Essays on the Production of Innovation

by

Ankur Chavda

B.S.E., University of Pennsylvania (1999)
B.S.E., University of Pennsylvania (1999)
M.S., Free University of Brussels (2003)
S.M., Massachusetts Institute of Technology (2016)

SUBMITTED TO THE SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN MANAGEMENT

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2019

Signature redacted

Signature of Author: __

Department of Management
May 3, 2019

Signature redacted

Certified by: __

Scott Stern
David Sarnoff Professor of Management
Thesis Supervisor

Signature redacted

Accepted by: __

Ezra Zuckerman Sivan
Alvin J. Siteman (1948) Professor of Entrepreneurship and Strategy
Deputy Dean
Faculty Chair, MIT Sloan PhD Program

Essays on the Production of Innovation

by

Ankur Chavda

Submitted to the Sloan School of Management
on May 3, 2019, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Management

ABSTRACT

Innovation is central to both the competitive strategy of many firms and gains in productivity that lead to economic growth. How do we improve our ability to produce innovations? This dissertation studies three aspects of this question: staged development, vertical integration and venture capital.

Staged development, common in many innovative settings including biotech and venture capital of ideas, involves partially funded an idea with the goal of learning more about that idea before additional funding is provided. My result suggests there are cases where committing to ideas, avoiding staged development, can lead to better outcomes. Staged development has the potential of distorting effort to an extent that outweighs any benefit provided by its implicit option value.

Research units pursuing innovations can either be integrated within the firm exploiting those innovations or kept as a separate entity. I find that integration leads to a higher rate of new innovations. Separating the research unit can reduce its appetite for risk, changing both the rate and direction of innovation.

Finally, uncertainty surrounds strategies to exploit innovations: new ideas by definition have not been tested by market forces. I show how venture capital plays a key role in resolving this uncertainty for entrepreneurs with new ideas. Specifically, venture capital provides the most value for entrepreneurs that are themselves the most uncertain about the underlying value of their ideas.

Thesis Supervisor: Scott Stern

Title: David Sarnoff Professor of Management *dedicated to Saar, for letting me be me*

ACKNOWLEDGMENTS

My committee was instrumental to the work in this thesis. Scott Stern has the unique ability to see the best version of new ideas; rather than focus on the incompleteness of my early drafts, Scott continuously encouraged me to pursue the underlying insights. This encouragement allowed me to execute a thesis of unconventional ideas as he gave me the leeway to explore questions that were initially hard to formulate and explain. Pierre Azoulay provided a different kind of patience, the patience of someone who knows where a student will end up but is willing to excuse the fumbles along the way. Pierre coached me on how to make a theoretical argument, grounded in empirical fact. I met Shane Greenstein early on in my PhD course work and was drawn to his enthusiasm for studying technological innovation. He mentored me on how to transition from industry and helped me translate the practitioner questions I had in mind for consumption by an academic audience. I will be forever indebted to each of my committee members.

I also greatly benefited from the community at and around MIT. Alessandro Bonatti, Mercedes Delgado, Bob Gibbons, Olenka Kacperczyk, Danielle Li, Hong Luo, Matt Marx and Ezra Zuckerman were all exceedingly generous with their time and advice. Michael Cusumano, Glenn Ellison, Joshua Gans, Parag Pathak and Neil Thompson were kind enough to train me as a research assistant, despite my relative inexperience. I could not have succeeded without the support of the TIES PhD family, both from Dan Fehder, Jorge Guzman, Joshua Krieger and Abhishek Nagaraj whom graduated before me and my own cohort, Danny Kim, Hyejun Kim and most especially Sam Zyontz.

There were also a group of people that led me to pursue a PhD in the first place. Sampath Kannan and Marc Jegers saw potential in me to contribute as a researcher. Darius Sankey and Simon Geoghegan prompted my transition from industry by encouraging me to explore alternatives.

Finally, none of this would have been possible without the sacrifices of my family. My children were relocated away from their friends. My spouse restarted her career to allow me to reinvent myself. Thank you.

Contents

1. Studying Innovation Through Creative Media

Innovation is easy to theorize about but difficult to study; empirical work on innovation suffers from a statistical challenge specific to innovation contexts. Asymptotic estimators rely on observations being drawn from the same distribution or at least distributions with similar properties. However, innovations are by definition novel and therefore a dataset of innovations may not be easily interpreted as observations from similar distributions. MacCormack, Verganti and Iansiti (2001) for example look at the importance of flexibility in new product innovations but have observations as different as a web browser and a search portal. Often this challenge leads researchers to narrow the scope of innovations, such as Trajtenberg's (1990) paper on CT scan patents. But this still lumps dissimilar objects: patents exist for both processing the information generated by the scan and the physics of generating that information using X-rays. It's not obvious both these types of innovations are drawn from the same distribution of potential innovations. Some studies do constrain the type of innovation enough to meet this challenge, as in Kacperczyk's (2012) paper on mutual funds, but are limited in the insights they can generate since their constraints limit the degree of innovation observed.

Creative media provide an opportunity to surmount this challenge in studying innovation. First, creative media often takes on a highly constrained, regular form. Music albums for example are typically around 45 minutes, classified into genres and are priced similarly. This structure naturally supports the argument that observations of media can be considered draws from a shared distribution. Secondly, creative media contain the trappings of innovation. Arrow (1962) describes outcome uncertainty as a key defining characteristic of innovations. Creative media also suffers from such uncertainty; 65% of television shows get cancelled after their first season (Ocasio, 2012), a clear indication of failure when it takes four seasons for a show's financiers to make a return on their investment (Bunn, 2002). Scotchmer's (1991) sense of innovation as cumulative is also present. The first every sci-fi family drama, 1963's *My Favorite Martian*, differed from anything before it and was a precursor to later shows like *Mork & Mindy*, *ALF* and *3rd Rock from the Sun*. *The Sopranos* mainstreamed drama shows with season long story arcs, leading to later shows like *Breaking Bad* and *Game of Thrones*.

Beyond this specific innovation challenge, creative media is fertile ground for general econometric research. Production functions are shared across incumbent firms, enabling firm level fixed effects. Outcome measures such as reviews and ratings exist which are independent of industry structure, allowing for the identification of the value of an innovation separate from the complementary assets necessary to exploit that innovation. Trade unions force detailed tracking of the roles and responsibilities of workers in creative media over their entire careers. Numerous technological and regulatory shocks exist in media industries which can potentially generate exogenous variations in the drivers of innovation. Yet outside papers such as Luo (2014), little work has been done exploiting creative media for empirical results on innovation. This dissertation aims to exploit this gap using television show development as a setting.

The first chapter, "The Downside of Experimentation: Evidence from Television Shows", studies the staged development of ideas, common in many innovative settings including biotech and venture capital. This form of experimentation has the obvious benefit of enabling many more ideas to be explored on a limited budget, selecting only the most promising for full investment. The chapter shows however that in some cases this can lead to worse average outcomes, stemming from complications due to the lack of commitment. Television shows provide a setting where experimentation has a very specific structure: the first episode of a television show is often produced before funding the rest of the first season. The fact that this structure is consistent across firms and time enables an econometric study of experimentation that would be difficult to replicate in another setting.

The second chapter, "Vertical Integration and the Direction of Innovation: Evidence from 1970's Television", considers how innovation produced by firms can be affected by their decision to vertically integrate. The chapter finds that when vertical integration is exogenously reduced, type of innovations pursued by firms shifts. In addition, the rate of new innovations brought to market drops. Television show production again provides an excellent setting for this research. The industry has an upstream creator of television shows, akin to a research unit, and a downstream broadcaster of those shows. Variation in whether the downstream firm has an ownership stake in the television show enables this study of vertical integration.

The third chapter, "Venture Capital Rents from Entrepreneurial Search", takes a different approach from the previous chapters. This chapter focuses on the heterogeneity in actions made by entrepreneurs facing similar opportunities. In contrast to a model where entrepreneurs simply hold different beliefs about the potential of those opportunities, adding uncertainty enables entrepreneurs and venture capitalists to bargain over resolving that uncertainty. Specifically, when entrepreneurs search for strategies to implement their ideas, they are willing to pay venture capitalists whom are better informed about when to stop or continue searching for new strategies. This type of theory would be difficult to test in a traditional entrepreneurial setting: a dataset would need observations of entrepreneurs with similar ideas but different beliefs about the potential of those ideas. But it is an example of the type of research that could be tested using the market for ideas in creative media as a proxy for traditional entrepreneurship.

1.1 References

Arrow, K. 1962. "Economic Welfare and the Allocation of Resources for Invention." in *The Rate and Direction of Inventive Activity: Economic and Social Factors,* pp. 609-625. Princeton, NJ: Princeton University Press.

Bunn, A., 2002. The Wasteland. *The New York Times Magazine.* June 23rd.

Kacperczyk, A., 2012. "Opportunity Structures in Established Firms: Entrepreneurship versus Intrapreneurship in Mutual Funds." *Administrative Science Quarterly,* 57:484-521

Luo, H., 2014. "When to sell your idea: Theory and evidence from the movie industry." *Management Science*, 60(12), pp.3067-3086.

MacCormack, A., Verganti, R. and Iansiti, M., 2001. "Developing products on 'Internet time': The anatomy of a flexible development process." *Management Science,* 47(1), pp.133-150.

Ocasio, A., 2012. "TV Success Rate: 65% Of New Shows Will Be Canceled (& Why It Matters)". *Screen Rant.* May 17th.

Scotchmer, S. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law." *Journal of Economic Perspectives*, 5(1): 29-41

Trajtenberg, M. 1990. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *RAND Journal of Economics*, 21(1): 172-187.

2. The Downside of Experimentation: Evidence from Television Shows

Ankur Chavda
April 2019

Abstract

Innovative firms often try out new ideas before fully investing in them as a kind of experimentation on those ideas. This experimentation generates an early signal of final outcomes, allowing potentially bad ideas to be terminated before those outcomes are realized. But not committing to ideas by retaining right to terminate can also be detrimental to outcomes, by for example attracting lower quality workers or shifting worker effort away from final outcomes towards passing the experimentation phase. In this paper I explore this tension, asking when does experimentation improve final outcomes. I test a theoretical model of experimentation against a dataset of television shows that both enables an estimate of the treatment effect of experimentation and allows a test for selection bias. I find evidence that experimentation may both handicap worker recruitment and adversely shift effort. This results in experimentation only improving final outcomes when it terminates enough bad ideas, otherwise experimentation is detrimental as its benefits are unable to overcome its downside: the lack of commitment.

2.1 Introduction

Biotech and venture capital firms are well known to invest in stages. Many ideas are funded for an early stage of development and fewer for later stages with the early stage information learned about an idea influencing whether later stages are funded. This kind of experimentation is a common characteristic of many innovative industries where final outcomes are uncertain. Such experimentation can winnow out potentially poor outcomes before they are realized, improving observed final outcomes (Bowman and

Hurry, 1993). Yet experimentation involves a lack of commitment (Ghemawat, 1991), suggesting the existence of downsides to experimentation: experimentation might make it harder to recruit high quality talent that prefers long term contracts (Holmstrom, 1983) or shift worker incentives towards improving the experimental signal to the determinate of final outcomes (Baker, 1992) among other potential downsides. A tension therefore exists between the benefits and drawbacks to experimentation with a natural research question: when does experimentation improve final outcomes and in the cases experimentation does not, why do final outcomes worsen?

I explore this research question in the setting of US television show production where television networks conduct experimentation on new shows through the staged financing of their development. Although the networks can make a "straight to series" order by directly funding a new show's entire first season of episodes, historically only the "pilot", the first episode, is ordered with the option of developing the rest of the season after evaluating the pilot. My main outcome variable is the quality of show episodes as rated by users of the Internet Movie Database, IMDb.

My empirical approach has four components. First, incumbent network became uncertain about whether piloting was optimal after Netflix's 2013 successful entry into television show production using an unconventional straight to series development process. This led to otherwise similar shows taking different production paths due to diversity in executive decision maker beliefs about the usefulness of piloting between incumbent networks. I use this uncertainty to craft a matching estimator to measure a treatment effect of piloting for new shows on incumbent networks. Second, Netflix's entry effectively cuts my dataset of incumbent shows into regimes, one where pilots was almost always used and another where a mixture of pilots and straight to series orders were used. This enables a difference in difference estimator I use to test for the presence of selection bias in my matching estimate based on predictions of my theoretical model of experimentation. Third, the likelihood a new show passes the pilot stage of development varies greatly with the previous track record of the creators of that new show. I use this variation to test heterogeneity in the overall treatment effect of piloting on final outcomes. Finally, IMDb users rate shows on an episode level, allowing me to observe the difference in outcomes between piloted and straight to series shows for the first episode relative to the rest of the first season. I exploit patterns in these outcomes for evidence on which commitment mechanisms could driving my matching estimates.

This paper's main contributions to the broader literature is to articulate, model and provide evidence for the tension between experimentation and commitment. My results suggest experimentation is worthwhile only when a high enough share of ideas is terminated after experimentation. Otherwise the lack of commitment to ideas causes average final outcomes to worsen. I find evidence this worsening is driven by worker incentives shifting towards passing the experimentation phase over final outcomes and an increased difficulty in recruiting high quality workers.

The rest of this paper proceeds as follows. Section 2 provides a review of previous literature related to my research question. Section 3 describes experimentation in television show production as well as how Netflix changed the incumbent's decision to experiment. Section 4 models the network's experimentation decision to generate testable hypotheses. Section 5 outlines the data used for this paper and Section 6 lays out my empirical strategy. Section 7 present this paper's results and Section 8 concludes with the broader implications of those results.

2.2 Literature Review

2.2.1 Benefits of Experimentation

The type of experimentation studied in this paper is a combination of the concepts presented in Bowman and Hurry (1993) and Thomke (1997). Bowman and Hurry's work details how firms can improve performance through incremental investments in resources under environmental uncertainty. By retaining the option to provide follow-on investment, firms can take advantage of environmental change favorable to a resource while restricting the losses on resources that are a poor fit for future environmental states. Although Thomke (1997) focuses on the learning possible from incremental investment rather than its option value, the environment is considered stable - the uncertainty lies in the potential of an investable resource, in his case new products under consideration by the firm. Taking these contributions together yields the proposition that staged investment, the kind of experimentation studied in this paper, can improve final observed outcomes poor potential outcomes are terminated before those outcomes are realized.

Variations of this type of experimentation process have appeared in other strains of literature. Models of sequential search for innovation as in Bhattacharya, Chatterjee and Samuelson (1986) usually have an unlimited number of investment stages with the innovation able to become finalized at any stage. Concerned with firms successfully transitioning in periods of technological change, Iansiti (2000) espouses the effectiveness of experimentation when conducting search activity in uncertain environments. Kerr, Nanda and Rhodes-Kropf (2014) apply the experimentation concept to describe the interplay between venture capital financing and entrepreneurship.

2.2.2 Drawbacks of Experimentation

In contrast to these papers that articulate how firms can benefit from experimentation, other research highlights how retaining the right to terminate a project can be detrimental to final outcomes. First, altering competitive behavior (Schmalensee, 1978; Ghemawat, 1993) is perhaps the canonical example of how experimentation can harm outcomes. Forgoing experimentation by committing to a project and relinquishing the right to terminate it could benefit a firm by inducing an industry equilibirum with fewer competitors. Second, experimentation could also affect outcomes due to misaligned incentives (Baker, 1992). Under experimentation, workers may be able to direct effort towards improving the probability that their project receives continued funding. If this effort is incongruent with final outcomes, the resulting shift in effort away from final outcomes could be detrimental to those outcomes. The innovation literature makes a similar argument that commitment despite early failures can encourage risker worker actions (Manso, 2011), leading to better outcomes for the firm (Lerner and Wulf, 2007; Azoulay, Graff Zivin and Manso, 2011). Third, worker aversion towards risk results in a preference for longer term contracts (Holmstrom, 1983). If in equilibirum the shorter contracts necessitated by experimentation do not compensate for this preference (Bonhomme and Jolivet, 2009), higher quality workers may select into firms that lack experimentation over ones that do. Finally, upfront planning on a project has been shown to correlate with better final favorable outcomes (Delmar and Shane, 2003). Under experimentation, the optimal level of planning for the entire project may not be justified (Milgrom and Roberts, 1990), leading to worsened outcomes.

2.2.3 Empirical Research on Experimentation

Prior empirical work on experimentation has focused on understanding an assumed positive effect on outcomes, rather than exploring the tradeoff from its use. Thomke (1998) and MacCormack, Verganti and Iansiti (2001), "MVI", are the closest papers to this one. Thomke shows positive correlation between the use of experimentation and existing of technology investments complementary to experimentation. MVI show positive correlation between low functionality prototypes and outcomes. Although MVI do not explicitly call the use of these prototypes as experimentation, prototypes are a form of staged development and therefore akin to the type of experimentation studied in this paper. This paper empirically differs from Thomke and MVI in several of ways.

First, by having enough comparable observations for statistical analysis and a standardized indicator of experimentation associated with each observation, this paper is the first to allow for a focused analysis of experimentation. MVI relies a small dataset of products as dissimilar as a web browser (Netscape 3.0) and an internet start page (My Yahoo!). Thomke's control and treatment groups differ in whether their technology makes experimentation easier to execute, not in whether experimentation is actually used. MVI experimentation indicator, how much of a product's final functionality was included in its initial prototype, could be inconsistently measured across products and could also be endogenous: the amount of functionality in an initial prototype could affect the initial feedback received and thereby the amount of additional functionality added after the prototype.

Second, this paper is the first to address selection bias: any observed difference in outcomes between experimentation and non-experimentation could be driven by an ex-ante selection of what is experimentation on. Thomke's observes projects only after an experimentation technology has been assigned and therefore has no way to control for the selection of projects into each technology. MVI does not have a source of randomization to overcome selection bias on the share of functionally included in an initial protype.

Third, my dataset includes variation in the value of experimentation, enabling me to uncover cases when experimentation could be detrimental. MVI is missing this type of variation altogether while Thomke lacks variation in the value of experimentation within observations that used experimentation.

Finally, unlike prior research on experimentation I provide evidence supporting the existence of specific mechanisms causing variation in the value of experimentation. Thomke mentions both error detection and design improvement as potential benefits of experimentation but cannot empirically distinguish between their effects on outcomes. MVI argues experimentation leads to the incorporation of early market feedback, improving the final product, yet lacks evidence that market feedback from prototypes affected final products in their setting.

2.3 Experimentation in Television

2.3.1 Staged Development in Television

As in many creative and innovative industries, uncertainty is ever present in the production of television shows. William Goldman's famous "Nobody knows anything" quote from the movie industry applies: 65% of television shows on the major networks get cancelled after their first season (Ocasio, 2012), a clear indication of failure when it takes four seasons for a show's financiers to make a return on their investment (Bunn, 2002). This failure rate exists despite a winnowing process that results in only a fraction of ideas for new shows getting aired on television.

Although this paper focuses on pilot stage of development as experimentation, television show production involves multiple development stages that are each a form of experimentation. Ideas for shows can come from a variety of places; a writer might come up with a plot idea for a new show or a network executive might wonder how a book would translate into a television series. The first stage of development for these ideas is drafting a log-line or synopsis (treatment) of the idea that captures the main setting and how the story might evolve over the course of a series. Treatments are pitched to network executives. However, at around five pages long, treatments are limited in do not provide a lot of information about whether the show will be successful (Luo, 2014).

Each year, a typical major network picks around a hundred treatments for the next stage of production, the creation of a script and a "bible", a full description of all aspects of the show's universe. Scripts typically cost on the order of $100k and provide the network with the option to fund continued development on the show idea. The script provides more information about the show idea than the treatment itself; the script order can be thought of as the first experiment done by the networks on the show idea.

Conditional on an attractive script, the networks can choose to engage in the third stage: the pilot. For each expected open slot for a new television show on their fall schedules, networks pick two to three scripts from its pool of optioned scripts to pilot. Piloting produces the first episode of the series at a cost of about two million dollars, an order or magnitude more expensive than the script. Although now somewhat controversial given Netflix's success with straight to series orders, piloting has a long history in the television industry; in 1951 CBS had Lucille Ball make a pilot for the *I Love Lucy* show even though she had played leading roles in both film and television.

The great majority of television produced today by networks other than Netflix and Starz continues to use pilots as experiments; Starz having adopted a straight to series production model long before Netflix entered the industry.

2.3.2 Netflix Changing Experimentation

Previously known as an online streaming distributor of film and television content, Netflix began funding new television production in 2011. Many have tried and failed to break into the exclusive club of prestige television show producers so by itself Netflix's funding of original constitute did not constitute a major event in the industry. For example, in 2001 A&E networks commissioned two scripted dramas which were cancelled by the end of 2002 because, in the words of one of the involved actors, "A&E was transforming from the premier intellectual cable network in America to one that airs Dog the Bounty Hunter on repeat (Farquharson, 2008)." Revealed preference is consistent with the idea that new entrants often fail: Netflix was initially only attracting show ideas that everyone else had passed on (Dawn, 2013).

The low expectations for Netflix contributed to why Netflix ordered all its shows straight to series. Netflix needed to prove their own commitment to television (Weisman, 2014); no one wanted to make a pilot for a new entrant that might quit the industry without making a single show. But low expectations were not the only reason for Netflix to order shows straight to series; technological change is often cited as another cause. Unlike traditional broadcast networks that could only broadcast a fixed number of shows each week during prime-time hours, Netflix faced no such capacity constraint. For Netflix, even a poor show was one more show that could attract subscribers on their service (Weisman, 2014); piloting has little value when a show was worth developing regardless of the information gained from the pilot. Finally, the lack of internal resources may have played a role: initially Netflix only had one person assigned to developing new shows (Weisman, 2014).

In 2013 Netflix surprised the industry when two of its shows, *Orange Is the New Black* and *House of Cards*, garnered Golden Globe and Emmy nominations. Afterwards, Netflix turned into a desirable place for creators to bring new show ideas (Adalian, 2013). As depicted in Figure 1, incumbent networks began skipping pilots for their own shows, both because Netflix's entry suggested commitment could pay off and in fear of losing shows to Netflix since show creators preferred commitment by the networks (Adalian, 2013; Brembilla, 2013). As 2013 was the year Netflix successfully broke into the prestige television industry, for the purposes of this paper I refer to 2013 as Netflix's true entry date rather than its initial funding of shows in 2011.

2.4 Modeling the Decision to Experiment

2.4.1 Optimal Experimentation Decision

The existing strategy literature is missing a formal model of the decision to experiment. The foundational models of real options apply to sequential decisions to invest across multiple stages (Roberts and Weitzman, 1981; McDonald and Siegel, 1986; Majd and Pindyck, 1987) and therefore focus on answering questions such as when it is optimal to stop investing rather than when to commit by investing for more than one stage at once. In contrast, the existing empirical work that does investigate variation in multi-stage investments (Thomke, 1998; MacCormack, Verganti and Iansiti, 2001) lacks a model and empirical approach accounting for the endogenous experimentation decision and therefore provides correlations rather than causal findings.

To fill this gap, I take advantage of the experimentation approach used by Nanda and Rhodes-Kropf, (2016) in the entrepreneurial finance literature. In their paper they illustrate the use of staging investments by venture capitalists to overcome the uncertainty inherent in startups. Partially funding a startup both generates a signal of the startup's viability and provides the VC with the option to further invest at later stages. In contrast, the VC can also choose to fully fund a startup without a signal. Committing to the startup may be worthwhile when the incremental cost of staging investments is greater than the value of the generated signal. Their modeled tradeoff between real options and commitment is not unique to venture capitalists; it more generally applies to any decision to experiment as a means of resolving uncertainty.

