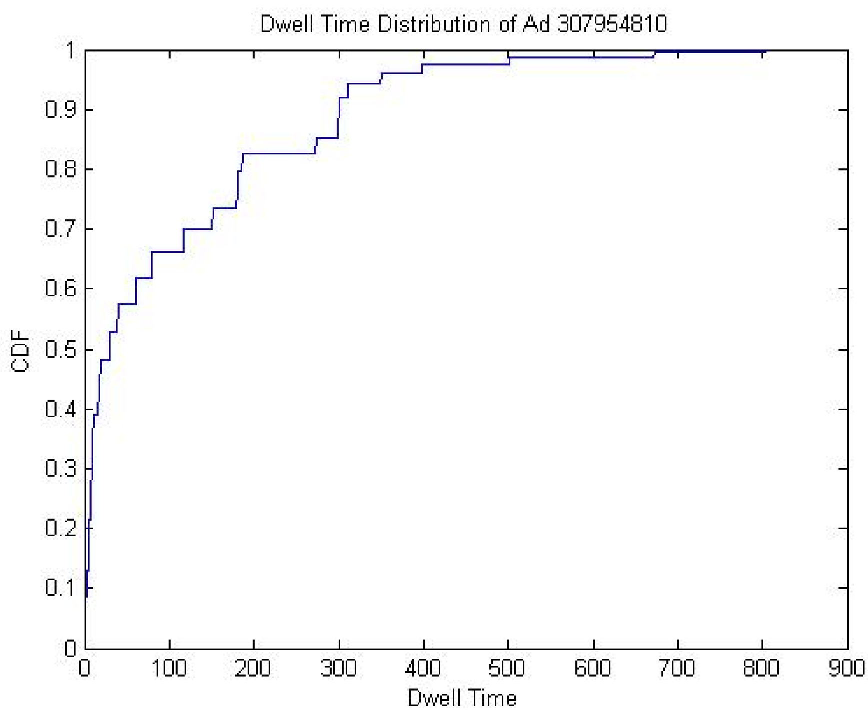


Figure C.2: *Estimated dwell time distribution of ad (2)*