For a particular project i, the decision maker an ex-ante non-negative expected belief θ_i about the outcome, drawn from a set of non-negative outcomes Θ. Executing on the project has fixed cost C associated with it, regardless of the outcome. For an additional cost e, the decision maker can generate a signal about the project which results in a posterior belief $y(\theta_i)$ about the project where y is a non-negative random value. The distribution of the posterior belief is dependent only on θ_i, has expected value $E[y(\theta_i)] = \theta_i$ and would always be worth doing if the cost of experimentation was zero:

$E[y(\theta_i)|y(\theta_i) > C] > \theta_i$. Subject to the payoff being positive, the decision maker would therefore maximize:

$$\max_{\lambda_i \in \{0,1\}} \lambda_i(\Pr(y(\theta_i) > C)\, E[y(\theta_i) - C|y(\theta_i) > C] - e) + (1 - \lambda_i)(\theta_i - C) \tag{1}$$

Equation 1 generalizes the model provided in Nanda and Rhodes-Kropf (2016). To include the concepts of the benefits and drawbacks of experimentation, I extend this model by added two terms in Equation 2. V_B represents the value of the benefits of experimentation over commitment, which for example can include incorporating learning from the experiment to improve the final outcome. In my setting this would be the case if piloted generated useful feedback from network executives that enables the creators of television shows to improve their shows. V_D represents the drawbacks to experimentation, if for example actors became harder to recruit, lowering final outcomes.

$$\max_{\lambda_i \in \{0,1\}} \lambda_i(\Pr(y(\theta_i) > C)\, E[y(\theta_i) + (V_B - V_D) - C|y(\theta_i) > C] - e) + (1 - \lambda_i)(\theta_i - C) \tag{2}$$

Definition 1: Define the *expected return from experimentation* as:

$$\Pr(y(\theta_i) > C)\, E[y(\theta_i) + (V_B - V_D) - C|y(\theta_i) > C]$$

Assumption 1: Assume $V_B - V_D < C$, so that not all projects have a positive expected return from experimentation.

Definition 2: Define the partial ordering of signal technologies $y' \geq y$ if and only if for any θ_i, V_B, V_D and C under A1, the expected return from experimentation is weakly greater for y' than for y.

As an example of this signal technology ordering, let $V_B - V_D = 0$, $y(\theta)$ have a 50% probability of being 0 and a 50% probability of being 2θ while $y'(\theta)$ has a 75% probability of being 0 and a 25% probability of being 4θ. Then for any θ and $C \geq 0$ the difference between the expected return from experimentation with y' versus y is $\frac{1}{4}C > 0$. The more precise signal y' can justify higher cost of experimentation e than the less precise signal y.

Theorem 1: Experimentation modeled in Equation 2 is increasing in y, V_B,$-V_D$, and $-e$ under A1.

Comment 1: The relationship between θ_i and experimentation modeled in Equation 2 with A1 is ambiguous.

Interestingly the intuition that ex-ante better projects will not use experimentation does not hold in Equation 2 without additional assumptions. For example, suppose at $\theta_i = 1$, y has a $2/3$ chance of being $3/2$ and a $1/3$ chance of being 0 while at $\theta_i = 2$, y has a $1/3$ chance of being 2 and a $2/3$ chance of being 0. If $c = 1/2$ and $e = 1/4$, then at $\theta_i = 1$ the payoff from direct investment is $1/2$ which is greater the payoff from experimentation, $5/12$. In contrast, at $\theta_1 = 2$ the payoff from direct investment is $3/2$ which is less than the payoff from experimentation, $19/12$. Experimentation happens with the higher value of $\theta_i = 1$ but not at the lower value of $\theta_i = 1$.

The example of Comment 1 illustrates the typical failure case where the signal technology gets better at detecting poor projects as θ_i increases. However, in many cases experimentation is increasing in θ_i leading to the decision in Equation 2 having two cutoffs points: θ' and θ''. θ' is the lower bound threshold for experimentation. Projects with ex-ante beliefs below θ' do not have a positive payoff after experimentation and will not be considered. θ'' is the upper bound threshold for experimentation. The experimental signal does not terminate enough projects when beliefs are above the θ'' cutoff to justify the cost of experimentation, yet these projects are profitable so they are executed directly without the experiment. All projects with ex ante beliefs between θ' and $\theta''$are experimented on and conditionally executed depending on the signal. Figure 2 plots the optimal decision to experiment for an example signal and cost parameters as well as what observed outcomes would be under the optimal decision. Although Comment 1 suggests there may not be a straightforward correlation between experimentation and ex-ante project quality, the rest of this paper will assume so as in this cutoff rule case to address potential concerns about selection bias in the results.

Assumption 2: $E[y(\theta)|y(\theta) > C] - \theta$ is decreasing in θ.

Assumption 3: If $\theta' > \theta$, then $y(\theta') >_{FOSD} y(\theta)$.

Theorem 2: Under A1-A3, experimentation modeled in Equation 2 is increasing in $-\theta$.

2.4.2 Observed Outcomes

Observed outcomes for non-experimentation projects in the cutoff case would be $E[\theta|\theta'' < \theta]$ and for experimentation projects $E[y(\theta) + (V_B - V_D) - C|y(\theta) > C, \theta' < \theta < \theta'']$. Suppose two sets of observed projects were drawn from different distributions, A and B. Due to the conditional expectation form of these payoffs, the likelihood ratio order would be a natural ordering for these distributions.

The likelihood ratio order has two important consequences for Equation 2's project outcomes. First if $A >_{LR} B$ then given any cutoff range for experimentation, $[\theta', \theta'']$, the probability of passing the experimentation stage should be higher for projects drawn from A than B. This can be used as supportive evidence about whether $A >_{LR} B$ holds.

> Theorem 3: Under A3 when $A >_{LR} B$ and observations for both distributions are drawn from the same range then the probability of A drawn observations passing the experimentation stage will be higher than for B drawn observations.

Second, we can potentially see that experimentation is worthwhile for draws taken from A but not B given a shared region of experimentation $[\theta', \theta'']$.

> Lemma 1: Under A1-A3, the difference between experimentation and non-experimentation outcomes is decreasing in θ_i

> Theorem 4: Under A1-A3, the expected difference between experimentation and non-experimentation outcomes is decreasing in the likelihood ratio ordering of distributions for θ_i.

Finally, what happens to the average outcomes of a portfolio of projects as a project is moved from experimentation to non-experimentation? If there is no net downside to experimentation the portfolio of shows always worsens. Experimentation is unable to filter out the moved project, lowering the average outcomes for the portfolio.

> Theorem 5: Let E be a set of experimented projects and N be a set of non-experimented projects. If $V_B - V_D$ is nonnegative, the average outcome across both sets E and N can only worsen if a project is moved from E to N.

2.4.3 Implications for Empirics

How might this model apply to a network's decision to pilot a particular television show idea i? Each idea likely varies in the network's beliefs about outcome θ_i. Pilots cost more than an average episode to produce due to for example the shorter-term labor contracts involved (Anonymous Emmy Nominated Producer, 2017). Evaluating the pilot is another expense the networks must consider (Bunn, 2002). e represents both these production and signal evaluation costs. From network's perspective, cost C is primarily a fee for the right to broadcast the show, generally 80% of the cost of the show's production. e and C's independence from i is justified by the fact that networks do not consider the cost of a specific script when making pilot or series order decision for a script (Anonymous ABC Executive, 2017). Although a show on HBO will cost far more than one on ABC, variance in cost between two scripts is lower within the same network. Cost is instead involved when determining the total number of scripts to pilot or order to series; in 2006 NBC for example chose to reduce the number of scripted dramas it ordered rather than attempt to reduce the cost per drama (Barnes, 2006).

The model provides a set of testable hypotheses. First, Netflix's entry is thought to have had several potential effects on the incumbent networks. Netflix could have changed how much the networks needed to compensate creators for experimenting, increasing cost e. Networks could have also decreased their beliefs about the net effect of experimentation on outcomes, either by dropping V_B or increasing V_D. Theorem 1 suggest any of combination of these potential consequences of Netflix's entry would have led to less piloting as illustrated in Figure 1. In addition, an uncompensated bargaining model would also produce the same result, see Appendix for details.

> Hypothesis 1: *Netflix's entry should have increased straight to series production by incumbents.*

Second, according to Theorem 4, when the distribution of a set of potential shows is "better" in terms of likelihood ratio, the improvement piloting has in outcomes over not piloting should decrease. The direct interpretation of the likelihood ratio order would be that as θ_i increases for a given i, the probability i was drawn from A relative to B also increases if $A >_{LR} B$. In my television context this is also equivalent to saying a) all network's utility from a show weakly increase in θ and b) regardless of the exact shape of a network's utility function, for any given target range $[a, b]$ of show quality θ, networks would prefer to take the average draw from distribution A over B. Based on Theorem 3, looking differences in the probability of a show passing the pilot phase is an indication of when a distribution could be better in this sense. Differentiating observations based on the probability of passing the pilot phase could uncover cases where piloting worsens outcomes.

> Hypothesis 2: *Piloting should be less beneficial to outcomes from distributions of shows that are more likely to pass the pilot phase.*

Third, from Theorem 5, the selection of shows into straight to series production based on ex-ante characteristic would have no effect on the portfolio of shows. This implies that if an improvement in portfolio outcomes is observed after Netflix's entry, more straight to series production did improve show outcomes.

> Hypothesis 3A: *If there is no net downside to piloting, a network cannot improve its portfolio of new shows by ordering more shows straight to series.*

> Hypothesis 3B: *If a portfolio's outcomes improve after increasing straight to series production, a net downside to piloting must exist, regardless of whether the networks are selecting ex-ante better shows for straight to series production.*

Finally, suppose evidence suggests a downside to piloting exists. The improvement of the first episode relative to the rest of the season can provide clues as to some underlying mechanisms behind this downside. Rather than model all possible mechanisms of how experimentation can affect outcomes, I characterize them into three broad categories. Those that that improve outcomes of the final stage relative the interim, experimentation stage, those that affect the outcomes of both stages equality and those that improve the interim stage relative to the final stage relative. The appendix provides simple models of each mechanism to justify its classification.

> Hypothesis 4A: *If the rest of the season improves relative to the first episode under straight to series production, piloting could be distorting incentives towards passing the pilot phase.*

> Hypothesis 4B: *If the first episode improves relative to the rest of the season, piloting could lower incentives to invest in the first episode or learning from the pilot feedback could be improving later episodes.*

> Hypothesis 4C: *If the first episode and later episodes are affected equally by straight to series productions, the mechanisms in 5A and 5B are less likely to be important relative to other mechanisms.*

2.5 Data and Measures

To empirically test the above hypotheses, I pool data from three sources: Film L.A., Gracenote, and the Internet Movie Database (IMDb)[1]. Film L.A. is a non-profit dedicated to facilitating film and television production in Los Angeles. They have a proprietary dataset which tracks the production of scripted US television starting at the pilot phase. Importantly, the dataset flags shows that were ordered straight to series phase, a variable crucial to this paper's analysis. Gracenote, a subsidiary of Nielsen Holdings, has a dataset provided commercially to the television industry. A record is made whenever a network makes a public announcement of investment in a show idea, by for example paying a writer to produce a script. Metadata is associated with each show such as genre and creators responsible for the show's production. IMDb, a subsidiary of Amazon.com, has a public dataset which includes ratings for shows that made it to a public airing on a network.

Joining these three datasets is non-trivial because of variances in a show's title, year of production and network across the datasets. Since for example Film L.A. creates a show record earlier in the show's production history than IMDb, the title used in Film L.A. may be a working title, different from the official, release title used in IMDb. In both databases, spelling errors can exist in their title fields. Film L.A. tracks shows by their development season while IMDb records the year of a show's first broadcast; a show broadcast early in 2014 according to IMDb might be labeled as part of the 2013 development season according to Film L.A. Film L.A. tracks which network each show was developed for while IMDb's distributors for each show is often incomplete. Similar issues appear when comparing records between Film L.A. and Gracenote or Gracenote and IMDb.

To build my matched dataset, I first treat the smaller, curated Film L.A. dataset my main set of observations. I then build lists of alternative titles for each television show and match these alternative titles across the three datasets using bigram matching to allow for spelling errors. I prioritize records with exact matches across these alternative titles, year and network, but allow for deviations in year and network when exact matches are not available. Table 1 provides summary statistics for the combined dataset.

[1] Information courtesy of IMDb (http://www.imdb.com). Used with permission.

The data is restricted to the incumbent networks that circa 2008 were consistently producing scripted television. This includes the prestige networks that would win Emmy or Golden Globe awards (ABC, NBC, FOX, CBS, HBO, FX, USA, Showtime and AMC) as well as other established networks known for creating original content (CW, Freeform, TNT, SyFy, Starz, and A&E). By restricting my dataset to these fifteen networks, my analysis is focused on the incumbents' reaction to Netflix's entry.

Since the Film L.A. dataset is only contains shows between 2008 and 2017, all observations of shows outside those years in the other datasets are not included. Data for 2017 is currently only partially available. The years used are season years; for example, the 2008 season year runs from September 2008 to August 2009.

The funnel from script to pilot to series is represented by the first few rows of Table 1. I restrict my data to show ideas that were developed in some way, either piloted or ordered straight to series. In the period prior to Netflix's entry, only 3% of shows were ordered straight to series, increasing to 15% after Netflix's entry. The genre and show length variables provide some indication of the stability of show types over my period of interest despite this change in production.

Based on Gracenote's data on the past work of a show's creators, I create a binary variable which indicates whether any of the show's creators have previously created a show that won a major Emmy or Golden Globe award. This constructed variable could be interpreted as a measure of a new show's uncertainty. A major concern for the networks is whether creators can "through skill and/or luck, manage to assemble all the right elements (cast, director, score, VFX, etc.) to perfectly execute the script they've written?" (Hawley, 2014); the creators with award-winning shows have a track record of having done so in the past.

I only observe outcomes for shows that were ordered to series, either by first being piloted or directly through a straight to series order. IMDb provides show ratings at both the show level and episode level for all broadcast shows. My primary outcome variable is an average of episode ratings for the show's first season. Table 2 shows strong correlation between IMDb ratings and two other measures of a show's success: the renewal decision made by the networks and breaking into the Nielsen Top 30. Renewals have historically been a strong indicator that a show met the network's internal metrics for success. Unlike raw viewership numbers, renewals factor in the value of reaching a specific demographic and the strength of a show's timeslot competition on rival networks. The connection between renewal, Nielsen viewership and a show's success is not completely mechanical; sometimes a show will be considered a success if it for example increases cable subscriptions despite having relatively low viewership and lacking plans for renewal (Thaxton, 2017; O'Connell, 2018). Unlike viewership numbers from Nielsen Media Research or a network decision to renew a show for another season, IMDb ratings are less mechanically linked to competitor outcomes and therefore relatively independent of market structure (Waldfogel, 2017). For example, decreasing piloting across all networks could have improved the quality of television shows without affecting viewership numbers or renewal rates if demand is inelastic.

2.6 Empirical Framework

2.6.1 OLS Estimator

An OLS approach could be used to estimate the relationship between piloting and outcomes.

$$FirstSeasonRating_{int} = \alpha_n + \delta_t + \beta Piloted_i + X_i + \varepsilon_{int} \tag{3}$$

In Equation 3, $FirstSeasonRating_{int}$ is the average IMDB rating of a show's first season episodes, X_i includes any show level controls while α_n and δ_t are network and year fixed effects. β is the coefficient of interest that indicates the improvement in show quality from piloting.

Based on the model defined in Equation 2, Equation 3 would estimate

$$\beta = E[y(\theta_i)|y(\theta_i) > C, \theta_i < \theta'', Z_{int}] - E[\theta_i|\theta_i > \theta'', Z_{int}] + V_B - V_D \tag{4}$$

with Z_{int} representing the controls a_n, δ_t and X_i. The selection bias manifests itself in the conditional terms $\theta_i \gtrless \theta''$; ex-ante higher θ_i shows would be produced straight to series as predicted in Hypothesis 2. If the controls Z_{int} are sufficient for the conditional independence assumption to hold, Equation 3 would estimate

$$\beta = E[E[y(\theta_i)|y(\theta_i) > C, \theta_i] - \theta_i]|Z_{int}] + V_B - V_D \tag{5}$$

In Equation 5, β measures the average treatment effect in changing production: how much are broadcast show outcomes improved by switching from piloting to a straight to series production model. The first term is positive since $y(\theta_i)$ is a garbling of θ_i and represents the attrition of shows that failed the pilot phase. As this attrition is a major benefit of piloting, it should be included in my estimate. Since this first term is positive, the estimate of β can only be negative when the net effect of piloting on outcomes, $V_B - V_D$, is sufficiently negative.

Hypothesis 2 predicts a heterogeneous effect depending on the distribution of θ within the piloting range of θ' to θ'': β is should be smaller when piloting is less likely to weed out bad show ideas. Ideally, we would like to find some covariate W_i such that $E[y(\theta_i)|W_i = 1] > E[y(\theta_i)|W_i = 0]$; W_i differentiates shows with a high or low probability of passing the pilot phase. However, distinguishing such shows may not trivial. Equation 2 suggests estimating the probability a show would pass the pilot phase based could be biased due to the selection of shows into piloting:

$$E[y(\theta_i)|W_i, \theta_i < \theta''] \neq E[y(\theta_i)|W_i]$$

This bias could be limited by looking the period between 2008 and 2013, before Netflix entered, when almost all shows were piloted if we assume the signaling technology hasn't changed over time.

$$E[y(\theta_i)|W_i, \theta_i < \theta'', t \leq 2013] \approx E[y(\theta_i)|W_i, t \leq 2013] \approx E[y(\theta_i)|W_i]$$

Figure 3 suggests likely candidate for a strong covariate W_i: it plots the share of piloted shows that pass the pilot phase by whether any of the show's creators had a previous award-winning show. In the time period prior to Netflix's entry, shows with these award-winning creators had a greater chance of passing the pilot phase. Although passing the pilot phase does not necessarily mean a positive signal was observed (or a negative signal when show does not pass the pilot phase), I assume passing the pilot phase is positively correlated with a positive signal. Therefore, I use Equation 7 to operationalize Hypothesis 2, expecting if piloting can worsen outcomes, I am most likely to observe that outcome using an interaction of piloting and shows with award-winning creators.

$$\begin{aligned} FirstSeasonRating_{int} = \alpha_n + \delta_t + \beta_S Piloted_i + \beta_A CreatorAward_i \\ + \beta_{SA} Piloted_i * CreatorAward_i + \varepsilon_{int} \end{aligned} \tag{7}$$

The two coefficients of interest are β_S and β_{SA}. β_S provides an estimate of Equation 4 when the benefit of piloting should be relatively higher. β_S should be greater than the average treatment effect measured by β in Equation 3. Conversely, β_{SA} estimates the change in Equation 4 when the benefit of piloting is lower: experimentation less useful in weeding out bad ideas. The sum $\beta_S + \beta_{SA}$ represents the overall effect of piloting on shows with creators having award-winning shows which should be lower than Equation 3's β. It is possible that β_S is positive while $\beta_S + \beta_{SA}$ is negative, which would suggest piloting is only useful for shows from creators without a strong track record.

2.6.2 Matching Estimator

Using Equations 3 and 7 to estimate the effect of piloting has two main drawbacks. First, it may not be reasonable to assume conditional independence holds; there could be unobserved variables causing correlation between the observed variables and error term, for example selection bias. To address this issue, I restrict my estimator to use incumbent network observations from 2014 to 2017. After Netflix's successful entry using a straight to series production, the incumbents held heterogenous beliefs about the value of piloting. Some network executives like at NBC embraced straight to series production more than others like at CBS which viewed it as a fad (Collins, 2014; Andreeva, 2014). This resulting in close counterfactual shows: a crime drama picked up at NBC would be ordered straight to series while an otherwise similar show at CBS would be piloted. I interpret my area of common support as the set of shows that had the possibility of either being piloted or ordered straight to series, depending on the network executives in place at any time. The estimator measures the average treatment effect of piloting on these types of shows.

Second, even if conditional independence holds, the linear assumption of OLS could be problematic. Saturating the control variables is not possible with my dataset's number of observations which can lead to selection bias. For example, suppose individual genres were included in Equation 3's X_i term but not combinations of genres and that sci-fi comedies were only observed as committed shows with award executives. Any estimate of β_{SA} could be picking up the fixed effect of a show being both a sci-fi and a comedy rather than measuring the return to experimentation. Matching estimators can mitigate this issue when observable covariates are balanced. The example's observations of sci-fi dramas would in theory be removed from a matching estimator because they would lead to imbalance in the share of sci-fi and comedy shows between the piloted and straight to series groups.

I use a propensity score based estimator for the average treatment effect (Hirano, Imbens and Ridder, 2003) since coarsened methods (Iacus, King and Porro, 2012) require a more favorable ratio of observations to covariates than exists in my dataset. Propensity score matching requires a two-stage estimator, first estimating how observables affect the probability a script will be piloted versus ordered straight to series:

$$StraightToSeries_{int} = \alpha_n + \delta_t + \gamma X_i + \varepsilon_{int} \tag{8}$$

Then observations are restricted to those where the predicted $\widehat{StraightToSeries}_{int}$ lies in a region of common support. Finally, Equations 3 and 7 are estimated within those common support observations, weighted by the inverse propensity score.

2.6.3 Portfolio Estimator

The matching estimator leaves selection on unobservables potentially unmitigated as a source of bias. However, Hypothesis 3 suggests a one-way test of selection bias: if piloting on average improves outcomes, then the decrease in piloting by incumbents induced by Netflix's entry could only have worsened incumbent outcomes. I use a difference in difference estimator at the network's portfolio level to test this hypothesis:

$$\begin{aligned} AverageFirstSeasonRating_{nt} \\ &= \beta_S ShareStraightToSeries_{nt} \\ &+ \beta_{SP} ShareStraightToSeries_{nt} * Post2013_t + \delta_t + \alpha_n \\ &+ \epsilon_{nt} \end{aligned} \tag{9}$$

Equation 9 approaches my dataset as a panel of network observations. For each year t, each network n releases a set of new shows. The average first season rating for those shows is my outcome variable $AverageFirstSeasonRating_{nt}$. Some of those new shows were ordered straight to series, while others were the result of a piloted, staged development process. The share of the network's shows that year that were ordered straight to series is reflected in the dependent variable $ShareStraightToSeries_{nt}$. Netflix's entry into the industry increased the incumbents share of straight to series orders, this time shock is represented by the $Post2013_t$ variable. Fixed effects for networks α_n control for the differences in average show quality across networks; for example, on average HBO shows tend to be rated higher than shows on the big four networks. Fixed effects for year δ_t control for time trends in show quality, if for example the industry is overall getting better at producing higher quality shows at the end of my time period relative to the beginning.

If β_{SP} is non-negative, straight to series orders must have some net benefit to the quality of shows regardless of whether networks are selecting ex-ante better shows for straight to series production. An important consideration in this empirical approach is whether the types of shows commissioned by incumbents was stable before and after Netflix's entry. I make three arguments supportive of this approach. First, as Table 1 indicates, observables for incumbent shows did not change after Netflix's entry even though the rate of piloting did change. This suggests a stability in the type of content being produced. Unlike HBO's success with the *Sopranos* a decade before, the types of shows produced by Netflix were not radically different from existing television shows. Second, until the 2018-2019 season, Netflix was unable to attract the creators working for the incumbent networks, many of which are on long term contracts (Littleton, 2017). The creators coming up with ideas for new shows on incumbent networks was consistent across the time periods. Third, if Netflix was able to attract incumbent creators, it's natural to expect the networks most at risk of losing talent to switch to straight to series production to retain creators. This works in my favor; the decrease piloting again should correlate with worse outcomes since networks that reduced piloting were the same ones losing creators. If outcomes instead improved, this both argues for a downside to piloting as well as stability in incumbent creators over time.

2.6.4 Mechanism Estimator

There are several mechanisms that could cause an improvement in outcomes when experimentation is avoided. One way of divining which mechanisms are more likely present in my setting is to break out the effect of ordering a show straight to series by each broadcast episode, as outlined in Hypotheses 4A to 4C.

$$\begin{aligned} &EpisodeRating_{eint} \\ &\quad = \lambda_e EpisodeNumber_e + \beta_e EpisodeNumber_e \times Piloted_i \\ &\quad + \alpha_n + \delta_t + \varepsilon_{eint} \end{aligned} \tag{11}$$

In Equation 11, i still represents a show but now e represents one of the show's episodes, specifically an episode that was part of a show's first season order: the follow-on series order if a show was originally piloted or the initial straight to series order if show was not piloted. $EpisodeNumber_e$ is the ordinal number of an episode's broadcast. λ_e picks up a trend for how the IMDB rating of a piloted show evolves over its episodes while β_e, the coefficient of interest, is how the quality of straight to series shows evolve differently than piloted shows.

2.7 Empirical Results

2.7.1 Show Level Results

Column 1 of Table 3 estimates Equation 3 without any fixed effects or covariates. Overall, we see piloting correlates with worse outcomes in my data. However, as observables are added in Column 2 the relationship attenuates. Using Oster's (2016) approach to adjusting estimates for selection on unobservable based on the amount of selection on observables, the bias adjusted point estimate relationship between piloting and ratings is positive, suggesting my OLS estimates suffer from strong selection bias.

To mitigate this selection bias, I use a propensity score matching estimator. Table 4 shows the propensity score estimate of Equation 8. The strongest predicator of straight to series is whether the show received independent funding. Such funding would enable a show to bypass the network's usual funding process of first piloting a show so it's natural for these shows to be ordered straight to series. The intuition behind my matching estimator is evident in the difference between Columns 1 and 2 as year and network fixed effects are added and the estimate's pseudo R^2 increases: in the period after Netflix's success, the industry had not converged on a reevaluation of straight to series show production. Some network executives were exploratory in increasing straight to series production more than others, leading to similar shows on different networks in different years receiving different types of production orders. The propensity score estimator is used to find a region of common support: shows that were not certain to be piloted or ordered straight to series but could have gone either way depending on the circumstances of the network at the time.

Figure 4 provides a visualization of which observations are being dropped for common support based on this propensity score estimate. The matching estimator focuses on the difference in outcomes between straight to series shows that were predicated to be piloted and piloted shows that control for those straight to series shows. Table 5 shows how the covariates used in the propensity score logistic prediction equation differ between the full sample and matched sample. Overall balance is improved and although differences between straight to series and piloted shows remain, none of the differences in means is statistically significant at the 10% level.

Returning to the matching results of Table 3, Column 3 again estimates Equation 3 without any fixed effect or covariates, now using inverse probability weights within the region of common support. As in Column 1, a negative correlation between piloting and IMDb ratings is exposed. However, unlike in Column 2, Column 4 shows this relationship strengthens as fixed effects and covariates are added, suggesting selection on unobservables is less of concern within this matched sample.

Equation 7's exploiting of a show's association with creators having prior award-winning shows is estimated in Column 3 of Table 6 using the matching estimator. The relationship between piloting and outcomes presented in Column 1 is revealed to be primarily driven by creators with award-winning shows. Piloting shows without such creators improves outcomes, consistent with the winnowing effect of experimentation dominating any drawbacks from the lack commitment. In contrast, piloting shows with strong creators lowers show ratings by a full standard deviation.

The matching results suggest the optimal strategy for shows within the area of common support is to order them straight to series when they have award-winning creators and pilot them otherwise. If an incumbent network were to undertake this optimal strategy, their overall renewal rate for new shows would increase by 5%, or about half a new show. Given the costs of new show development and typical advertising revenue per successful show, this would amount to an incremental $25M in profit for the network.

2.7.2 Network Level Results

Column 1 of Table 7 estimates Equation 9. According to Hypothesis 4B, a positive coefficient *Share StS * Post 2013* would indicate reducing piloting worsened outcomes, consistent with the matching approach measuring an unbiased estimate of the treatment effect of piloting. Although the result in Column 1 lacks statistical significance, it is directionally consistent with Hypothesis 4B.

Figure 6 provides a visualization of how Netflix's entry affected production at outcomes at incumbent networks. The left panel plots the majority of networks in my dataset while the right panel breaks out the big four US networks separately. These big four, ABC, NBC, CBS and Fox, are materially different from the rest of the U.S. networks. They each fund over a dozen pilots per year, have a long track record of airing original programming, share a history of transitioning from over-the-air broadcasting with affiliate networks to reaching most of their audience through cable, and directly compete for prime-time advertising spend. They also have a much higher rate of using creators with previous award-winning shows, as depicted in Figure 5.

Figure 6 suggests a strongly heterogenous effect on outcomes across incumbent networks. The typical network's portfolio of new shows was unchanged when Netflix induced less piloting. However, the big four markedly improved the portfolio quality after Netflix's entry. Column 2 of Table 8 provides a triple difference version of Equation 9, effectively breaking out the *Share StS * Post 2013* variable of Column 1 by whether the network was one of these big four networks. Column 2's *Share StS * Big Four * Post 2013* is both positive and significant, again consistent with the matching estimator in that reducing piloting would be benefit for shows with award-winning creators as common for the big for networks.

Note this result cannot simply be the result of Netflix forcing the big four networks to improve their shows as a high-quality new entrant, otherwise Column 2's *Big Four * Post 2013* coefficient would have the positive result rather than the *Share StS * Big Four * Post 2013* coefficient.

2.7.3 Mechanism Evidence

Figure 7 plots the change in episode quality from piloting using a matching estimator based on Equation 11 that sets the first episode as the base quality level. The later episodes seem to increase in quality over the season, suggesting the mechanism behind the downside of experimentation may involve distorting effort. Piloting could incentivize a show's cast and crew to focus on passing the pilot phase rather than make the highest quality show possible. This trend also is suggestive that the earlier estimates of the difference in show or portfolio outcomes due to piloting are not driven solely by selection bias; selection into piloting based on ex-ante show attributes would likely affect all shows equality without such a heterogenous effect between episodes.

The trend depicted in Figure 7 could help explain why the incumbent television networks did not make more use of straight to series before Netflix's entry, given that increasing straight to series production seems to have improved outcomes. Since the early days of television, advertisers have pre-paid for slots on new television shows, the pre-payment being used to fund the development of those shows. In return, the networks guarantee to deliver a certain number of viewers to those advertisers. When a new show falls short, the networks must provide advertisers with "make goods"; free advertising on their network until the gap is closed (Vogel, 2015). A new show that underperforms therefore eats away at a network's revenue by forcing the network to give away advertising it would otherwise charge for, leading the network to kill shows early that miss targets (Angwin and Vranica, 2006). The networks bias their evaluation of a broadcasted show's success towards the outcome of the show's first episode and therefore miss the benefits commitment provides in improving later episodes. Even if this gain from commitment was known to the networks prior to Netflix's entry, they may have lacked the ability to revise relational contracts with advertisers to measure quality later in the season and realize those gains from commitment (Gibbons and Henderson, 2012).

Finally, Figure 8 presents the relationship between the background of a show's creators, whether a show was piloted, and whether at least one of the two main actors on the show had previous experience as a main actor. The development process of television shows typically follows a timeline where the decision to pilot or order a show straight to series is done before casting. Figure 8 illustrates that creators with stronger backgrounds attract more of these veteran actors, supportive of anecdotal evidence that veteran actors are preferred in television show production (Anders, 2015). Figure 8 also shows piloted shows in general have a harder time attracting such talent relative to straight to series orders, which supports anecdotal evidence that actors prefer for long term contracts (Robinson, 2018). This preference for long term contracts is another potential mechanism behind the drop in IMDb ratings when shows with strong creators are piloted.

2.8 Conclusion

In this paper I use piloting process used in the television show industry to study the value of experimentation through staged development. Using an estimator that matches piloted shows with shows that were not piloted, my main result is that skipping piloting is positively correlated with outcomes for shows with creators that had strong track records. Netflix's entry as shock to the decision to pilot, I corroborate this result with a network level analysis, suggesting my main result may not be driven solely by selection bias and a causal interpretation could be warranted. Furthermore, I provide evidence on two mechanism behind such results. Improved outcomes from skipping piloting may derive from changed incentives: a show's creators no longer overinvest in the pilot to ensure their show passes the pilot phase. Or the improved outcomes could derive from being able to attract higher quality talent: actors prefer long term contracts over short ones.

These results can apply broadly to several areas within strategy. One is the inherent tension between the strategic approaches of real options and commitment. As small investments which, depending on what is learned over time, can optionally lead to a larger, follow-on investments (Bowman and Hurry, 1993), pilots can be thought of as a real option on a television show. Real options theory predicts pilots would unconditionally improve observed outcomes as potentially poor outcomes are filtered out at early stages: having been developed in financial contexts, real options approaches lacks a mechanism whereby the act of experimenting can itself affect outcome (Adner and Levinthal, 2004). Theories describing the consequences of commitment (Ghemawat, 1991) fill this gap, predicting piloting can the adversely affect outcomes. My empirical results show this tension is not just a theoretical one: neither real option nor commitment approaches to strategic decision making should be applied without considering their joint consequences.

Whether to evaluate research projects over short or long-time horizons is a key question for research on the innovation production function (Manso, 2011; Azoulay, Graff Zivin and Manso, 2011). When funding an idea for a new television show is cast in terms of a short-term pilot investment versus a long term full season investment, my results contribute to this innovation literature. One of the mechanisms I uncover in my setting has a close parallel in existing work on research environments: short term incentives can distort the actions taken by researchers (Manso, 2011). My other mechanism, worker preference for long term contracts, may be an important but less understood part of the innovation production function, perhaps hinted at with worker preference for publication (Stern, 2004).

Entrepreneurship is a third area where my results apply, where early stage venture capital investments are thought of as experiments (Kerr, Nanda and Rhodes-Kropf, 2014). My results suggest venture capitalists should fund startups created by an unproven groups of recent college grads for short periods of time while funding experienced serial entrepreneurs for a longer period. Although there is some potential loss in the lack of experimentation on serial entrepreneurs, this loss could be outweighed by incentivizing the entrepreneur to focus on the startup's long-term success rather than the requirements to meet its next short-term funding goals.

This paper's focus on commitment is timely as experimentation approaches espoused through movements like Lean Startup and implemented with anecdotal success in the technology industry begin to get transferred to other industries. The preference for experimentation before making large investments could be shifting funding away from necessary innovations that require long term investments which preclude experimentation, as in the clean energy industry (Gaddy, Sivaramand O'Sullivan, 2016; Burger, Murray, Kearney and Ma, 2018) and the pharmaceutical industry (Budish, Roin and Williams, 2016). The preference could also be forcing experimenting that is harmful to final outcomes, by for example starting new schools under the expectation that failure is expected, with long term consequences to the students enrolled in those schools (Duane, 2018). Experimentation may have well served the technology industry with its low cost of experimentation and perhaps low upside to commitment. However, these parameters may not exist in other industries. The use of approaches like Lean Startup need to be evaluated on a case by case basis as to not overlook the downside of experimentation: the lack of commitment.

2.9 References

Adalian, J., 2013. "Straight to Series: The Networks' Big-Money Bet to Skip Pilots." *Vulture*. November 27th.

Adner, R. and Levinthal, D.A., 2004. "What is not a real option: Considering boundaries for the application of real options to business strategy." *Academy of management review, 29*(1), pp.74-85.

Anders, C., 2015. "The Whole Crazy Process of Creating A TV Show, From Pitch to Pilot." *Gizmodo*. January 23rd.

Andreeva, N., 2014. "CBS' Nina Tassler: 'Pilot Season Is Not Perfect but Works For Us'". *Deadline Hollywood*, January 15th.

Azoulay, P., Graff Zivin, J.S. and Manso, G., 2011. "Incentives and creativity: evidence from the academic life sciences." *The RAND Journal of Economics, 42*(3), pp.527-554.

Baker, G.P., 1992. "Incentive contracts and performance measurement." *Journal of political Economy, 100*(3), pp.598-614.

Barnes, B., 2006. "Reality TV: NBC Universal to Slash Costs in News, Prime-Time Programs." *The Wall Street Journal*. October 19th.

Bhattacharya, S., Chatterjee, K., and Samuelson, L., 1986, "Sequential research and the adoption of innovations", *Oxford Economic Papers* 38, 219–243.

Bonhomme, S. and Jolivet, G., 2009. "The pervasive absence of compensating differentials." *Journal of Applied Econometrics, 24*(5), pp.763-795.

Bowman, E.H. and Hurry, D., 1993. "Strategy through the option lens: An integrated view of resource investments and the incremental-choice process." *Academy of management review, 18*(4), pp.760-782.

Paola Brembilla (2013), "Straight-to-Series in Broadcast Television: Causes, Issues and Consequences", in *Media Mutations 6 - Modes of Production and Narrative Forms in the Contemporary TV Series* (Conference Proceedings)

Budish, E., Roin, B.N. and Williams, H., 2016. "Patents and research investments: Assessing the empirical evidence." *American Economic Review*, *106*(5), pp.183-87.

Bunn, A., 2002. "The Wasteland." *The New York Times Magazine.* June 23rd.

Burger, S., Murray, F., Kearney, S., and Ma, L.. "The Investment Gap that threatens the planet." *Stanford Social Innovation Review.* Winter 2018.

Collins, S. 2014. "TV Pilots are still on the networks' radar." *Los Angeles Times.* April 2nd.

Dawn, C.C, 2013. "Ted Sarandos upends Hollywood with Netflix revolution", *Los Angeles Times.* August 25th.

Delmar, F. and Shane, S., 2003. "Does business planning facilitate the development of new ventures?". *Strategic management journal*, *24*(12), pp.1165-1185.

Duane, D. 2018. "How the Startup Mentality Failed Kids in San Francisco." *Wired*, Conde Nast, 10 July

England, P., Farkas, G., Kilbourne, B.S. and Dou, T., 1988. "Explaining occupational sex segregation and wages: Findings from a model with fixed effects." *American Sociological Review*, pp.544-558.

Farquharson, V., 2008. "Whole lot of Chaykin going on; The actor is more than kind to his new sitcom, but he's not so fond of a certain other Canadian series." *National Post.* August 21

Fontenay, C.C. and Gans, J.S., 2014. "Bilateral bargaining with externalities." *The Journal of Industrial Economics*, *62*(4), pp.756-788.

Gaddy, B., Sivaram, V. and O'Sullivan, F., 2016. "Venture Capital and Cleantech: the wrong model for clean energy innovation." *In MITEI-Working Paper No. 2016-06.*

Ghemawat, P., 1991. *Commitment.* Simon and Schuster.

Hawley, Noah. (2014) "'Fargo' Showrunner on a Straight-to-Series Order: It's the 'Holy Grail' (Guest Column)." *The Hollywood Reporter.* 20 Feb.

Holmstrom, B., 1983. "Equilibrium long-term labor contracts." *The Quarterly Journal of Economics*, pp.23-54.

Iansiti, M., "How the Incumbent Can Win: Managing Technological Transitions in the Semiconductor Industry," Management Science, Vol. 46, No. 2, February 2000, pp. 169-185.

Kerr, W.R., Nanda, R. and Rhodes-Kropf, M., 2014. "Entrepreneurship as experimentation." *Journal of Economic Perspectives*, *28*(3), pp.25-48.

Kogut, B., 1991. "Joint ventures and the option to expand and acquire." *Management Science*, *37*(1), pp.19-33.

Lerner, J., and Julie W., 2007. "Innovation and Incentives: Evidence from Corporate R&D." *Review of Economics and Statistics* 89(4): 634-644.

Littleton, C., 2017. "What Shonda Rhimes' Move to Netflix Means for Broadcast TV." *Variety.* Aug 14th.

Luo, H., 2014. "When to sell your idea: Theory and evidence from the movie industry." *Management Science*, *60*(12), pp.3067-3086.

MacCormack, A., Verganti, R. and Iansiti, M., 2001. "Developing products on 'Internet time': The anatomy of a flexible development process." *Management Science*, *47*(1), pp.133-150.

Milgrom P., Roberts J. (1990) "The economics of modern manufacturing: Technology, strategy, and organization." *American Economic Review.* 80(3):511–528.

Nanda, R. and Rhodes-Kropf, M., 2016. "Financing Entrepreneurial Experimentation." *Innovation Policy and the Economy*, *16*(1), pp.1-23.

Ocasio, A. 2012. "TV Success Rate: 65% Of New Shows Will Be Canceled (& Why It Matters)" *Screen Rant.*

O'Connell, Michael, 2018. "Showtime Boss Talks Subscriber Growth, 'Twin Peaks' and Mark Halperin." *The Hollywood Reporter.* January 6th.

Roberts, K. and Weitzman, M.L., 1981. "Funding criteria for research, development, and exploration projects." *Econometrica: Journal of the Econometric Society*, pp.1261-1288.

Robinson, T., 2018. "What Keeps You Alive shows the importance of art in horror films." *The Verge.* Aug 24rd.

Schmalensee, R., 1978. "Entry deterrence in the ready-to-eat breakfast cereal industry." *The Bell Journal of Economics*, pp.305-327.

Shaked, M. and Shanthikumar, J.G., 2007. *Stochastic orders.* Springer Science & Business Media.

Stern, S., 2004. "Do scientists pay to be scientists?". *Management Science*, 50(6), pp.835-853.

Thaxton, R., 2017. "Twin Peaks Breaks Showtime Subscription Records, Yet Notches Mediocre Ratings." *Pastemagazine.com.* May 23rd.

Thomke, S.H., 1998. "Managing experimentation in the design of new products." *Management Science,* 44(6), pp.743-762.

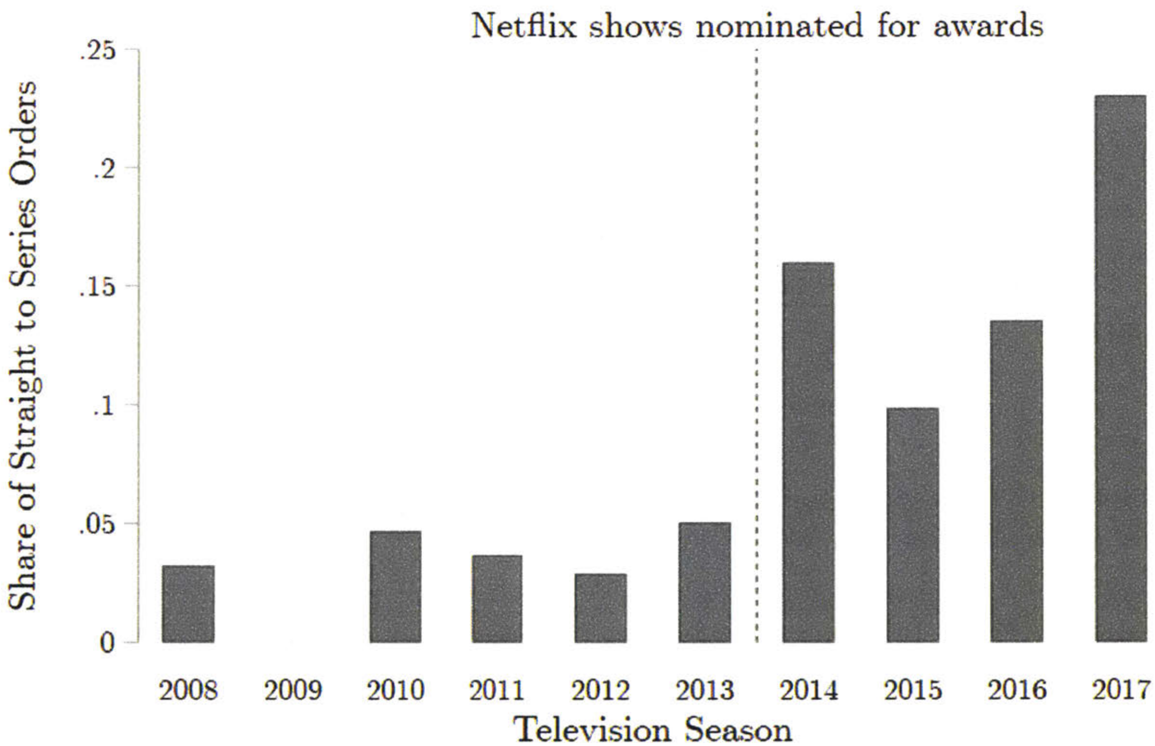
Waldfogel, J., 2017. "The random long tail and the golden age of television." *Innovation Policy and the Economy, 17*(1), pp.1-25.

Weisman, A., 2014. "The 3 Reasons Netflix Decided on Its Straight-To-Series Viewing Model." *Business Insider.* June 25th.

2.10 Figures and Tables

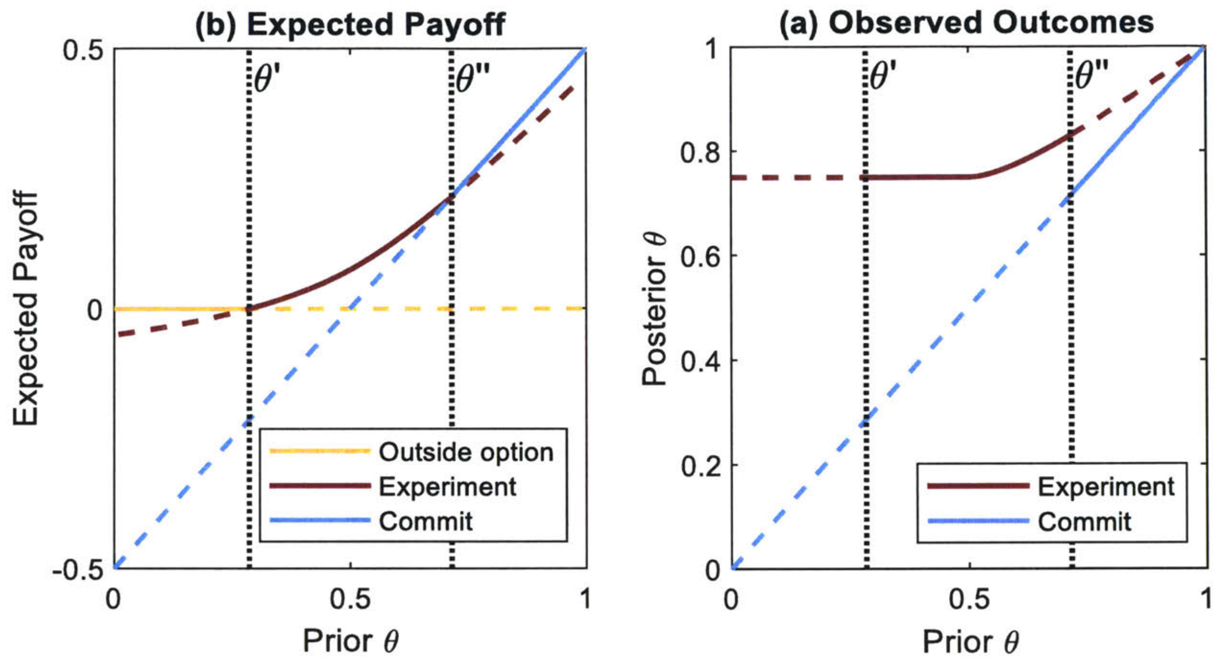
2.10.1 Figures

Figure 1. Share of Shows Ordered Straight to Series



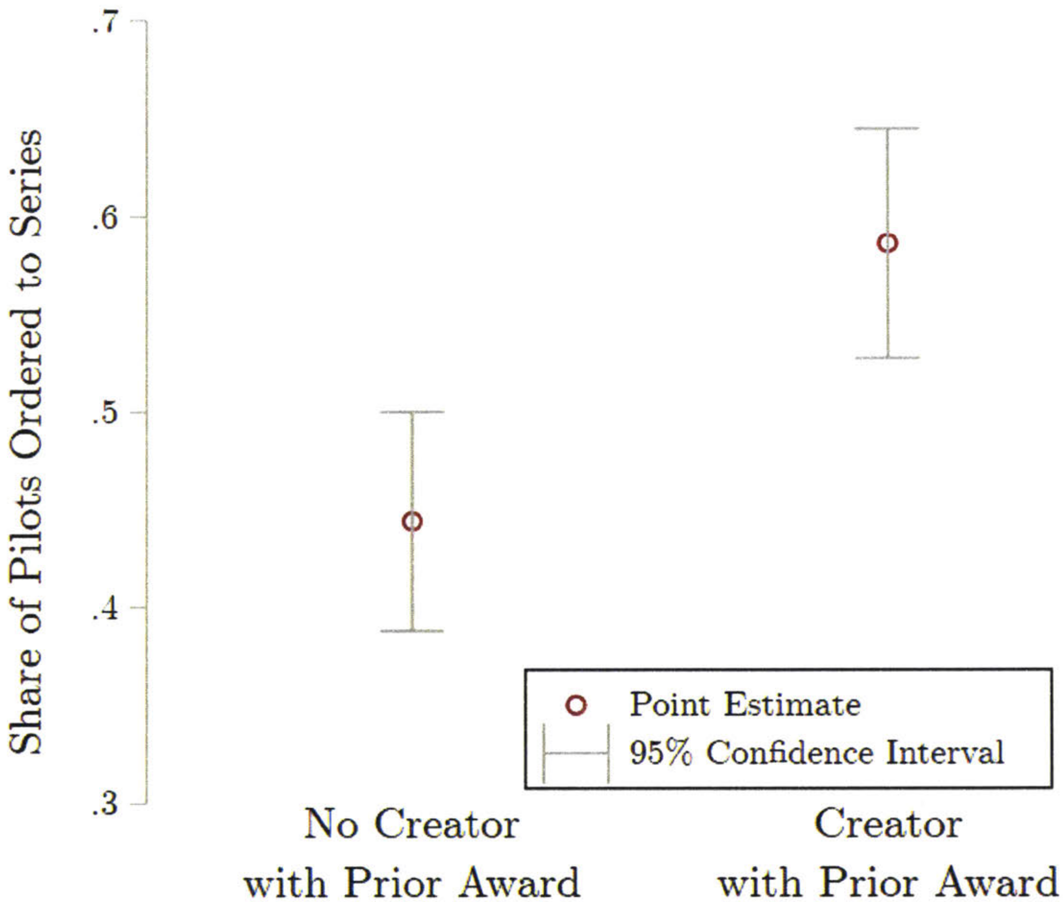
Scripted shows on incumbent networks that were either piloted or ordered directly to series.

Figure 2. Example Payoffs from Experimental Decision Making



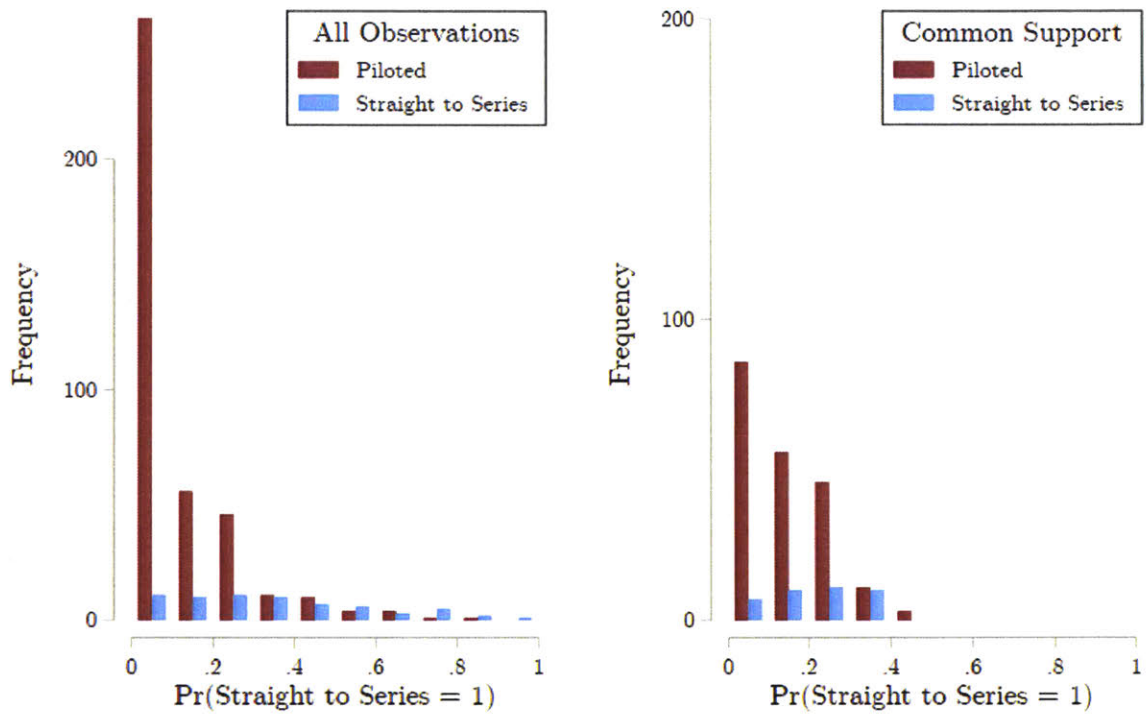
Assumes θ domain is $[0,1]$, project cost $C = 0.5$, experimental cost $e = 0.05$, $V_B - V_D = 0$ and a signal technology whose posterior distribution is mixture of uniform densities in ranges $[0, \theta]$ and $[\theta, 1]$ with mixture weights set so unconditional expected value of signal is θ. Panel (a) plots expected payoff conditional on prior belief θ. Optimum decision envelope represented by solid line. Panel (b) plots the observed outcomes under each production process with the optimum decision areas represented by solid lines.

Figure 3. Rate of Shows Passing Pilot Phase



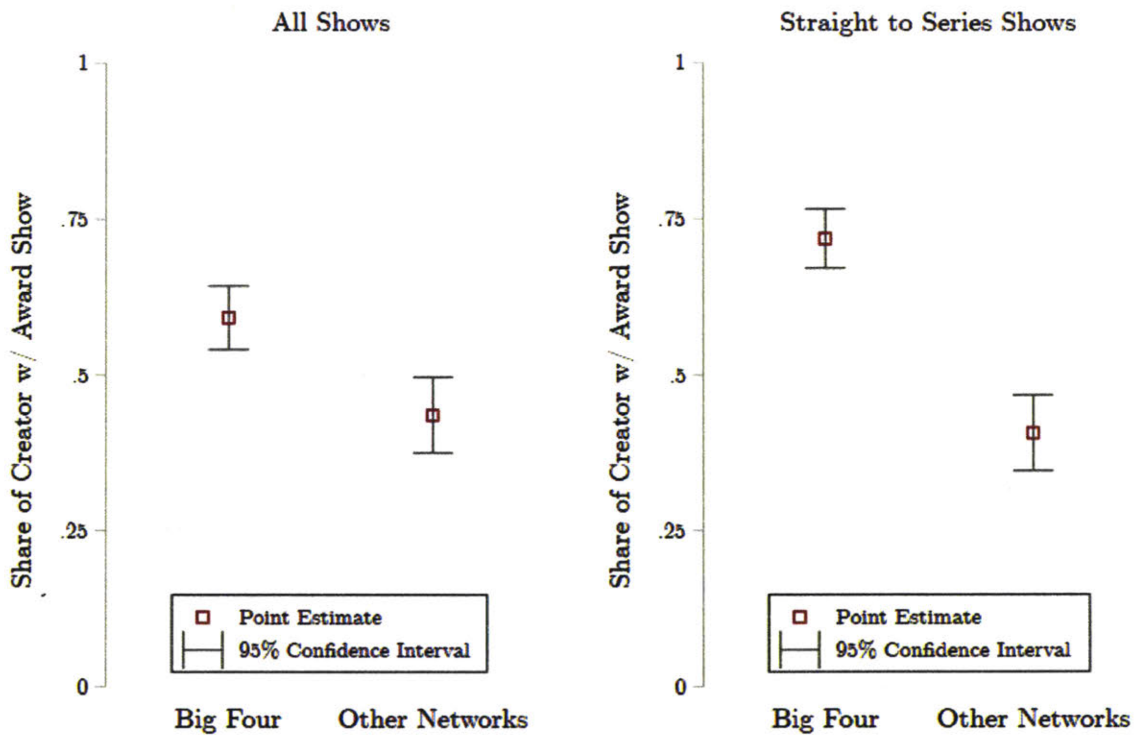
Piloted shows on incumbent networks from 2008 to 2013. Shows having creator with prior award have at least one creator with a previous show that one a major Emmy or Golden Globe award.

Figure 4. Prediction of Straight to Series for Inverse Probability Weighting



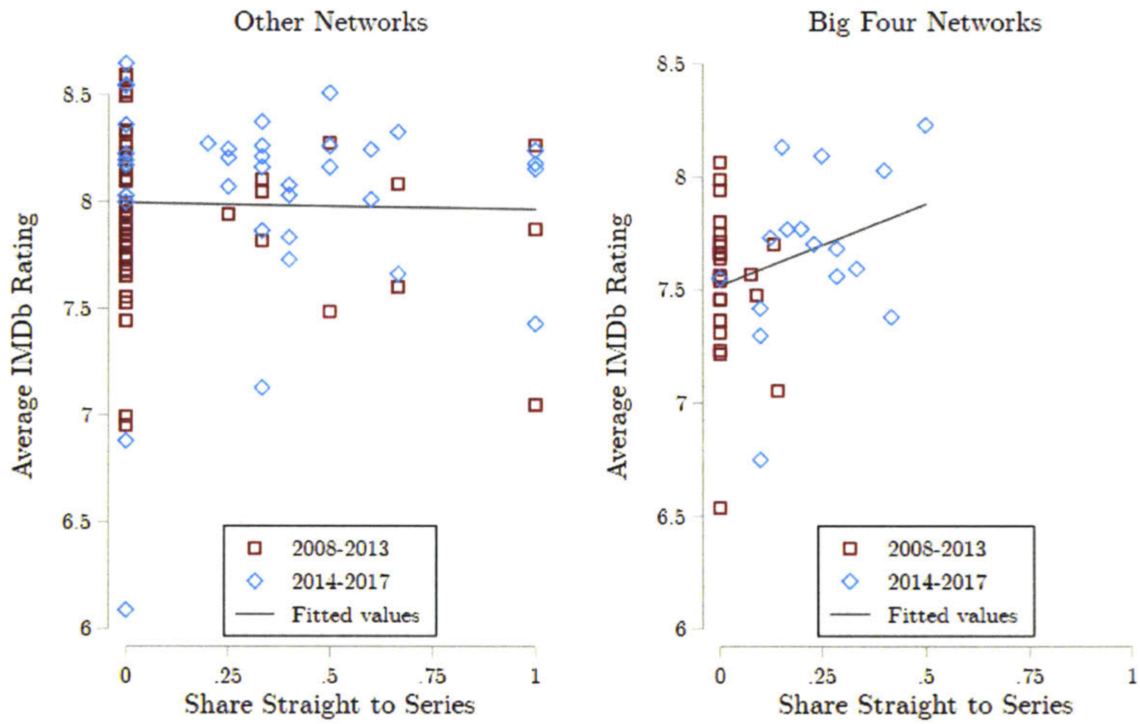
Distribution of predicted probability a piloted show will be ordered straight to series, based on incumbent network observations from 2014 to 2017. Common support cutoffs set to 5^{th} percentile of straight to series predicted probabilities and 95^{th} percentile of piloted predicted probabilities.

Figure 5. Share of Shows with Creators having previous Award-Winning Show



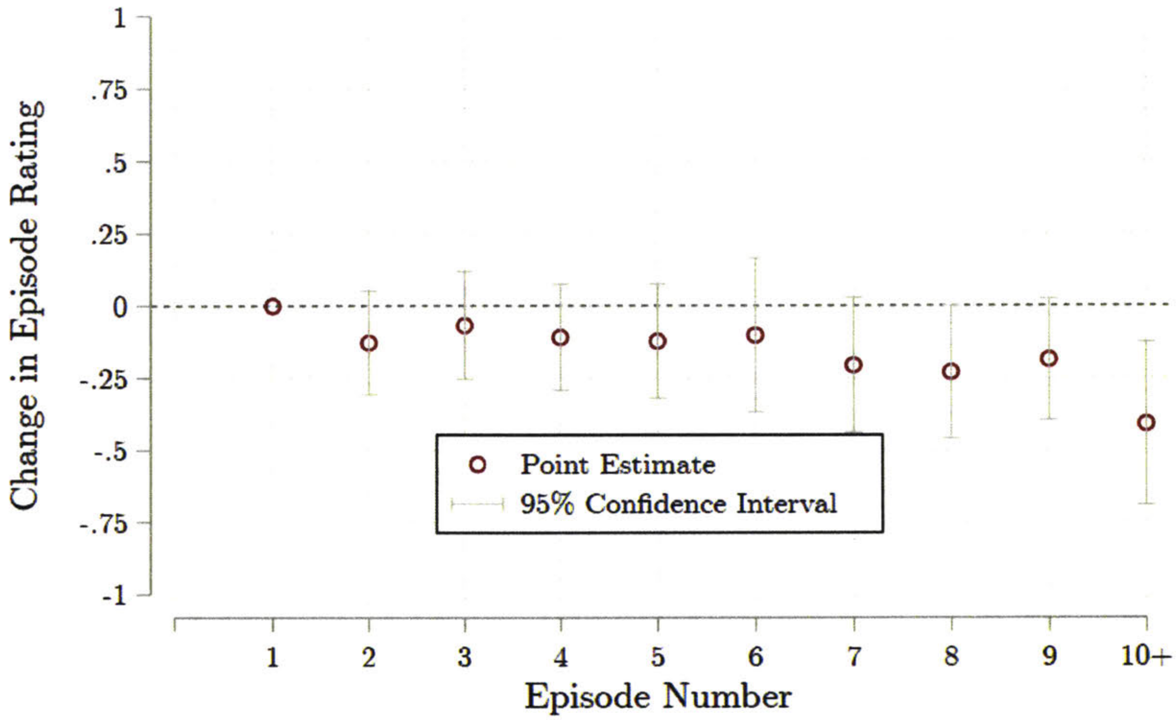
Point estimates and confidence intervals for the share of ordered shows with a creator that had a previous show which won a major Emmy or Golden Globe award. Big four networks are ABC, NBC, CBS and Fox. Observations restricted to US incumbent networks from 2008 to 2017.

Figure 6. Portfolio Relationship between Straight to Series and IMDb Ratings



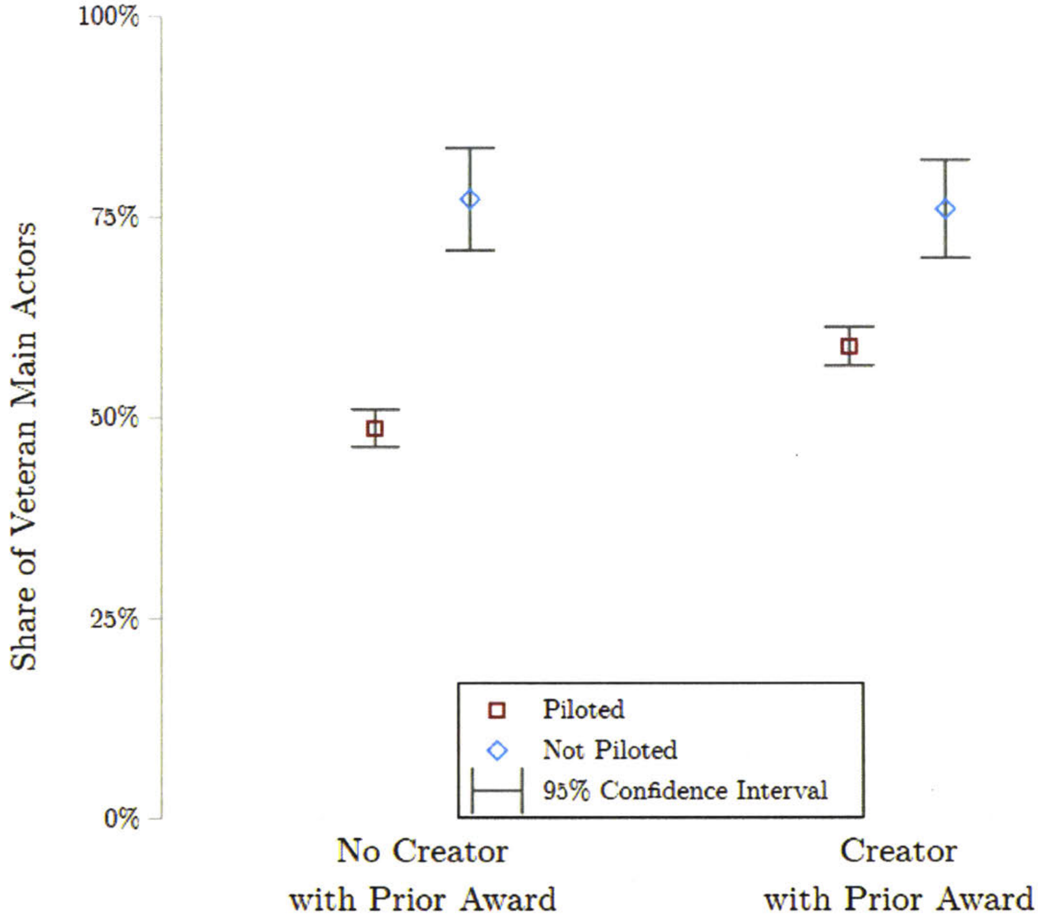
Plots of each network's yearly new show average IMDb rating against share ordered straight to series. Big four networks are ABC, NBC, CBS and Fox. Plot points and fitted value line use US incumbent networks for 2008 to 2017.

Figure 7. By Episode Relative Change in IMDb Ratings for Piloted Shows



Difference between piloted and straight to series shows episode with first episode set as baseline. Observations of ordered shows between 2014 and 2017 on incumbent networks weighted by inverse propensity score to match straight to series and piloted shows over area of common support, clustered at the show level.

Figure 8. Relationship between Veteran Actors, Piloting and Creators w/ Awards



Piloted shows on incumbent networks from 2008 to 2013. Shows having creator with prior award have at least one creator with a previous show that one a major Emmy or Golden Globe award.

2.10.2 Tables

Table 1. Summary Statistics Before and After Netflix's Entry

	2008-13	2014-17	Overall
Piloted or ordered shows	620	502	1122
Average per year	103.3	125.5	112.2
Piloted shows	599	425	1024
Average per year	99.8	106.3	102.4
Piloted shows ordered to series	317	200	517
Average per year	52.8	50.0	51.7
Straight to series ordered shows	21	77	98
Average per year	3.5	19.3	9.8
Is a drama	58%	57%	57%
Hour long show	59%	60%	59%
Has creator with prior award	47%	49%	48%
Ordered shows	338	277	615
Average per year	56.3	69.3	61.5
Mean IMDb first season rating	7.7	7.8	7.8
Standard deviation in IMDb first season rating	0.7	0.8	0.7
Renewed past initial order	64%	56%	60%

Table 2. Correlation between IMDb Rating and Other Outcome Measures

	(1) Renewed or Extended	(2) Nielsen Top 30 Show
1st Season IMDb Rating	0.181***	0.0579***
	[0.0262]	[0.0167]
Constant	-0.802***	-0.356***
	[0.206]	[0.126]
Network Portfolios (N)	522	309
Adj. R-Squared	0.122	0.0760

OLS estimation with robust standard errors. Star levels: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$. Data is restricted to shows from 2008 to 2016 on incumbent networks, since renewal and Nielsen data missing for 2017. Includes year and network fixed effects. 1st Season Rating is the average episode IMDb rating for a show's initial series order. Renewed or Extended is an indicator for whether a show was either renewed for another season or extended past its original order. Nielsen Top 30 Show indicates whether the show broke into the Nielsen Top 30 highest viewed programs for the season. Renewal correlation is for all incumbent networks while Nielsen correlation is just for the big four networks since they are the only networks with enough viewership to potentially enter the top 30.

Table 3. Estimators of Treatment of Piloting on IMDb Ratings

	OLS		Matching	
	(1)	(2)	(3)	(4)
	1st Season Rating	1st Season Rating	1st Season Rating	1st Season Rating
Piloted	-0.229**	-0.169*	-0.188	-0.232*
	[0.0950]	[0.0953]	[0.131]	[0.116]
Creator w/ Award		0.0630		0.287**
		[0.0917]		[0.129]
Constant	7.989***	7.658***	8.024***	7.667***
	[0.0769]	[0.161]	[0.140]	[0.226]
Year FE		X		X
Network FE		X		X
Bias Adj. Piloted		1.243		-0.340
Shows (N)	263	263	144	144
Deg. of Freedom	51	51	43	43
Adj. R-Squared	0.016	0.115	0.009	0.214

Robust standard errors in brackets. Star levels: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$. Clustered by network * year. Data is restricted to shows on incumbent networks from 2014 to 2017. Matching columns use an inverse propensity score based weighting over a region of common support to match similar straight to seriers and piloted shows. Bias adjusted piloted estimate assumes unobservables relate to piloted similarly to FE and Creator w/ Award observables as in Oster (2016).

Table 4. Logistic Prediction of Straight to Series Decision

	(1) Straight to Series	(2) Straight to Series
Funded independently of major studios	2.611***	3.137***
	[0.630]	[0.753]
Funded by broadcasting network	0.408	0.272
	[0.362]	[0.411]
Creator with Prior Show on Same Network	0.236	0.192
	[0.346]	[0.381]
Filmed outside Los Angeles	-0.0360	-0.609
	[0.627]	[0.723]
Genre: SciFi	0.985*	1.194**
	[0.520]	[0.557]
Genre: Comedy	-0.495	-0.867*
	[0.456]	[0.506]
Genre: Crime	0.905*	0.922*
	[0.482]	[0.513]
Genre: Supernatural	0.670	0.999*
	[0.513]	[0.549]
Genre: Thriller	0.649*	0.577
	[0.387]	[0.466]
Genre: Action	-0.941	-0.989
	[1.086]	[1.049]
Year FE		X
Network FE		X
Shows (N)	460	438
Pseudo R^2	0.158	0.243

Robust standard errors in brackets. Star levels: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$.
Logistic estimation of straight to series used in IPW matching. Data is restricted to shows on incumbent networks from 2014 to 2017.

Table 5. Result of Matching on Observables for Incumbent Shows 2014-2017

All Observations

	Piloted		Straight to Series		
	N	Mean	N	Mean	Diff
Funded independently of major studios	426	0.09	77	0.38	0.29***
Funded by broadcasting network	426	0.62	77	0.53	-0.08
Creator with Prior Show on Network	426	0.73	77	0.65	-0.08
Filmed outside Los Angeles	426	0.52	77	0.71	0.19***
Genre: Science Fiction	421	0.04	75	0.13	0.09***
Genre: Comedy	421	0.47	75	0.20	-0.27***
Genre: Crime	421	0.05	75	0.12	0.07**
Genre: Supernatural	394	0.05	66	0.11	0.06*
Genre: Thriller	421	0.09	75	0.19	0.09**
Genre: Action	421	0.04	75	0.03	-0.02

Common Support

	Piloted		Straight to Series		
	N	Mean	N	Mean	Diff
Funded independently of major studios	202	0.12	38	0.21	0.09
Funded by broadcasting network	202	0.61	38	0.66	0.05
Creator with Prior Show on Network	202	0.74	38	0.68	-0.06
Filmed outside Los Angeles	202	0.74	38	0.74	0.00
Genre: Science Fiction	202	0.06	38	0.08	0.02
Genre: Comedy	202	0.28	38	0.18	-0.09
Genre: Crime	202	0.09	38	0.11	0.02
Genre: Supernatural	202	0.08	38	0.05	-0.03
Genre: Thriller	202	0.15	38	0.21	0.06
Genre: Action	202	0.04	38	0.03	-0.02

Star levels for t-tests: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$

Table 6. Heterogeneity in Treatment Effect of Piloting by Creators w/ Awards

	(1) 1st Season Rating	(2) 1st Season Rating	(3) 1st Season Rating
Piloted	-0.231**	-0.232*	0.221
	[0.114]	[0.116]	[0.169]
Creator w/ Award		0.287**	0.596***
		[0.129]	[0.165]
Piloted * Creator w/ Award			-0.773***
			[0.268]
Constant	7.837***	7.667***	7.435***
	[0.197]	[0.226]	[0.269]
Shows (N)	144	144	144
Deg. of Freedom	43	43	43
Adj. R-Squared	0.188	0.214	0.279

Robust standard errors in brackets. Star levels: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$. Includes year and network FE. Clustered by network * year. Data is restricted to shows on incumbent networks from 2014 to 2017. All estimates use an inverse propensity score based weighting over a region of common support to match similar straight to seriers and piloted shows as well as account for attrition in piloted shows not ordered to series.

Table 7. Effect of straight to series on network's portfolio of new shows

	(1) Average IMDB rating	(2) Average IMDB rating
Share Straight to Series	-0.147	-0.173
	[0.156]	[0.177]
Share StS * Post 2013	0.622	0.543
	[0.391]	[0.429]
Share StS * Big Four		-1.079
		[0.750]
Big Four * Post 2013		-0.318
		[0.251]
Share StS * Big Four * Post 2013		2.773**
		[1.016]
Constant	7.544***	7.559***
	[0.0831]	[0.0862]
Network Portfolios (N)	131	131
Deg. of Freedom	14	14
Adj. R-Squared	0.153	0.171

Robust standard errors in brackets. Star levels: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$. Observations are of each network's yearly porfolio of new shows. Includes year fixed effects. Fixed effect estimator used clustered at network level. Data is restricted to shows from 2008 to 2017 on incumbent networks.

2.11 Appendix

2.11.1 Mathematical proofs

Theorem 1

Consider the solution to the decision problem:

$$\max_{\lambda \in \{0,1\}} \lambda(\Pr(y(\theta) > C)\, E[y(\theta) + (V_B - V_D) - C | y(\theta) > C] - e) + (1 - \lambda)(\theta - C)$$

With the definition of signal ordering as in D2 and assume A1. Then by the monotone maximization theorem, experimentation ($\lambda = 1$) is increasing in y, V_B,$- V_D$, and $-e$:

- The space of possible choices, $\{0,1\}$, is trivially a lattice.
- The space of possible parameter values y, C, V_B, V_D e and $\theta_i \in \mathcal{R} \times \mathcal{R}^+ \cup \{0\} \times \mathcal{R}^+ \cup \{0\} \times \mathcal{R}^+ \cup \{0\} \times \mathcal{R}^+ \cup \{0\} \times \mathcal{R}^+ \cup \{0\}$ is a poset under the usual $\mathcal{R}^N$ order.
- The set of possible choices is trivially non-decreasing in parameter values, as it is always $\{0,1\}$.
- The objective function is trivially supermodular in the set of choice variables, given there is a single choice variable.
- The objective function displays increasing differences in λ, y, V_B, $-V_D$ and $-e$.

Theorem 2

As in Theorem 1, we need to show increasing differences between $-\theta$ and λ. This amounts to showing:

$$\Pr(y(\theta') > C)\, E[y(\theta') + V_B - V_D - C | y(\theta') > C] - \theta' + C$$
$$\leq \Pr(y(\theta) > C)\, E[y(\theta') + V_B - V_D - C | y(\theta) > C] - \theta' + C$$

Which simplifies to:

$$\Pr(y(\theta') > C)\, E[y(\theta') + V_B - V_D - C | y(\theta') > C] - \theta' + C$$
$$\leq \Pr(y(\theta) > C)\, E[y(\theta) + V_B - V_D - C | y(\theta) > C] - \theta + C$$

A1 gives us that $V_B - V_D - C$ is negative and therefore A2 implies

$$\Pr(y(\theta') > C)\,(V_B - V_D - C) \leq \Pr(y(\theta) > C)\,(V_B - V_D - C)$$

The remainder of the inequality follows from A3.

The other variables do not have the correct increasing difference structure to make a more general comparative statics stamen as in Theorem 1. For example, increasing differences exist between θ and V_B under A3 since:

$$\Pr(y(\theta') > C)\,(V'_B - V_B) \geq \Pr(y(\theta) > C)\,(V'_B - V_B)$$

Theorem 3

$A >_{LR} B$ implies $A|A \in S >_{FOSD} B|B \in S$ given A and B are both drawn from any measurable subset S of Θ, see 1.C.6 in Shaked and Shanthikumar (2007), from here on referred to as SS. Since buy A3 $\theta' > \theta$, $y(\theta') >_{FOSD} y(\theta)$, Theroem 1.A.6 in SS implies $y|A >_{FOSD} y|B$. This means the $\Pr(y(\theta) > C|\theta \in A) > \Pr(y(\theta) > C|\theta \in B)$, the desired result.

Lemma 1

First, A1-A3 imply the observed difference in payoffs between experimentation and non-experimentation are decreasing in θ. For $\theta' > \theta$ A2 states:

$$\begin{aligned}\Pr(y(\theta') > C)\, E[y(\theta') + (V_B - V_D) - C|y(\theta') > C] - \theta' + C \\ \leq \Pr(y(\theta) > C)\, E[y(\theta) + (V_B - V_D) - C|y(\theta) > C] - \theta + C\end{aligned}$$

A1 and A3 allows us to remove the C terms.

$$\begin{aligned}\Pr(y(\theta') > C)\, E[y(\theta') + (V_B - V_D)|y(\theta') > C] - \theta' \\ \leq \Pr(y(\theta) > C)\, E[y(\theta) + (V_B - V_D)|y(\theta) > C] - \theta\end{aligned}$$

Since by A3 $\Pr(y(\theta) > C) \leq \Pr(y(\theta') > C)$

$$\begin{aligned}\Pr(y(\theta') > C)\, E[y(\theta') + (V_B - V_D)|y(\theta') > C] - \theta' \\ \leq \Pr(y(\theta') > C)\, E[y(\theta) + (V_B - V_D)|y(\theta) > C] - \theta\end{aligned}$$

Then splitting θ terms

$$\begin{aligned}\Pr(y(\theta') > C)\, [E[y(\theta') + (V_B - V_D)|y(\theta') > C] - \theta'] \\ - (1 - \Pr(y(\theta') > C))(\theta' - \theta) \\ \leq \Pr(y(\theta') > C)[E[y(\theta) + (V_B - V_D)|y(\theta) > C] - \theta]\end{aligned}$$

Since $-(1 - \Pr(y(\theta') > C))(\theta' - \theta)$ is negative and $\Pr(y(\theta') > C)$ is positive we get the desired result

$$E[y(\theta') + (V_B - V_D)|y(\theta') > C] - \theta' \leq E[y(\theta) + (V_B - V_D)|y(\theta) > C] - \theta$$

Theorem 4

Suppose A and B are two distributions with $A >_{LR} B$. Define $\psi(\theta) \equiv E[y(\theta) + (V_B - V_D)|y(\theta) > C] - \theta$, the difference between experimented and non-experimented outcomes. Let S be the range of Θ used to compare experimentation versus non-experimentation outcomes. Since $A >_{LR} B$ we know $A|S \in \Theta >_{FOSD} B|S \in \Theta$ for any subset S of Θ. By Lemma 1, $\psi(\theta)$ is decreasing. This implies our required result

$$E_A[\psi(\theta)|S] \leq E_B[\psi(\theta)|S]$$

Theorem 5

Let project i be a project moved from experimentation to non-experimented production. Suppose afterwards, there are m projects still developed through experimentation. Let A be the average value of these project. For the non-experimented projects, let there be n projects other than i, with an average value B. The portfolio of shows then worsens if

$$\frac{1}{m+n+1}[mA + nB + E[y(\theta_i) + (V_B - V_D)|y(\theta_i) > C]] > \frac{1}{m+n+1}[mA + nB + \theta_i]$$

This simplifies to $E[y(\theta_i)|y(\theta_i) > C] + (V_B - V_D) > \theta_i$. Since $E[y(\theta_i)] = \theta_i$, this will be true unless $V_B - V_D$ is sufficiently negative. Moving a project from experimentation to non-experimentation production can only improve outcomes when the drawbacks from experimentation outweigh the benefits from experimentation: when commitment has value.

In the case where the experimented project did not pass the experimentation phase, outcomes would worsen whenever:

$$\frac{1}{m+n+1}[mA + nB + \theta_i|\theta_i < C] < \frac{1}{m+n}[mA + nB]$$

Since a project would only be in set A or B if the expected outcome was greater than cost C, we know:

$$\frac{1}{m+n+1}[mA + nB + C] \leq \frac{1}{m+n}[mA + nB]$$

Since the specific project θ_i did not make the cut into experimentations as post experiment the posterior belief was less than C, we can complete the proof:

$$\frac{1}{m+n+1}[mA + nB + \theta_i|\theta_i < C] < \frac{1}{m+n+1}[mA + nB + C] \leq \frac{1}{m+n}[mA + nB]$$

Theorem 6

First, I argue the entry of a network has no bargaining power effect on the decision to experiment under standard cooperative bargaining models. I model bargaining through Shapely value, which is appropriate for bi-lateral relationships (Fontenay and Gans, 2014). Based on interviews with show creators, writers approach the networks serially with an idea. Writers rank networks based on fit with their vision and support for their show conditional on the other shows in the network's portfolio. The number of networks in their consideration set is usually small; perhaps two or three networks are a strong fit for a particular show in a particular season. Writers then approach each network serially. When negotiations fail with a network, writers need to tweak their idea for the next network; an aggressive cop drama originally pitched to a cable network like HBO would be moderated for a broadcast network like ABC. Bi-lateral bargaining between agents in a graph is a reasonable model for this type of linear negotiation process.

The value allocated to each player i in a collation game is a weighted sum of

$$v(S \cup \{i\}) - v(S)$$

where S is a set of players not including i and $v(S)$ is the highest total sum of payoffs that can be generated by S through cooperation. Because there is only one bilateral contract at a time that can exist between the writer and a network, we can write $v(S \cup \{i\})$ as

$$\max[v(S), \max_{\lambda^i \in \{1,0\}} (\lambda^i \hat{V}_p^i + (1 - \lambda^n)\hat{V}_s^i)]$$

Network i selects whether or not to pilot, λ^i. The expected payoff from piloting for network i is $\hat{V}_p^i$ and the payoff from deciding to skip the pilot is $\hat{V}_s^i$.

A subgame perfect equilibirum is reached if all networks always pick the higher of $\hat{V}_p^i$ or $\hat{V}_s^i$. When one or both of $\hat{V}_p^i$ and $\hat{V}_s^i$ is above $v(S)$, picking the larger one results in the highest payoff. If both are below $v(S)$, network i received no value from the term so picking the larger is still weakly preferred. So regardless of the number of other networks or the payoffs available to those networks, firm i will only consider whether $\hat{V}_p^i$ or $\hat{V}_s^i$ is larger. If the network is picking piloting over skipping the pilot prior to entry, it suggests the network views the piloting payoff as higher. Furthermore, the network should continue picking the piloting after entry since it remains optimal to do so if the coalition value is not affected by entry.

Next, what about if bargaining in television deviates from the standard assumptions, by for example exhibiting uncompensated differentials? Let V_E be expected joint payoff from experimentation, inclusive of costs of production and V_C be the joint payoff from commitment. Let $V_E > V_C$, so that experimentation results in a larger overall payoff. The payoff is split with $1 - \alpha$ going to the network and α going to the show creator. To model uncompensated differentials, let κ represent a private cost to experimentation born by the show creator that is not included in V_E. Before Netflix's entry, all the networks prefer to experiment since $V_E > V_C$. Each creator's outside option is the payoff from experimentation on another network, so a subgame perfect equilibirum is for the networks to always experiment by funding pilots.

Once Netflix enters and only commits, the decision calculation of the incumbent networks changes to reflect that the creator's outside option may include a payoff V_C that crucially does not cause them to incur private cost κ. Now when experimenting, the incumbent networks need to provide a larger share of payoff V_E to the creator when experimenting, represented by Δ.

$$\max_{\lambda\in\{0,1\},\Delta\in[0,1-\alpha]} \lambda(1-\alpha-\Delta)V_E + (1-\lambda)(1-\alpha-\Delta)V_C$$

$$s.t.\ \lambda\big((\alpha+\Delta)V_E-\kappa\big) + (1-\lambda)(\alpha+\Delta)V_C \geq \alpha V_C$$

When the network chooses to commit, Netflix as an outside option is not superior from the perspective of the creator so Δ can be 0. However, under experimentation, Δ may need to be positive so that its rational for the creator to prefer the network over Netflix. In the case when the optimal $\Delta^* \in [0,1-\alpha]$, this means the private cost κ gets factored into the decision making of the network, so the optimization problem becomes:

$$\max_{\lambda\in\{0,1\}} \lambda(V_E-\kappa) + (1-\lambda)V_C$$

This does not solve the identification issue since variation in bargaining power is not affected the network's decision. However, when α is high its possible that $\Delta^* > 1-\alpha$, making it not possible for the network to sufficiently compensate a creator to experiment: the network would be forced to commit instead. This would lead to shows with the same expected payoff to have different production processes due to the level of creator bargaining alone.

2.11.2 Mechanism Models

Commitment improves later episodes relative to first episode

Consider an example of a mechanism that would improve the final stage relative to the interim stage, where experimentation distorts effort towards passing the interim stage and commitment removes that distortion. A basic multitasking model with two actions can represent the level of effort placed in the first episode a_1 and subsequent episodes a_2. The payoff to the show's creator the total quality of the show, here equal to the sum of effort $a_1 + a_2$, minus the cost of that effort, $a_1^3 + a_2^3$; a cubic cost function is used rather than the traditional quadratic to ensure cost grows fast enough for later modifications to have finite solutions.

$$\max_{a_1,a_2 \in \mathcal{R}^+} a_1 + a_2 - (a_1^3 + a_2^3)$$

The solution to this model is symmetric with equal levels of effort applied to a_1 and a_2. This symmetry also holds if rather than receiving the payoff $a_1 + a_2$ with certainty, the show creator only gets that payoff with probability $p \in (0,1)$; the network can decide not to broadcast the show.

$$\max_{a_1,a_2 \in \mathcal{R}^+} p * (a_1 + a_2) - (a_1^3 + a_2^3)$$

Effort levels become asymmetric when the pilot influences the probability of receiving the payoff. Assuming for the moment that $a_1 \in [0,1]$, optimizing

$$\max_{a_1,a_2 \in \mathcal{R}^+} a_1 * (a_1 + a_2) - (a_1^3 + a_2^3)$$

leads to a higher level of effort a_1 than a_2 as applying effort to the first episode not only improves the payoff $a_1 + a_2$ but also increases the probability of receiving that payoff.

Combining the two models by letting $\gamma \in [0,1]$ represent the relative importance of pilot effort to influence passing the pilot stage, the creator optimizes

$$\max_{a_1,a_2 \in \mathcal{R}^+} (p * (1-\gamma) + a_1\gamma) * (a_1 + a_2) - (a_1^3 + a_2^3)$$

At $\gamma = 0$, effort is balanced between a_1 and a_2 but a gap grows with $a_1 > a_2$ as γ increases. When $p = 1$ and $\gamma = 0$, we can think of this model as one of commitment, with a distortion of effort towards the pilot as γ increases to represent the use of experimentation. Of course, in my setting we would naturally expect greater effort exerted in a_2 versus a_1 due to the sheer quantity of later episodes versus the first one, but we can think of the above model as representing a normalized effort that is balanced when the first episode does not influence the probability of receiving the payoff.

Commitment affects all episodes equally

In a planning model, one can imagine planning time a and an execution time b as a share of total time spent in development.

$$\max_{a,b \in [0,1]} a + b - a^2 - b^2 \quad s.t.\ a + b = 1$$

The optimal allocation a^*, b^* of time between these two tasks results in the best possible show. Suppose piloting a show constricts the amount of time that can be spent on planning before production to c, since the first episode must be produced according to the network's pilot schedule.

$$\max_{a,b\in[0,1]} a + b - a^2 - b^2 \quad s.t.\ a + b = 1, a \leq c$$

If $a^* > c$, then piloting results in the suboptimal amount of time in planning. Switching to commitment would improve the quality of the show by benefiting all episodes.

A similar pattern of improvement would occur for many of the theories relating to commitment. All episodes would improve if the cast and crew of televisions shows preferred the stability of a straight to series show and were uncompensated for the disutility of piloting. Changes in the risk profile of projects principles or agents would again affect all episodes equally.

Commitment improves first episode relative to later episodes

A task model could have the first episode improve relative to the rest of the season under straight to series production if the effort was sequentially decided and creators felt effort did not influence passing the pilot phase, for example if creators believed network executives lacked the ability to judge pilot quality. Then the first episode gets lower effort under piloting because it may never get broadcast while subsequent episodes get higher effort. With a straight to series order, all episodes get this higher level of effort since creators know their work will be seen by an audience, so it's the first episode that benefits the most from the switch from real options to commitment.

$$\max_{a_1,a_2\in\mathcal{R}^+} pa_1 + a_2 - a_1^2 - a_2^2$$

A model where piloting helps creators improve the quality of the show could have the similar comparative statics if an overall improvement in all episodes from straight to series is also present. Suppose there is some characteristic about a show that is the creator's control which determines the quality of the show. In order for a show to be viewed as high quality, the level of that characteristic should match the desires of the audience, for example a drama show might need to have some action elements but too much action could detract from the plot and hurt the show's ratings.

Let θ be the level of the characteristic preferred by the audience, distributed F_θ. The network knows θ but the creator does not. Each creator has a belief γ about what the audience would like, distributed F_γ. Suppose there are two outputs, the first episode and the rest of the season, and the show's value is the sum of quadratic loss from the difference between audience taste and creator beliefs for those two outputs:

$$E[-(\gamma-\theta)^2]+E[-(\gamma-\theta)^2]$$

Now allow the network to adjust γ to equal θ by piloting with cost e. The network then maximizes whether to pilot $\lambda = 1$ based on the trade-off between piloting cost and the expected loss from the rest of the season:

$$\max_{\lambda \in (0,1)} \lambda(E[-(\gamma-\theta)^2|\theta]-e)+(1-\lambda)\,(E[-(\gamma-\theta)^2|\theta]+E[-(\gamma-\theta)^2|\theta])$$

The network would decide to pilot when $E[(\gamma-\theta)^2|\theta] > e$.

In this model if the distributions are non-degenerate, the quality the episodes beyond the first will be better for piloted shows; non-piloted shows have a quality loss of $E[(\gamma-\theta)^2|\theta]$ while piloted shows have no quality loss. More interestingly, the reverse is true for the first episode. Since piloted shows have $E[(\gamma-\theta)^2|\theta] > e$, non-piloted shows must have $E[(\gamma-\theta)^2|\theta] < e$; $E[(\gamma-\theta)^2|\theta,\lambda = 1] > E[(\gamma-\theta)^2|\theta,\lambda = 0]$ so the first episode of non-piloted shows are better. Experimentation occurs on the relatively "bad" shows so the first episode of the piloted shows is worse. Experimentation improves those "bad" shows, so the subsequent episodes of the piloted shows are better. This paired with a benefit of commitment such as improved planning or higher quality workers could show an overall positive effect from straight to series production driven primarily by an improvement in the quality of the first episode.

2.11.3 Uncompensated Bargaining Model

Beyond showing that pilots were unnecessary to create a critically acclaimed television show, there is evidence that Netflix, as an entrant committed to skipping pilots, changed the bargaining relationship between incumbent networks and content creators (Adalian, 2013). However, a shift in bargaining power away from the networks doesn't necessarily mean networks are going to order shows straight to series more often. The appendix contains a more formal mathematical argument, but under a standard cooperative bargaining model everyone is always better off making decisions that maximize the total value generated by the group. Before Netflix's entry, the networks would only be selecting piloting if the expected payoff from piloting is higher than from straight to series. Because piloting generates more value and the network's decision does not affect the fraction of value allocated to itself, the networks should always prefer piloting. Netflix's entry just affects the share of surplus shared with show creators and the networks will always be better off with a share of the larger piloting surplus than the smaller straight to series surplus. Therefore, a change in bargaining power after Netflix's entry would not be able to explain an increase in straight to series production in a standard cooperative bargaining model.

But a standard cooperative bargaining model may not apply in this setting. Based on interviews with content creators, many prefer a straight to series order over a pilot order. A straight to series order ensures their work will be seen by the public whereas "if you do a pilot and it doesn't get picked up, you don't have anything", to quote *Sex and the City* director Allison Anders (Bunn, 2002). This preference for straight to series stands in contrast to the actions of the networks, which overwhelmingly chose to pilot shows. An equilibirum could have existed where all the incumbent networks piloted and left the straight to series preference of creators uncompensated (England, Farkas, Kilbourne and Dou, 1988). By committing to straight to series production, Netflix would have changed this equilibirum upon entry, increasing the rate of straight to series orders by incumbents. The networks would have been forced to compensate creators for their disutility from piloting since the networks would lose creators to Netflix without this compensation. If the piloting payoff on a particular show prior to Netflix's entry is not too much higher than the straight to series payoff, adding compensation to creators for piloting can make the network prefer the straight to series payoff after Netflix's entry.

> Assumption 4: Bi-literal bargaining occurs between project agents and the principles that fund the projects and that the shared payoff from experimentation is higher than from non-experimentation.

This incremental shift in payoff is not the only mechanism that could increase straight to series orders in such a model. When a show's creator has high bargaining power, the network is already giving a large share of show rents to the creator. They may be constrained from giving an even larger share once Netflix enters to compensate for piloting; the creator's show might be considered a loss leader to draw in larger audiences for the other shows on the network. In this case, the network would switch to straight to series after Netflix's entry because simply because compensating for piloting is not possible.

This intuition suggests bargaining could be an alternative explanation to the decision process modeled in Equation 2 for observed network behavior. Depending on how bargaining power correlates with a show's quality, θ_i, Netflix's entry could have various effects on outcomes. Specifically, no correlation would explain a shift towards straight to series production without a corresponding drop in straight to series show quality.

Theorem 6: Let there be a private cost to agents from experimentation that is excluded from the shared payoff calculation. Then under A4 there exists a subgame perfect equilibirum where all projects are experimented on. In addition, if one principle deviates and never experiments, agent bargaining power will determine which projects get experimented on for the other principles.

3. Vertical Integration and the Direction of Innovation

Ankur Chavda

March 2019

Abstract

The decision to vertically integrate is an important optimization decision made by firms. However, this decision not only affects the firm itself, it also influences the firm's industry as the relationships between firms is changed. This paper is an empirical study of how vertical integration impacts an industry, specifically the set of new products developed each year: the direction of innovation. Television shows can either be financed independently of the show's broadcast network or partially funded by the show's broadcast network; this variation in funding changes the owner of the television show and is therefore a form of vertical integration. Using a regulatory shock that restricted the networks' incentives to fund television shows, I find a drop in vertical integration contemporaneous with a shift away from dramas and an overall decrease in the introduction of new show genre combinations. My results demonstrate how organizational form affects an industry's rate and direction of innovation.

3.1 Introduction

The theory of the firm addresses the make-or-buy decision from the point of view of the firm: given the tradeoffs to vertical integration, what firm boundary optimizes profit and what types of contracts should mediate that boundary (Coase, 1937)? However, theory suggests this integration decision should also affect the innovative output of the firm, by for example through a property rights (Aghion and Tirole, 1994), incentive (Holmstrom, 1989) or adaptation (Manso, 2011) mechanisms. Although the type of integration selected has been shown to affect firm performance (Novak and Stern, 2008; Forbes and Lederman 2010), integration's affect the range or rate of new products brought to market has been missing in the empirical literature.

Two challenges exist in testing for a causal relationship between the decision to vertically integrate and resultant level of innovation. First, in general random or quasi-random variation in contracting is rare; exogenous technological innovations are often used as a shock to test theories of vertical integration (Baker and Hubbard, 2003; Rawley and Simcoe, 2013), the reverse of the desired causal structure. Second, empirical estimators have asymptotic properties that require observations drawn from an identical distribution yet innovate products are by definition different from each other. Contracting variation needs to be found in a setting where innovation occurs but is constrained so observations are comparable enough for statistical analysis.

To meet these challenges, I use the production of television shows in the United States as my empirical setting. Each year the industry produces dozens of new shows that vary by characteristics such as genre. I consider a show innovative if the show's genre is novel compared to past shows. Funding for show development can either come from the networks that broadcast the show or from independent sources. When the network that broadcasts a show is also at least a partial owner of the show, I consider the show vertical integrated. This type of vertical integration was common in television's early history until 1970, when the Federal Communications Commission ("FCC") prohibited the major US networks from holding ownership stakes in any of television shows broadcast during the evening hours from 7PM to 11PM ("primetime"). I exploit this shock to investigate how vertical integration affects the genre of shows produced, using Internet Movie Database (IMDb) data complimented by qualitative evidence from interviews with television executives.

Using a difference and difference estimator, I have three main empirical findings. First, I demonstrate the broadcast restrictions instituted by the FCC changed the financing of television; the networks did not simply continue funding show development and sell off ownership for any shows they broadcast. Second, this change in financing coincided with a shift in the types of shows created by the industry; dramas were far more common after the FCC rule change than before. Third, the rate of new show genres tried out by the industry dropped as well, suggestive of a decrease in innovation within the industry.

This paper's empirical results provide two contributions to the literatures on vertical integration and innovation. First, I show the consequences of vertical integration decisions are not limited to the profitability of firms. They have a real impact on the types of products that get created in an industry with long term consequences for innovation. Second, I illustrate an underappreciated driver of innovation in an economy: the contracting relationship between firms.

The rest of the is paper is organized as follows. Section 2 provides an overview of the previous related literature on vertical integration and describe this paper's contribution. Section 3 describes television show production and how it mirrors innovation in other industries. Section 4 details my dataset and measures used in my analysis. Section 5 outlines my empirical approach. Section 6 presents my results. Section 7 concludes with a discussion of how this paper's results connect with the broader conversation on the direction of innovation.

3.2 Television Show Production

The production of television shows involves three entities: a supplier, a buyer and a financier. Production companies are the suppliers responsible for the nuts and bolts of creating a show: hiring cast and crew, leasing set space and securing any necessary IP rights. The television networks are the buyers of the shows. They pay a license fee for the right to broadcast a show twice. This license fee does not offset all the costs of production: television shows are deficit financed. Hence the need for a financier. The financier hopes to profit from a show in later airings; the television networks may pay for additional broadcast rights, a process known as syndication.

A network's investment in a show is staged, first by taking an option on the idea, second by ordering a script, then the first episode, called the pilot, as the idea progresses in development. Finally, the network may choose to order for a show's first season, typically somewhere between 10 and 24 episodes. Depending on the success of the show, the network will order additional seasons of the show. When the total number of episodes for a show approaches approximately 100 episodes, syndication becomes possible. Syndicated shows can be highly profitable. Financiers expect to lose money on the average show and recoup those loses on the few successful shows that make it into syndication.

Sometimes one of the financiers for a show is also the buyer, the form of vertical integration studied in this paper. The parent company of the networks have a financing arm that can fund the shows broadcast on the company's network. The network refers to this financing arm as its "sister studio".

3.2.1 Vertical Integration

Show financing by a network's sister studio has the potential to mitigate two coordination challenges of television show production: the selection of show ideas and the show specific investments made by the networks.

First, vertical integration affects the information flow between the network and the production company, albeit in a different manner than outlined by Arrow (1975). When the networks are shopping for new show ideas, they often have specific needs for their next season's schedule, e.g. a police drama, but are hesitant to share that information with production companies. Networks are concerned that production companies will mask their true creative intent, e.g. a police comedy, in order to appeal to the network's need. As one television executive I interviewed stated, this leads to a "Tetris"-type negotiation problem: even if a production company has the right show idea to fit into the network's schedule needs, it may pitch on of its other, mis-matched ideas to the network. Integration is one solution to this problem; the network shares its scheduling needs with its sister studio and production companies share their show ideas with the same entity if that sister studio agrees to fund those show ideas, effectively turning the sister studio into an information broker.

Second, the networks better internalize the value of the television show under integration. Without integration, the networks receive advertising revenue at the cost of the licensing fee for broadcasting the show. The level of marketing invested in the show and the decision to renew the show for another season is made solely on the advertising profitability. Under integration, the network's sister studio receives a portion of the syndication fees generated by the show conditional on enough episodes being ordered by the network. The decision maker's ownership of the underlying asset could improve joint surplus.

As a downside, integration could lead to worse quality television being created. During an interview, a network executive described how although piloting decision were made at the network level, the choice of which pilots were ordered to series balanced the needs of the network with its sister studio. Although the best pilots were still made into shows, this meant that sister studio funded shows were favored when deciding between more uncertain pilots. Also, from the production company's standpoint, a tradeoff exists when a sister studio is used for financing: the odds of getting a show on that sister studio's network improve but are lowered for the other networks.

3.2.2 Innovation

Television shows as a setting shares many characteristics with other innovative industries. As in many creative and innovative industries, outcome uncertainty (Arrow, 1962) is ever present in the production of television shows. William Goldman's famous "Nobody knows anything" quote from the movie industry applies: 65% of television shows on the major networks get cancelled after their first season (Ocasio, 2012), a clear indication of failure when it takes four seasons for a show's financiers to make a return on their investment (Bunn, 2002). This failure rate exists despite a winnowing process that results in only a fraction of ideas for new shows getting aired on television. It is also not yet clear that the big data capability of entrants like Netflix and Amazon has reduced this uncertainty. To quote Netflix's Chief Content Officer: "the data just tells you what happened in the past. It doesn't tell you anything that will happen in the future" (Adalian, 2013). Amazon is shifting away from its original data driven decision approach because as one Amazon executive puts it: "we're getting chewed up" (Fritz and Flint, 2017).

This high failure rate leads to a similar hit driven, staged financing model as seen in venture capital's funding of technology innovations. Kerr, Nanda and Rhodes-Kropf (2014) describe the staged financing used in the venture capital industry as a kind of experimentation on new, innovative ideas. Television shows are also staged: the networks will first order a script, then pilot episode, then first season, and finally later seasons depending on the quality of the product at each stage.

Finally, television shows are cumulative (Scotchmer, 1991). The first every sci-fi family drama, 1963's *My Favorite Martian*, differed from anything before it and was a precursor to later shows like *Mork & Mindy*, *ALF* and *3rd Rock from the Sun. The*

Sopranos mainstreamed drama shows with season long story arcs, leading to later shows like *Breaking Bad.*

3.2.3 1970 FCC Rule Change

The 1969 election of Richard Nixon as US president led to a new chairman at the FCC, Dean Burch. Both the White House and the FCC's new leadership was concerned with the concentrated control of media held by the big three broadcast networks (Covington 1994). The following year the FCC enacted two major reforms to the way television operated in the US that restricted the networks' control over broadcast content. One reform, the Prime Time Access Rule, limited the amount of primetime local television stations could devote to their affiliated network's programming. For example ABC's owned and operated television station WABC-TV in New York could not show more than three hours of content from the ABC national network during the 7:00 to 11:00 primetime block each evening. This paper focuses on the impact of the other reform, the Financial Interest and Syndication Rules ("fin-syn rules").

The fin-syn rules mandated that the networks sell off syndication rights to any shows they broadcasted during primetime hours. Although a studio affiliated with a network could still finance a show, syndication rights needed to be sold before broadcasting the show. In effect the FCC rule triggered a change in the incentives to vertically integration, see Figure 1. Although in 1983 the FCC drafted a tentative rule change that would have relaxed the fin-syn rules, they remained unmodified until the early nineties. In 1990 a Fox petition to alter rules triggered a chain of events that led to their full repeal in 1995 (Herskovitz 1997). This paper focuses on the initial 1970 rollout of the fin-syn rules as its source of exogenous variation.

3.3 Theoretical Approach

3.3.1 Reduction in Vertical Integration

The fin-syn rules altered the incentives for the networks to fund their broadcasted television shows but did not directly restrict that funding. Institutional characteristics suggests a property-rights model could be appropriate to understand the effect on network funding.

A show's revenue streams can be broken into two parts: the revenue generated from its first broadcast and the revenue generated from subsequent broadcasts. The second part of the revenue stream only materializes if the network renews the show for enough seasons; the renewal can be thought of as an asset specific investment. When the network does not own the show, it only captures a part of the first revenue stream based on the advertising generated during the show's initial broadcast. The network makes the renewal decision for the show based on a portion of the total potential value of the show. When the network does own the show, the decision is based on the show's entire revenue stream; it considers not only the advertising revenue from initial broadcast but also the potential for the show to generate additional syndication profits from later broadcasts.

Although the network could have still funded a show and then resold the rights before broadcast, under a property-rights model there would be no gain in doing so. The network effectively sells ownership of the asset before making the asset specific investment; it can only decide whether to renew the show after the show's first broadcast reveals the revenue potential for the show. Assuming a pure financial entity would provide more efficient funding for a show than a television network, this implies networks will not finance shows after the fin-syn rules would reduce the practice of networks financing television.

> Hypothesis 1A: Under a property rights model, the share of television shows funded by the broadcasting network drops after the fin-syn rules are enacted.

In contrast, under an incentive model there would still be gains to integration. The network could use financing to ensure the production company's efforts aligned with their programing needs. Prior to broadcast, the network would sell off its ownership of the show.

> Hypothesis 1A: Under an agency model, the share of television shows funded by the broadcasting network is unaffected by the fin-syn rules.

Since the networks are no longer taking the second source of revenue into account when deciding to a renew a show, shows are in effect weakly less profitable for them. This should lead to a reduction in the level of renewals in the industry under either model.

Hypothesis 2: The share of television shows renewed for a second season by the broadcasting network drops after the fin-syn rules are enacted.

3.3.2 Effect on Industry Production

Independently financed shows have a different risk profile from network financed shows. An independently financed show has convinced a network to renew the show without providing the network with an upside from re-broadcast revenue. The types of shows independently financed are therefore likely to be characteristically different from network financed shows. Arrow (1975) suggests another mechanism for a similar effect: the types of shows that were vertically integrated required more information flow between supplier and buyer to execute properly. Those show types would be harder to execute without vertical integration. As the industry shifts away from network financing of television, either argument suggests the type of content broadcast should also shift away from the characteristics common in network financed shows.

Hypothesis 3A: After the fin-syn rules are enacted, the type of television shows created shifts towards the type of shows independently funded prior to the rules.

However, another set of assumptions might produce the opposite effect. If the networks viewed the type of shows that were vertically integrated as high value, its possible they maintained industry dominance in those types of shows. Independent funding was spent on other types of shows. In this case, the relaxation of vertical integration could lead to independently funded shows to start entering this product market, increasing these types of shows relative to others.

Hypothesis 3B: After the fin-syn rules are enacted, the type of television shows created shifts towards the type of shows vertically integrated prior to the rules.

Vertical integration could have allowed for greater risk on the part of show creators. If so, a reduction in vertical integration should have reduced the level of risk the industry was willing to accept. Shows would appear more similar to successful shows in the past.

Hypothesis 4A: Television shows after the fin-syn rules were less innovative than before.

With the major networks no longer funding shows, an alternate set of financier decision makers could have selected new show ideas, leading to an increased diversity of shows.

Hypothesis 4B: Television shows after the fin-syn rules were more innovative than before.

3.4 Data and Measures

To empirically test the above hypotheses, I pool data from two sources: the Internet Movie Database (IMDb) and Wikipedia. IMDb, a subsidiary of Amazon.com, has a public dataset which includes several key variables for my analysis, including the broadcasting network, funding sources as well as show genre. Wikipedia is used to determine which shows were broadcast during the primetime period for their initial run as well as calculate the number of links a show has.

The vertical integration status of a show is a constructed variable in my dataset. There are three different potential methods for me to construct this variable given the sources I have available. One is to mark shows who had a funding source matching its broadcast network in IMDb. A second is to use IMDb's copyright information to match copyright holders with the broadcasting network. A third method is to use Wikipedia's information on funding sources and broadcast network rather than IMDb.

Goolsbee (2007) is an earlier paper that used variation in show ownership during primetime from 2000 to 2004. Goolsbee's data source was the Internet Television Almanac, a book published on the television industry which covers all broadcast shows. Figure 2 contrasts the vertical integration measure used by Goolsbee against the potential measures for this paper. If the Goolsbee data and the IMDB are two independent measures of the same phenomenon, the figure suggests that the IMDB measure of vertical integration is perhaps the most accurate one in our dataset so this paper focuses its results on the IMDB measure.

Table 1 provides summary statics for my dataset, which spans the years 1964 to 1975. Many shows in this period were created for sale directly to affiliate networks rather than for the major networks. Within the major networks, vertical integration was common and dropped from 1970 onwards. The overall mix of shows produced seems otherwise stable in this time period as suggested by the genre, rating and renewal information presented.

IMDB has genres associated with each television show. By 1952, 26 different genres had been observed in scripted television shows on the major broadcast networks. Since then no new genre was observed on major broadcast networks, only different combinations of the different genres. The use of genre as a measure for innovation was pioneered in film by Mezias and Mezias (2000) while equating innovation with new combinations of film genres was first done by Perretti and Negro (2007). Although previous research has used genre as a measure of content diversity in television (van der Wurff and van Cuilenburg 2001; Kennedy 2002; Bielby and Bielby 2003; Tsourvakas 2004), this paper is the first to use genre to measure innovation in television. The space of potential combinations does not seem to be fully explored yet. Although the most common number of individual genres a show can have is 3, only 474 of the possible 18279 combinations of 3 base genres have been used in TV shows in my full dataset to date. Table 2 provides some examples of shows marked as new genre combinations prior to the shock.

3.5 Empirical Framework

To test Hypothesis 1 in my dataset, I use a difference-in-difference estimator on cross sectional data. Each show i was first broadcast year t on network n:

$$\begin{aligned} VerticallyIntegrated_{int} &= \alpha_n + \delta_t + \beta_B BigThreePrimetime_i + \beta_P Post1969_t \\ &+ \beta_{BP} PrimetimeBigThree_i * Post1969_t + \varepsilon_{int} \end{aligned} \tag{1}$$

In Equation 1, $VerticallyIntegrated_{int}$ is an indicator for shows that IMDb has the same entity for its original distributor and one of its funding sources. $BigThreePrimetime_i$ is an indicator for shows that were originally broadcast on one of the big three networks during primetime hours. $Post1969_t$ is an indicator for the years 1970 to 1975. α_n is a fixed effect representing five networks: ABC, NBC, CBS, PBS and a dummy network for all shows IMDb lists as not originally distributed by the others. δ_t is a year fixed effect. The coefficient of interest in Equation 1 is β_{BP}. It should capture whether the big three networks shifted their primetime programing towards independent funding.

The treated group in my natural experiment are television shows broadcast on the three major US networks during primetime. A natural control for this group might be the non-primetime shows on those same networks. However, there are two issues with using non-primetime shows as a control group. First, the complementarity of vertical integration decisions (Novak and Stern, 2009) could cause non-primetime shows to be indirectly treated by the FCC shock; the networks could shut down their financing divisions altogether if the reduction in primetime financing reduced their economies of scale enough to make financing unprofitable. Second, the types of shows created in primetime are different than non-primetime: daytime shows are more likely to be soap operas for example. I therefore also include the two major UK television networks of the period, BBC and ITV, as part of my control group in addition to the non-primetime shows of the major US networks.

I attempt to cluster all my estimators at the network by primetime level, which is consistent with the existing difference-in-difference best practice of clustering at the level of control and treated groups invariant of time (Bertrand, Duflo and Mullainathan, 2004). Unfortunately, I only have primetime data for the big three networks in the US. In practice this means my UK control observations are only clustered at the network level.

For Hypothesis 2, a similar specification is used:

$$\begin{aligned} Renewed_{int} = \alpha_n + \delta_t + \beta_B BigThreePrimetime_i + \beta_P Post1969_t \\ + \beta_{BP} PrimetimeBigThree_i * Post1969_t + \varepsilon_{int} \end{aligned} \tag{2}$$

Equation 2's dependent variable is $Renewed_{int}$. The coefficient of interest, β_{BP}, measures how much this renewal rates was affected by the change in show financing.

For Hypothesis 3A and 3B, I use the genre drama as my dependent variable:

$$\begin{aligned} IsDrama_{int} = \alpha_n + \delta_t + \beta_B BigThreePrimetime_i + \beta_P Post1969_t \\ + \beta_{BP} PrimetimeBigThree_i * Post1969_t + \varepsilon_{int} \end{aligned} \tag{3}$$

Table 3 provides the intuition why. Prior to the change in rules, dramas were more common among vertically integrated shows versus independently funded shows. Coefficient β_{BP} in Equation 3 tests whether the probability a show is a drama increased or decreased after the rules took effect. A decrease would suggest the industry shifted towards independent type shows, supportive of Hypothesis 3A. An increase suggests Hypothesis 3B, a shift towards the type of shows that used to be vertically integrated.

To test Hypothesis 4A and 4B, I use the genre measure of innovation:

$$NewGenreCombination_{int} = \alpha_n + \delta_t + \beta_B BigThreePrimetime_i + \beta_P Post1969_t + \beta_{BP} PrimetimeBigThree_i * Post1969_t + \varepsilon_{int} \tag{4}$$

In Equation 4, $NewGenreCombination_{int}$ indicates whether the specific set of genres for show i was novel at time t in the market for network n. The two markets I use are US and UK, which effectively means that for example a US broadcast show is considered novel if no other US network had broadcast a show of that genre before. A negative result supports Hypothesis 4A, the level of innovation dropped in the industry when vertical integration was curtailed. A positive result support Hypothesis 4B, innovation flourished with less financial control over the product by the networks.

3.6 Results

3.6.1 Main Findings

Table 4 estimates Equation 1. Once controls are added for genre and network, the estimate suggest an approximately 13% drop in the share of primetime shows on the big three networks that were partially funded by those networks. This is consistent with magnitude of the effect plotted in Figure 1: the change in vertical integration from about 25% to 10% of shows was specific to those primetime shows as opposed to a general industry wide effect. Figure 3 plots the point estimates by year for the difference between the treatment and control groups to check whether the parallel trends assumption holds. Although the estimates are noisy, a pre-trend does not seem to exist from 1964 to 1969. Overall this evidence is supportive of Hypothesis 1; when the benefits to integration we removed, the frequency of integration dropped.

Table 5 estimates Equation 2. The overall first year renewal rate for the industry was about 35%, see Table 1. Primetime shows on the big three networks traditionally had a 15% higher renewal rate than other shows. Figure 4 plots the point estimates by year. Again, no pre-trend is apparent although the point estimates are noisy. After the rule change, this difference in renewal rate disappeared and may have even become negative, supportive of Hypothesis 2. The drop in show financing meant the networks were less likely to earn syndication revenue from renewed shows. As upside from renewing shows was reduced, the networks lowered their renewal rate.

Table 6 estimates Equation 3. Prior to the fin-syn rules, about 30% of all primetime shows on the big three networks were dramas. After the fin-syn rules reduced the level of vertical integration, this increased to 45%. As shown in Table 3, primetime dramas on the big three networks were primarily independently funded prior to 1970. Figure 5 plots the point estimates by year, indicating no pre-trend. Hence the results are supportive of Hypothesis 3A, the restriction in vertical integration led the industry to produce more shows similar to the independently funded shows of the past.

A drop in new combinations of genres is documented in Table 7's estimate of Equation 4, supportive of Hypothesis 4A. Table 1 suggests an overall rate of new genres per year of 15%. This dropped to basically zero for primetime shows on the big three networks after the FCC rule change. Figure 6 plots the yearly point estimates for the treatment effect. Here the base year, 1969, seems to be lower than the previous point estimates. Since the other difference-in-difference estimates lack this feature, I argue the 1969 estimate is due to the stochastic nature of innovation; 1969 was lower than previous years by chance. Table 8 shows historically the network funded shows on the big three networks had a higher rate of new genre combinations. The reduction in vertical integration perhaps made the industry risk adverse; the propensity to try out a new genre combination dropped during primetime hours on the major networks relative to the rest of the industry.

3.6.2 Robustness Checks

Prime Time Access Rule

The year after the fin-syn rules were enacted, the Prime Time Access Rule, "PTAR", was implemented. PTAR reduced limited the number of primetime hours the major networks could provide programing for their affiliate stations, effectively reducing primetime from 7:30PM-11:00PM to 8:00PM-11:00PM. The networks chose to curtail the 7:30PM slot since primetime viewership was lowest at that time.

To test for any confounding effect from this rule change, I rerun my analysis using only data from 8:00PM to 11:00PM over my entire time period for the effected networks. I find these shortened primetime results are consistent with my main results.

3.7 Conclusion

This paper provides an empirical study of a case where supplier financing by the buyer, a form of vertical integration, is exogenously restricted. In my setting the buyers are the television networks and the suppliers are the creators of television shows. I first show a regulatory change lowered the incidence of vertical integration for some networks, consistent with property-rights models. I then show this shift in financing was contemporaneous with a change in the product mix provided by the industry: more dramas were produced while overall show quality and innovation rates fell.

I interpret my results as evidence of the importance of vertical integration for the stream of innovations produced by an industry. In my setting, the products are television shows and innovations are new versions of those products. By shifting the mix of product types that were created in my setting, the restriction on vertical integration effectively altered the set of future innovations: the direction of innovation. There could also be a drop in the rate of innovation in my setting as the quality of new shows decreased after the restriction to vertical integration.

The finding that organizational form matters to the direction of innovation is broadly applicable to the field of innovation. Research on the innovation has generally focused on the effect of industry structure and incentives on the rate of innovation, looking at relationship between competition and R&D investment (Aghion et al., 2005; Segal and Whinston, 2007) or the incentives for such investment (Aghion and Tirole, 1994; Budish, Roin and Williams, 2015). My contribution is showing how organizational form can alter innovation in an industry. The strategic decisions entrants make to partner or compete has wider consequences for the direction of innovation (Gans and Stern, 2003).

3.8 References

Adalian, J., 2013. "Straight to Series: The Networks' Big-Money Bet to Skip Pilots." *Vulture.* November 27th.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt. (2005). "Competition and Innovation: An Inverted-U Relationship." *Quarterly Journal of Economics,* 120, 701-728.

Aghion, P. and Tirole, J., 1994. "The management of innovation." *Quarterly Journal of Economics,* 109(4), pp.1185-1209.

Anderson, E. and Schmittlein, D.C., 1984. "Integration of the sales force: An empirical examination." *The Rand Journal of Economics*, pp.385-395.

Arrow, K.J., 1975. "Vertical integration and communication." *The Bell Journal of Economics*, pp.173-183.

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources for Invention." *In The Rate and Direction of Inventive Activity: Economic and Social Factors*, pp. 609-625. Princeton, NJ: Princeton University Press.

Azoulay, P., 2004. "Capturing knowledge within and across firm boundaries: evidence from clinical development." *The American Economic Review*, 94(5), 1591-1612.

Baker, G.P. and Hubbard, T.N., 2003. "Make versus buy in trucking: asset ownership, job design, and information." *The American Economic Review,* 93(3), pp.551-572.

Baker, G.P. and Hubbard, T.N., 2004. "Contractibility and asset ownership: On-board computers and governance in US trucking." *The Quarterly Journal of Economics,* 119(4), pp.1443-1479.

Bertrand, M., Duflo, E. and Mullainathan, S., 2004. "How much should we trust differences-in-differences estimates?." *The Quarterly Journal of Economics, 119*(1), pp.249-275.

Bielby, W.T. and Bielby, D.D., 2003. "Controlling prime-time: Organizational concentration and network television programming strategies." *Journal of Broadcasting & Electronic Media, 47*(4), pp.573-596.

Budish, E., Roin, B.N. and Williams, H., 2015. "Do firms underinvest in long-term research? Evidence from cancer clinical trials." *American Economic Review, 105*(7), pp.2044-85.

Bunn, A., 2002. "The Wasteland." *The New York Times Magazine.* June 23rd.

Coase, R.H. "The Nature of the Firm," *Economica.* Vol. 4(16), Nov. 1937, pp. 386-405.

Cohen, W. and D. L. Levinthal, 1990. "Absorptive Capacity: A New Perspective on Learning and Innovation," *Administrative Science Quarterly*, 35, 1, 128-152.

Covington Jr, W.G., 1994. "The financial interest and syndication rules in retrospect: History and analysis." *Communications and the Law*, 16, p.3.

Cusumano, M. A., & Nobeoka, K. 1992. "Strategy, structure and performance in product development - observations from the auto industry." *Research Policy*, 21(3), 265-293.

Forbes, S.J. and Lederman, M., 2009. "Adaptation and vertical integration in the airline industry." *The American Economic Review,* 99(5), pp.1831-1849.

Forbes, S.J. and Lederman, M., 2010. "Does vertical integration affect firm performance? Evidence from the airline industry." *The RAND Journal of Economics,* 41(4), pp.765-790.

Fritz, B. and Flint, J. 2017. "Where Amazon Is Failing to Dominate: Hollywood." *Wall Street Journal.* 6 Oct

Gans, Joshua S., and Scott Stern. 2003. "The Product Market and the Market for Ideas: Commercialization Strategies for Technology Entrepreneurs." *Research Policy* 32: 333-350

Gibbons, Robert, "Four formal(izable) theories of the firm?" *Journal of Economic Behavior & Organization,* 2005

Goolsbee, A., 2007. "Vertical integration and the market for broadcast and cable television programming." *Federal Communications Commission Media Ownership Study*, 9.

Grossman, S.J. and O.D. Hart. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *The Journal of Political Economy.* Vol. 94(4), Aug. 1986, pp. 691-719.

Herskovitz, M.L., 1997. "The repeal of the financial interest and syndication rules: The demise of program diversity and television network competition." *Cardozo Arts & Entertainment Law Journal*, 15, p.177.

Joskow, P.L., 1985. "Vertical integration and long-term contracts: The case of coal-burning electric generating plants." *Journal of Law, Economics, & Organization,* 1(1), pp.33-80.

Kennedy, R.E., 2002. "Strategy fads and competitive convergence: An empirical test for herd behavior in prime-time television programming." *The Journal of Industrial Economics*, *50*(1), pp.57-84.

Kerr, W.R., Nanda, R. and Rhodes-Kropf, M., 2014. "Entrepreneurship as experimentation." *The Journal of Economic Perspectives,* 28(3), pp.25-48.

March, J. G. and H. Simon. *Organizations.* New York: John Wiley 1958.

Masten, S.E., 1984. "The organization of production: Evidence from the aerospace industry." *The Journal of Law and Economics,* 27(2), pp.403-417.

Mezias, J.M. and Mezias, S.J., 2000. "Resource partitioning, the founding of specialist firms, and innovation: The American feature film industry, 1912–1929." *Organization Science*, 11(3), pp.306-322.

Monteverde, K. and Teece, D.J., 1982. "Supplier switching costs and vertical integration in the automobile industry." *The Bell Journal of Economics*, pp.206-213.

Novak, S., & Stern, S. (2008). "How does outsourcing affect performance dynamics? Evidence from the automobile industry." *Management Science,* 54(12), 1963-1979.

Novak, S. and Stern, S., 2009. "Complementarity Among Vertical Integration Decisions: Evidence from Automobile Product Development." *Management Science*, 75: 1257-85.

Ocasio, A. 2012. "TV Success Rate: 65% Of New Shows Will Be Canceled (& Why It Matters)" *Screen Rant.*

O'Connell, Michael, 2018. "Showtime Boss Talks Subscriber Growth, 'Twin Peaks' and Mark Halperin." *The Hollywood Reporter.* January 6th.

Perretti, F. and Negro, G., 2007. "Mixing genres and matching people: A study in innovation and team composition in Hollywood." *Journal of Organizational Behavior*, *28*(5), pp.563-586.

Rawley, E. and Simcoe, T.S., 2013. "Information technology, productivity, and asset ownership: Evidence from taxicab fleets." *Organization Science*, 24(3), pp.831-845.

Scotchmer, Suzanne. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law." *Journal of Economic Perspectives,* 5(1): 29-41

Segal, Ilya, and Michael Whinston. 2007. "Antitrust in Innovative Industries." *American Economic Review,* 97(5): 1703-1730.

Shaked, A. and Sutton, J., 1982. "Relaxing price competition through product differentiation." *The Review of Economic Studies*, pp.3-13.

Thaxton, R., 2017. "Twin Peaks Breaks Showtime Subscription Records, Yet Notches Mediocre Ratings." *Paste*. May 23rd.

Tsourvakas, G., 2004. "Public television programming strategy before and after competition: The Greek case." *Journal of Media Economics*, *17*(3), pp.193-205.

van der Wurff, R. and van Cuilenburg, J., 2001. "Impact of moderate and ruinous competition on diversity: The Dutch television market." *The Journal of Media Economics*, *14*(4), pp.213-229.

Waldfogel, J., 2017. "The random long tail and the golden age of television." *Innovation Policy and the Economy*, *17*(1), pp.1-25.

3.9 Figures and Tables

3.9.1 Figures

Figure 1: Share of Primetime Vertically Integrated Shows on Big Three Networks

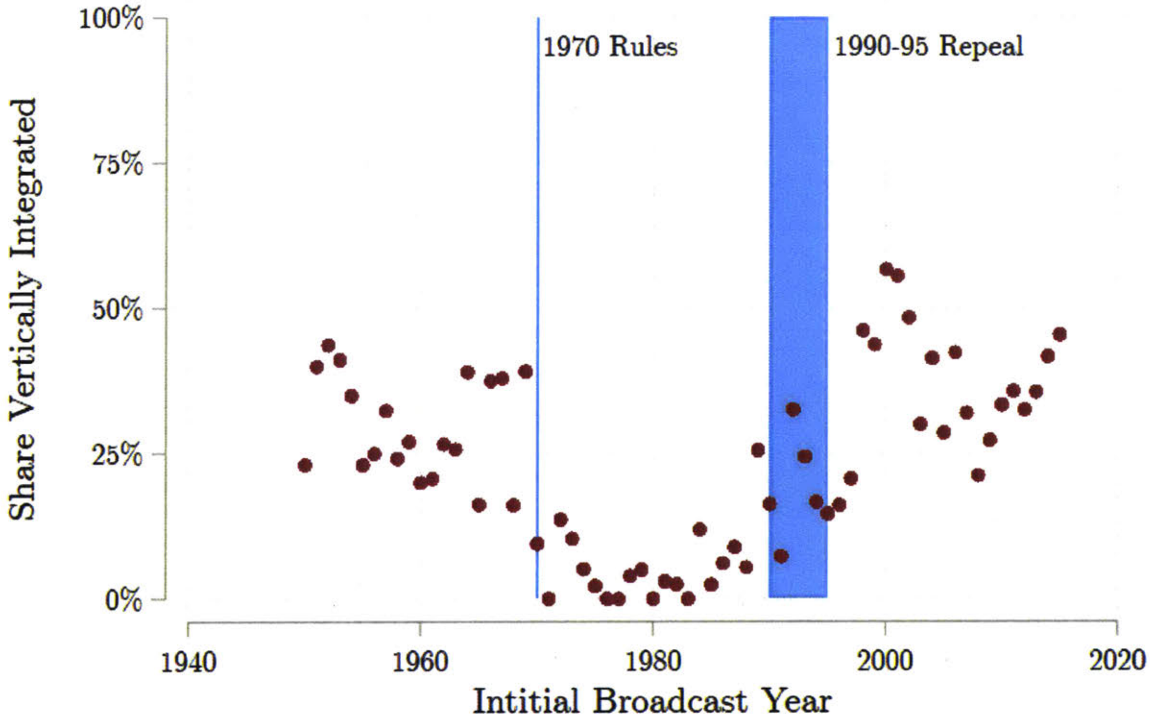


Figure 2: Measures of Television Show Vertical Integration

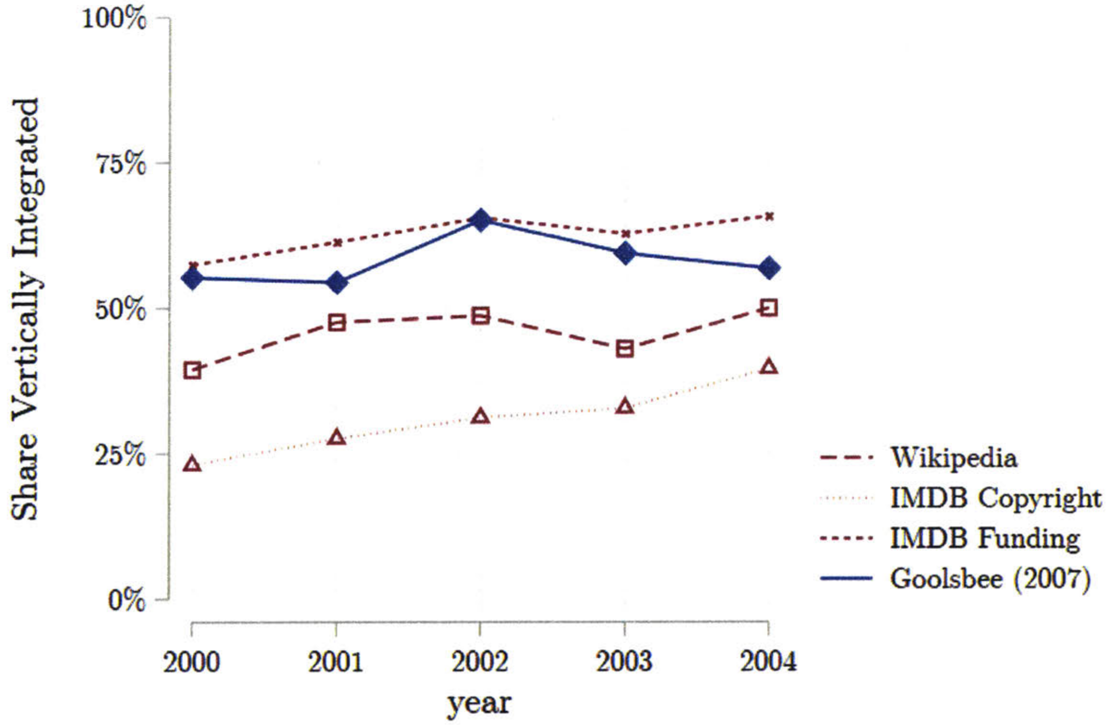


Figure 3: Difference-in-Difference Estimator of Vertical Integration

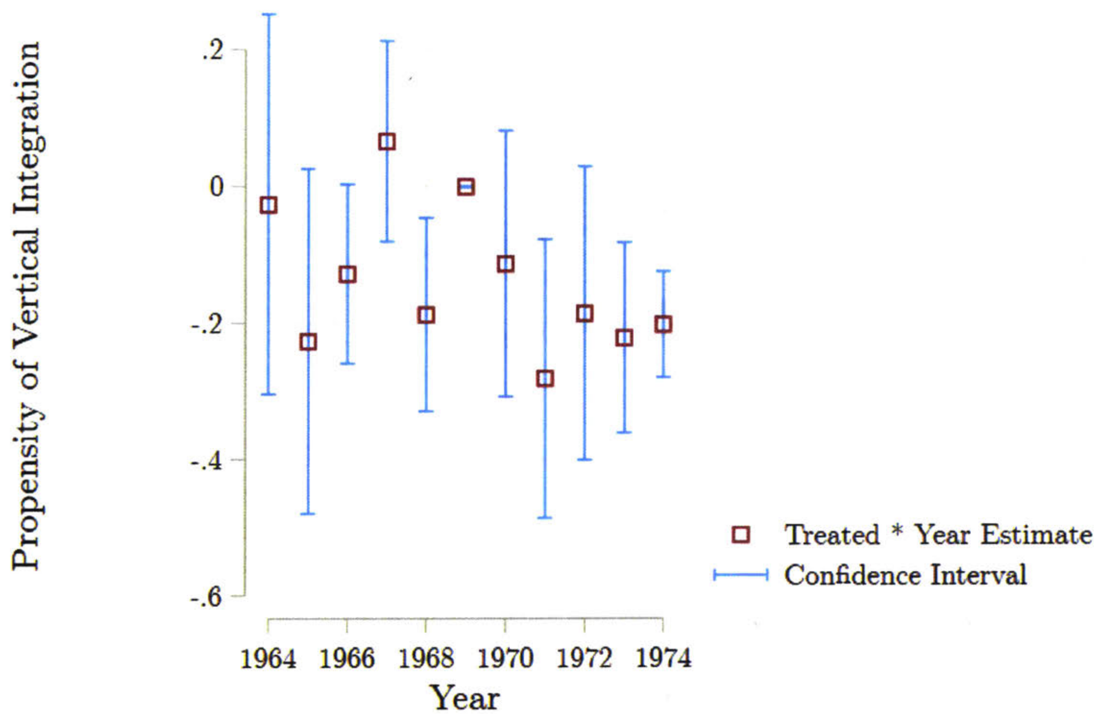


Figure 4: Difference-in-Difference Estimator of Renewal Rates

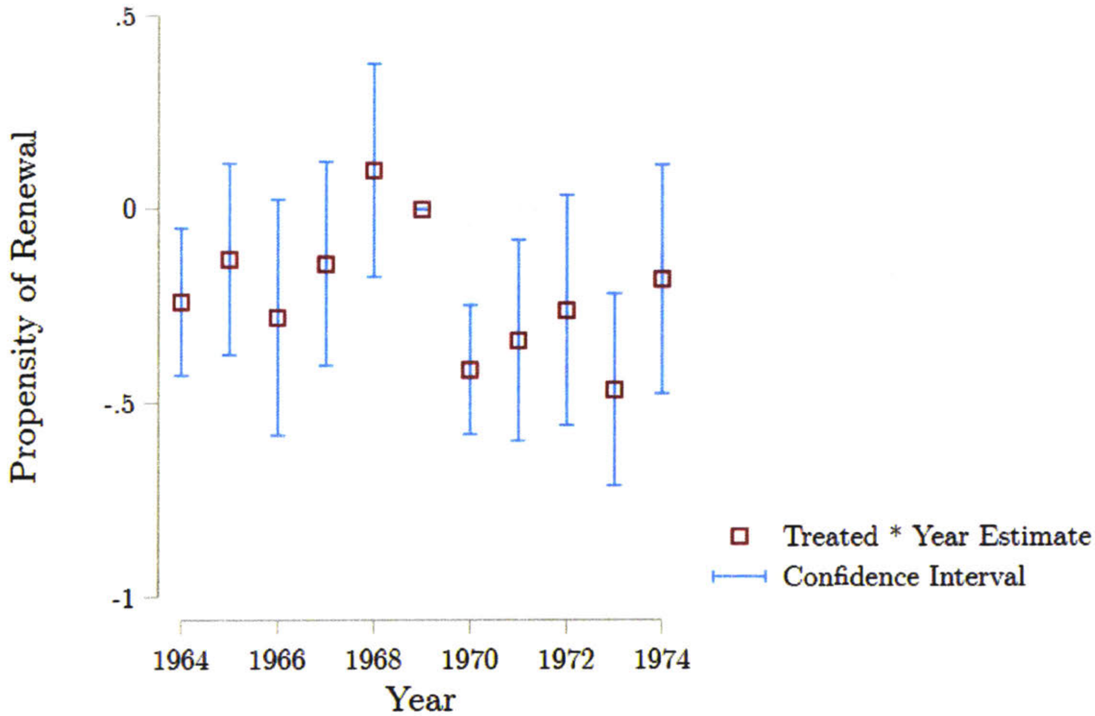


Figure 5: Difference-in-Difference Estimator of Dramas

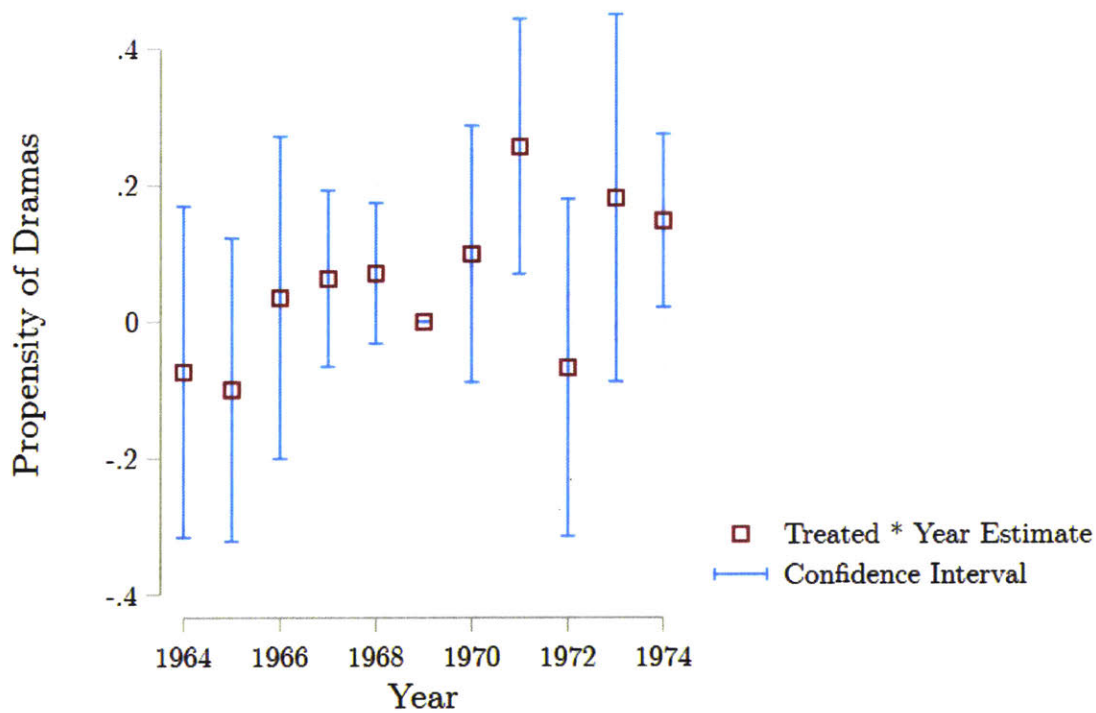
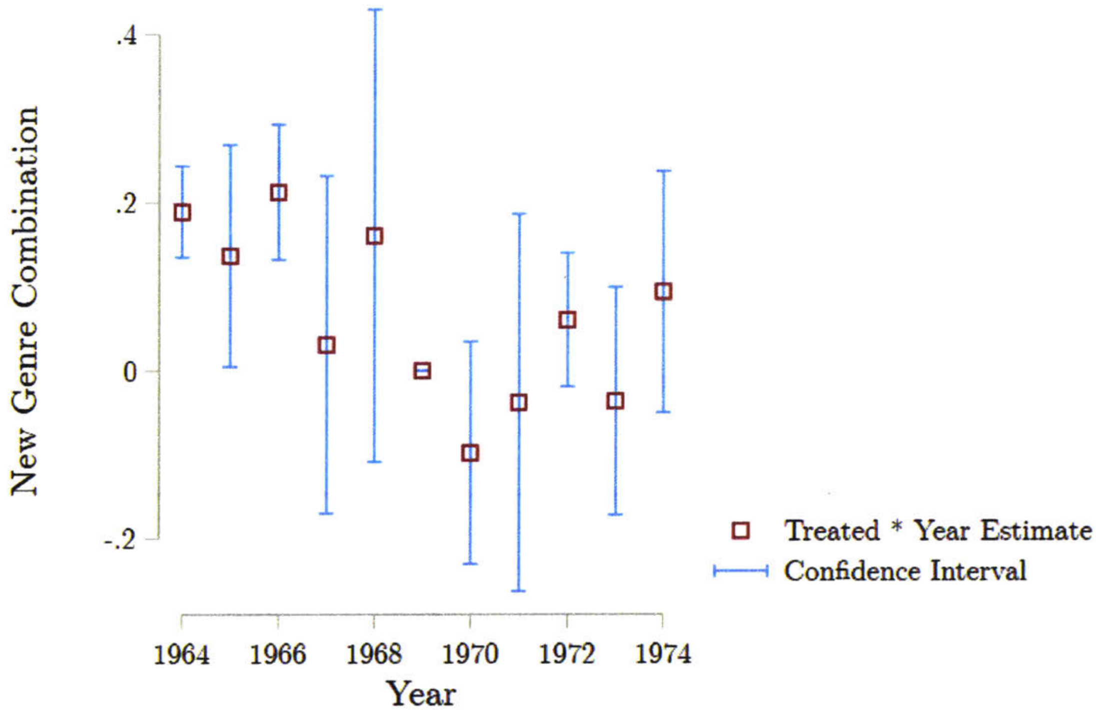


Figure 6: Difference-in-Difference Estimator of New Genre Combinations



3.9.2 Tables

Table 1: Summary Statistics

	1965-1969	1970-1975	Overall
Total New Shows	809	591	1400
Average per Year	134.8	118.2	127.3
Share Vertically Integrated	54%	35%	46%
Big Three Network Shows	356	275	631
Average per Year	59.3	55.0	57.4
Share Vertically Integrated	37%	19%	29%
Primetime Big Three Network Shows	200	129	329
Average per Year	33.3	25.8	29.9
Share Vertically Integrated	30%	9%	22%
Is a Comedy	26%	31%	28%
Is a Drama	36%	38%	37%
Is a Family Show	20%	19%	20%
Is a Musical	15%	9%	13%
Renewed Past First Season	34%	35%	35%
New Genre Combination	17%	14%	16%

Table 2: Primetime New Genre Combinations on Big Three Networks in 1966

Title	Network	Genres
ABC Stage 67	ABC	Drama, Comedy, Family, Musical
Batman	ABC	Half Hour, Comedy, Action, Family, Crime, Adventure, Fantasy, SciFi, Mystery
Daktari	CBS	Family, Adventure
Mission: Impossible	CBS	Drama, Action, Thriller, Crime, Adventure
Star Trek	NBC	Action, Adventure, SciFi, Mystery
Tarzan	NBC	Action, Adventure
The Girl from U.N.C.L.E.	NBC	Action, Adventure
The Green Hornet	ABC	Half Hour, Action, Crime, SciFi
The Man Who Never Was	ABC	Half Hour, Drama, Action
The Rat Patrol	ABC	Half Hour, Drama, Action, Adventure, War
The Time Tunnel	ABC	Action, Adventure, SciFi

Table 3: Popular Genres by Show Vertical Integration Status

	Independen t	Integrated
Is a Drama	75%	25%
Is a Comedy	68%	32%
Is a Family Show	65%	35%
Is a Music Show	56%	44%

Share of funding sources for primetime shows by genre prior to 1970 on the big three networks. Only genres associated with greater than 10% of shows included. Independent shows did not have the original broadcast network listed as a funding source while integrated shows did have.

Table 4: Estimating Change in Vertical Integration

	(1) Vertically Integrated	(2) Vertically Integrated	(3) Vertically Integrated
Big Three Primetime	-0.321	-0.116**	-0.121**
	[0.235]	[0.0377]	[0.0373]
Post 1969	-0.194***	-0.0849**	
	[0.0170]	[0.0317]	
Big Three Primetime * Post 1969	-0.0204	-0.134**	-0.128**
	[0.0440]	[0.0549]	[0.0541]
Constant	0.621**	0.944***	0.902***
	[0.228]	[0.0247]	[0.0236]
Genre FE		X	X
Network FE		X	X
Year FE			X
Shows (N)	1400	1400	1400
Deg. of Freedom	9	9	9
Adj. R-Squared	0.112	0.553	0.556

Robust standard errors in brackets. Star levels: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$.
Observations are year's new shows. OLS estimator used clustered at network * primetime level.
Data is restricted to shows from 1964 to 1974.

Table 5: Estimating Change in Renewal Rates

	(1) Renewed	(2) Renewed	(3) Renewed
Big Three Primetime	0.136***	0.149***	0.154***
	[0.0368]	[0.0375]	[0.0374]
Post 1969	0.0571***	0.0478***	
	[0.0103]	[0.0117]	
Big Three Primetime * Post 1969	-0.200***	-0.188***	-0.192***
	[0.0315]	[0.0227]	[0.0224]
Constant	0.309***	0.222***	0.276***
	[0.0268]	[0.0279]	[0.0208]
Genre FE		X	X
Network FE		X	X
Year FE			X
Shows (N)	1400	1400	1400
Deg. of Freedom	9	9	9
Adj. R-Squared	0.00808	0.0257	0.0251

Robust standard errors in brackets. Star levels: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$.
Observations are year's new shows. OLS estimator used clustered at network * primetime level.
Data is restricted to shows from 1964 to 1974.

Table 6: Estimating Change in Genre

	(1) Is Drama	(2) Is Drama	(3) Is Drama
Big Three Primetime	-0.0493	-0.000678	-0.00558
	[0.0357]	[0.0241]	[0.0222]
Post 1969	0.0207	0.0335	
	[0.0213]	[0.0204]	
Big Three Primetime * Post 1969	0.147***	0.135***	0.132***
	[0.0280]	[0.0268]	[0.0242]
Constant	0.269***	0.311***	0.304***
	[0.0355]	[0.00786]	[0.0270]
Network FE		X	X
Year FE			X
Shows (N)	1400	1400	1400
Deg. of Freedom	9	9	9
Adj. R-Squared	0.00612	0.0118	0.0139

Robust standard errors in brackets. Star levels: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$.
Observations are year's new shows. OLS estimator used clustered at network * primetime level.
Data is restricted to shows from 1964 to 1974.

Table 7: Estimating Change in New Genre Combinations

	(1) New Genre Combination	(2) New Genre Combination	(3) New Genre Combination
Big Three Primetime	0.0540	-0.0382	-0.0428
	[0.0405]	[0.0526]	[0.0517]
Post 1969	0.00418	-0.00600	
	[0.0229]	[0.0212]	
Big Three Primetime * Post 1969	-0.129***	-0.121***	-0.120**
	[0.0378]	[0.0363]	[0.0372]
Constant	0.156***	0.0967***	0.0501
	[0.0304]	[0.0277]	[0.0484]
Genre FE		X	X
Network FE		X	X
Year FE			X
Shows (N)	1400	1400	1400
Deg. of Freedom	9	9	9
Adj. R-Squared	0.00444	0.0497	0.0514

Robust standard errors in brackets. Star levels: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$.
Observations are year's new shows. OLS estimator used clustered at network * primetime level.
Data is restricted to shows from 1964 to 1974.

Table 8: New Genre Combinations by Show Financing

	New Genre Combination
Independently Financed	22%
Network Financed	28%

Share of new genre combinations by financing used prior to 1970 on the big three networks. Independent shows did not have the original broadcast network listed as a funding source while integrated shows did have.

4. Venture Capital Rents from Entrepreneurial Search

Ankur Chavda, Joshua Gans, Scott Stern
March 2019

Abstract

Entrepreneurs in high growth industries face a unique form of uncertainty in their search for strategies to execute their ideas: the underlying distribution of potential outcomes is unknown. This uncertainty creates an opportunity for venture capitalists to extract value in certain cases by resolving that uncertainty and improving the search prospects for entrepreneurs. This paper models the optimal search problem faced by entrepreneurs and finds the value generated by venture capitalists is non-monotonic in the best strategy discovered so far by an entrepreneur. Our results suggest the rents captured by venture capital may be driven by selection of a specific kind of entrepreneur: one with a great idea but poor strategy for executing that idea.

4.1 Introduction

Improving the entrepreneurial search process for new ideas is not only relevant to entrepreneurs themselves but also for broader economic growth (Schumpeter, 1942). Scholars have long studied entrepreneurial search to provide recommendations about how to discover new ideas (Gaglio and Katz, 2001; Fiet and Patel, 2008). However, although prior research acknowledges the uncertainty present in the search for new ideas, it fails to fully account for the implications of that uncertainty, specifically the learning produced by the entrepreneurial search process and how venture capitalist can substitute for that learning. Entrepreneurs frequently explore new ideas whose potential is unknown. Searching for strategies based on these ideas not only results in the discovery of strategies themselves, but information about the idea's potential. The lack of this information at the beginning of the process can lead the search for strategies to go on for too long or stop prematurely. Hence the opportunity for venture capitalists to perform arbitrage by providing private information about the true distribution of strategy outcomes for an idea.

As an illustrative example, consider the process used by Jeff Bezos in 1994 to found Amazon, when the potential for commerce over the internet was unknown. Before launching an online bookstore, Bezos considered several strategies for launching an online retailer, including selling clothing and music, any of which would have been considered a viable business opportunity at the time and therefore worthy of execution (Fiet, 2007). Bezos however realized the existence of so many good strategies in e-commerce suggested the potential existence of even better strategies, finally landing on retailing books as Amazon's initial business strategy. Entrepreneurial search here did not stop with the first good strategy, rather the discovery of a good strategy lead to more search: if at first you do succeed, try again.

This paper provides a search model accounting for learning in its optimal stopping rule that predicts behavior exhibited in the Amazon example, where the existence of good strategies triggers additional search. This stands in contrast to existing models (Fiet, Piskounov and Patel, 2005) where the discovery of good strategies can only terminate entrepreneurial search. The model also suggests it may be optimum for entrepreneurs to return to old, discarded strategies in some cases, another departure the current literature on entrepreneurial search.

The rest of this paper proceeds as follows. Section 2 provides a literature review of optimal stopping rules applied to innovation and entrepreneurial contexts. Section 3 uses an example to illustrate our main findings. Section 4 provides our generalized results. Section 5 concludes with a discussion of the general applicability of this paper.

4.2 Prior Literature

The concept of uncertainty has a long history in the entrepreneurial literature; Knight (1921) essentially defines an entrepreneur as the economic agent responsible for decision making under unmeasurable uncertainty, a type of uncertainty where the distribution of outcomes is unknown. Kirzner (1997) argues the discovery of entrepreneurial opportunities under this uncertainty is core to the functioning of an economy; entrepreneurship acts as the mechanism that brings markets into equilibrium as arbitrage opportunities are turned into new ventures. Crucially, Kirzner allows for mistakes along the way: entrepreneurs can be wrong about the expected return of their actions due to the existence of uncertainty.

Shane and Venkataraman (2000) point out that uncertainty triggering the existence of opportunities and the discovery of those opportunities is insufficient for a complete theory of entrepreneurship; an understanding of the decision to execute on those opportunities is also needed: a model for entrepreneurial search under uncertainty. Current models of entrepreneurial search lack a satisfactory approach to uncertainty. Fiet, Piskounov and Patel (2005) for example use a bandit approach where the entrepreneur selects which among various sets of opportunities to search. Learning about uncertainty is incorporated by the entrepreneur abandoning search within a set of opportunities after multiple bad draws. However, there are two fundamental mismatches between a bandit setup and our understanding of how entrepreneurs search in practice. First, an entrepreneur's unique backgrounds enable them to uncover unique opportunities (Shane, 2000) and must decide on strategies to execute an uncovered opportunity as opposed to pick one among several opportunities (Gans, Stern and Wu, 2016). Second, bandit theory's major results optimize the average value of the discoveries, see Robbins (1952) or Feldman (1962). Entrepreneurs of course care less about the average quality of all their ideas and more about the singular potential of their best idea.

Our paper is therefore grounded in the sequential analysis first developed by Wald (1945) for classified war time quality control efforts (Wallis, 1980). Wald's approach was tailored to the case where observations were made in sequence and a decision to continue needed to be made after every observation. Sequential analysis is a stronger analogy for the entrepreneurial decision making process than bandit theory. Once an entrepreneur has an idea, she has to consider all the different possible ways to execute that idea before deciding to fully commit to the idea and stop experimenting. This nests well within Wald's framework.

Stigler (1962) was an early user of sequential analysis in economics. He studied job search by workers that knew the distribution of possible wages. Later work relaxed the assumption of knowledge about the distribution in both labor market search models like McCall (1970) and models of consumers searching for the lowest price as in Rothschild (1974). Kohn and Shavell (1974) provide various comparative statics in a general setting where the optimal stopping rule when learning takes place and has no closed form solution. The show stopping happens at lower draws when search costs increase and at higher draws if the mean of the distribution is raised. Talmain (1992) provided an explicit solution to Rothschild's search under uncertainty model, allowing for a searcher to accept a wage that was previously rejected which would not happen in models under certainty. Papers such as Adam (2001) that attempt to derive a more general stopping rule for sequential decision making with uncertainty make assumptions that are unsatisfactory for our stetting. Specifically in Adam's paper, cases were a high draw causes the searcher to greatly raise expected gains from future draws are ruled out.

We are particularly interested in how entrepreneurial search under uncertainty relates to the venture capitalists that fund them. Hsu (2006) identified a relationship between VC financing and the type of strategy used by a startup: VC backed startups were more likely to cooperate with incumbent firms. This goes beyond the monitoring and selection activities conducted by the VC stemming from their role as principles (Kaplan and Strömberg, 2001). A core question in entrepreneurial finance is whether the presence of VC's has any effect on an economies rate of innovation or growth; we extend this literature by showing how VC's influencing the entrepreneurial search for strategies can add economic value.

4.3 Motivating Example

4.3.1 Searching Under Uncertainty

An entrepreneur has an idea for a new startup. She can come up with multiple strategies to execute that idea or alternatively take her outside option and not do anything with the idea. Let her outside option be normalized to zero. She does not know whether the idea is a normal idea or an exceptional idea, but she does know that regular ideas tend to have strategies with payoffs distributed $\mathcal{N}(1,1/2)$ while exceptional ideas have strategies with payoffs distributed $\mathcal{N}(3,1/2)$. Suppose she believes there is a 25% chance her idea is exceptional. Figure 1 plots the distribution of strategies for each of these types of ideas as well as the mixture that results from her beliefs.

4.3.2 Non-Monotonic Search Behavior

Each time the entrepreneur searches a new strategy to understand its payoff, she incurs a cost of 0.1. She must weigh the potential gain from discovering a new strategy for her idea with a higher payoff than her current best strategy against this cost of search. If all ideas were regular, a simple cutoff rule for search would apply. As long as the best strategy found was less than about 1.25, continue searching. Once strategy is found with greater payoff, stop. However, when there is a possibility of having an exceptional idea, this simple rule is no longer optimal. Searching not only changes the best strategy found to date, but also can change beliefs about whether the distribution is regular or exceptional.

Suppose her first search leads to a discovery of a strategy with payoff 1.5. As seen in Figure 2, since the 1.5 draw is unlikely to have come from the exceptional distribution, the entrepreneur is fairly certain her idea is a regular one. This causes her to stop searching for new strategies; a 1.5 payoff is relatively high for a $\mathcal{N}(1,1/2)$ distribution so searching again is not worth incurring the cost of search. In contrast, a discovery of a strategy with payoff 2.5 suggests she likely has an exceptional idea. This induces more search since she is likely to find a better strategy when strategies are distributed $\mathcal{N}(3,1/2)$. Finally, if a 3.5 strategy had been found she would stop searching immediately. Although the entrepreneur still thinks her idea is exceptional, finding a better strategy than 3.5 even for an exceptional idea is unlikely.

These example draws capture the non-monotonic nature of search when there is uncertainty in the underlying distribution. Figure 3 plots the expected value of her idea depending on the first strategy payoff discovered as a way to illustrate the search or stop decision. Five distinct regions exist. In region A, there is a high probability the idea is regular but the discovered strategy is not good enough so search continues. The expected value of the idea is flat because stopping is will not occur until a strategy of at least 1.25 is found. In region B, the idea is still believed to be regular. However, now the discovered strategy is good relative to that distribution, so stopping occurs with entrepreneur executing the discovered strategy; the value of the idea is in fact the value of the discovered strategy. In region C there is an increasing belief that the idea might be exceptional, causing the curvature seen in the line plot. Search continues based on the chance the idea is exceptional since relative to an exceptional idea, the discovered strategy is poor. In region D there is certainty that the idea is exceptional and again the discovered strategy is relatively poor, so search continues as in region A with a regular idea. Finally, in region E the discovered strategy is good even relative to an exceptional idea so search stops. In summary search occurs in regions A, C and D but stops in B and E; the decision to search is non-monotonic in the payoff of discovered strategy.

4.3.3 Revert to Previous Strategy

When ideas only follow the regular distribution; searching for new strategies will terminate once one with a high enough payoff is found. Consequently, the strategy executed will always be the last one discovered. This does not hold when there is a possibility of an exceptional distribution as well. Suppose the entrepreneur's first discovered strategy has a payoff of 2.25, placing it in region C of Figure 3. There is enough probability of the idea being exceptional to warrant further search. Figure 4 reflects the potential next draw after the first strategy of value 2.25 was discovered. Regions C, D, and E of Figure 4 have a similar interpretation as in Figure 3. However, unlike in region A of Figure 3, in region F of Figure 4 stopping occurs because the entrepreneur is certain the 2.25 was just a great draw from the regular idea distribution. The entrepreneur goes back to an old strategy after searching for a new one, a behavior that would never occur without uncertainty about the underlying idea distribution.

4.3.4 Entrepreneur Beliefs

An entrepreneur's initial beliefs about an idea's distribution can dramatically affect the stopping decision. In the above example, if the entrepreneur's first strategy had a payoff of 1.5, she would terminate search. However, a different entrepreneur with higher confidence in the idea being exceptional could believe the 1.5 strategy was a poor draw from the exceptional distribution rather than a good draw from the regular distribution. As illustrated in Figure 5, an entrepreneur that thought her idea had a 90% probability of being exceptional would continue searching even after the 1.5 strategy was discovered. Entrepreneurs with low confidence in their ideas may miss out on an idea's true potential if that idea is exceptional.

4.3.5 Expert Advice

Given the potential importance beliefs have for search behavior, entrepreneurs will sometimes be willing to pay for certainty about the underlying distribution. Suppose a venture capitalist is an expert in the entrepreneur's idea space and can with certainty tell the entrepreneur whether the idea is regular or exceptional. The entrepreneur's willingness to pay for such advice is dependent on her current best option and belief about the idea, as depicted in Figure 6.

Our entrepreneur started off with an outside option of zero and a 10% belief the idea was exceptional. In this case, regardless of whether the idea is exceptional or regular, she will want to search for a strategy. The value of that first strategy will provide a lot of information about the underlying distribution. Conditional on her always searching regardless of the underlying idea distribution, its not worth paying for knowledge of the true distribution. Similarly, if her best discovered strategy is a 3.5, she will stop searching regardless of the idea distribution, so she is not willing to pay to know the true distribution.

When her best discovered strategy has a 2.5 payoff, she is willing to pay almost up to the cost of search when she has a 10% belief of an exceptional idea. Knowing the true distribution enables her to avoid a suboptimal decision, in this case stopping search in the 10% case when the idea is actually exceptional. As her beliefs change with the same best discovered strategy, the amount she will pay also changes. When she has a high belief the distribution is exceptional, she also believes the expert will tell her the distribution is exceptional and does not expect to change her search behavior; she sees little value in paying for knowledge of the true distribution. Similarly, when she has a very low belief the distribution is exceptional, she views the chances that the distribution is exceptional too low to justify paying a high price for that knowledge.

Advisors such as venture capitalists in this setting provide the most value when both entrepreneurs are uncertain about the underlying distribution and when they would make a different search decision depending on the true distribution.

4.4 Entrepreneurial Search Model

Consider the case where an entrepreneur is searching for strategies with recall and there is uncertainty in the distribution of strategy payoffs. At each stage i, the entrepreneur can choose to stop and receive payoff of the best strategy discovered so far or continue searching for another strategy at cost c. The distribution of strategy payoffs at each stage are not independent and not identical, represented by random variables X_i whose probability distributions form a set $\mathcal{P}$. The entrepreneur's problem is maximizing the payoff y by selecting the number of stages searched:

$$y = \sup_n[\max\{x_1, \dots x_n\} - nc] \tag{1}$$

Assumption 1: There exists a random variables $\bar{X}$ and $\underline{X}$ with distributions such that $\bar{X}$ first order stochastically dominates all marginal distributions $\Pr(X_i = x_i | X_1 = x_1, \dots, X_{i-1} = x_{i-1})$ generated from $\mathcal{P}$ and $\underline{X}$ is dominated by all marginal distributions generated from $\mathcal{P}$.

Assumption 2: $\bar{X}$ and $\underline{X}$ have finite first moments.

Assumption 3: $\bar{X}$ has a finite second moment.

Theorem 1: Under A1-A2, $y < \infty$. Under A1-A3, $E[y] < \infty$.

All proofs are in the appendix. The intuition behind Lemma 1's proof is that $\bar{X}$ and $\underline{X}$ bound all the possible distributions X_i can have and since $y < \infty$ if draws were taken only from these bounding distributions, $y < \infty$ must be true for X_i draws. Similar argument holds for the $E[y] < \infty$ result.

Comment 1: The distributions of $\bar{X}$ and $\underline{X}$ can be constructed by taking the upper and lower envelopes of CDF's from the set $\mathcal{P}$.

Corollary 1: Theorem 1's results hold when Equation 1 is modified to include a finite outside option.

Theorem 1 provides us with general conditions when the value of the entrepreneur's idea, $E[y]$, is finite: the distribution of strategies needs to have finite variance. Next, we can reformulate the search problem as a staged value function for results on the search decision made at each stage. Suppose the set of distributions $\mathcal{P}$ is comprised of a single family of distribution characterized by parameter α with corresponding pdf

$f(x;\alpha)$. Additionally, suppose α updates to $\pi(\alpha,x)$ after a search results in an item with value x.

Definition (Idea Value Function): Define the expected value of an idea with the current best strategy having payoff z and strategy distribution belief parameter α as:

$$v(z,\alpha) = \max\{z, \int v(max\{z,x\}, \pi(\alpha,x)) \cdot f(x;\alpha)dx - c\} \tag{2}$$

Assumption 4: The family of distributions $f(x;\alpha)$ displays first order stochastic dominance in α.

Assumption 5: There exists a supremum distribution $\bar{\alpha}$ with a finite first and second moments such that for all α, $f(x;\alpha) \underset{\text{FOSD}}{\leq} f(x;\bar{\alpha})$ and an infimum distribution $\underline{\alpha}$ with a finite first moment such that α, $f(x;\underline{\alpha}) \underset{\text{FOSD}}{\leq} f(x;\alpha)$.

Assumption 6: The function $\pi(\alpha,x)$ is weakly increasing in α and x.

Lemma 1: Under A4-A6, the expected value of an idea is weakly increasing in payoff z and parameter α, as well weakly decreasing in cost c.

Corollary 2: Under A4-A6, an entrepreneur with beliefs $a' \leq a^*$ where a^* is her idea's true distribution will undervalue her idea while an entrepreneur with beliefs $a'' \geq a^*$ will overvalue her idea.

In a search problems with certainty about the distribution outcomes, there exists a cutoff point above which stopping is optimal and below which search is optimal, see for example Chow and Robbins (1963). However, when the distribution is uncertain, it's possible for this cutoff rule to fail.

Assumption 7: Given current belief α there exists discoverable strategies with payoffs y', and y'' such that $y' < y''$, $\alpha' \equiv \pi(\alpha,y') < \alpha'' \equiv \pi(\alpha,y'')$ and $f(x;\alpha') \underset{\text{FOSD}}{<} f(x;\alpha'')$

Theorem 2: Under A4-A7, there exists a discovered strategy payoff and search cost such that it is optimal to stop after discovery of a lower payoff strategy but continue searching after discovery of a higher payoff strategy.

Under uncertainty search has two effects on the expected gain from additional search: a best payoff effect and a learning effect. The best payoff effect can only make additional search less valuable; if a newly discovered strategy has a higher payoff than any previously discovered strategy, the entrepreneur is less likely to conduct further search simply because possibility of finding a better payoff than the new best has been reduced. When the distribution of payoffs is known with certainty this is the only effect of search, resulting in the simple cutoff stopping rule. However, when there is uncertainty the learning effect also is present. Search provides information about the likelihood strategies follow a particular distribution. For expositional purposes A7 used in Theorem 2 effectively isolates the learning effect but more generally when the payoff of a strategy suggests a poor draw from a good distribution rather than a good draw from a poor distribution, the learning effect can outweigh the best payoff effect.

Proposition 1: Under Theorem 2, the best strategy payoff upon termination of search may not be the last discovered strategy.

Theorem 3: Under (A4)-(A7) when $f(x;\alpha'') \underset{\text{FOSD}}{<} f(x;\bar{\alpha})$, the optimal search decision is non-monotonic in the payoff of discovered strategy.

Proposition 2: Under Theorem 3, search can stop at the discovery of a low payoff strategy, continue at a medium payoff strategy, and again stop if a high payoff strategy is drawn.

Assumption 9: Let $f(x;\alpha)$ take the form $\int g(x;\theta)h(\theta;\alpha)d\theta$ where θ is an unknown parameter and $h(\theta;\alpha)$ represents beliefs about the distribution of that parameter conditional on hypermeter α.

Assumption 10: For a given cost c, best strategy payoff z, and belief α, there exists θ' such that $h(\theta';\alpha) > 0$ where it is optimal to continue searching and there exists θ'' such that $h(\theta'';\alpha) > 0$ where it is optimal to stop searching.

Theorem 4: Under A4-A10, the entrepreneur will always be willing to pay to know the true value of θ.

Theorem 4 is an example of the convexity of Bayesian risk: knowing the true distribution allows the entrepreneur to optimize around that distribution and improve outcomes. A10 guarantees there is a case where the entrepreneur would make a different search decision based on what is the underlying distribution. That ensures there is value in knowing the true distribution since the entrepreneur can avoid the wrong decision for that particular search stage.

4.5 Discussion

This paper contains four main implications for entrepreneurial search. First, the decision to continue search is non-monotonic under uncertainty. An entrepreneur that uncovers a mediocre strategy might stop and execute that strategy, while another that discovers a superior strategy might decide to continue searching in hopes of finding an even better one. Second, search does not always terminate by taking the last discovery. An entrepreneur whose most resent search resulted in a strategy worse than the previous one might stop and execute the previous strategy, realizing the distribution of outcomes is worse than she originally believed. Third, initial beliefs matter for entrepreneurial search. An entrepreneur under or overestimates the true distribution of strategies for her idea will stop sub-optimally. Fourth, resolving that uncertainty has value. An entrepreneur would be willing to pay for information about the true distribution of strategies for her idea.

The first result argues against the straightforward application of search models in other settings to the entrepreneurial phenomena. The median wage of employees is publicly available in a firm's SEC filings; a job seeker has some ability to judge whether a salary offer is competitive or not. This stands in stark contrast to entrepreneurs where there is a great deal of uncertainty about an idea's potential. An existing target industry's size or profitability may be irrelevant if the strategy executed involves reshaping the industry's structure or spanning different industries. In contrast to traditional models of competition, the relative success of startup should not necessarily dissuade further entrepreneurial activity; perhaps better strategies for the same idea exist.

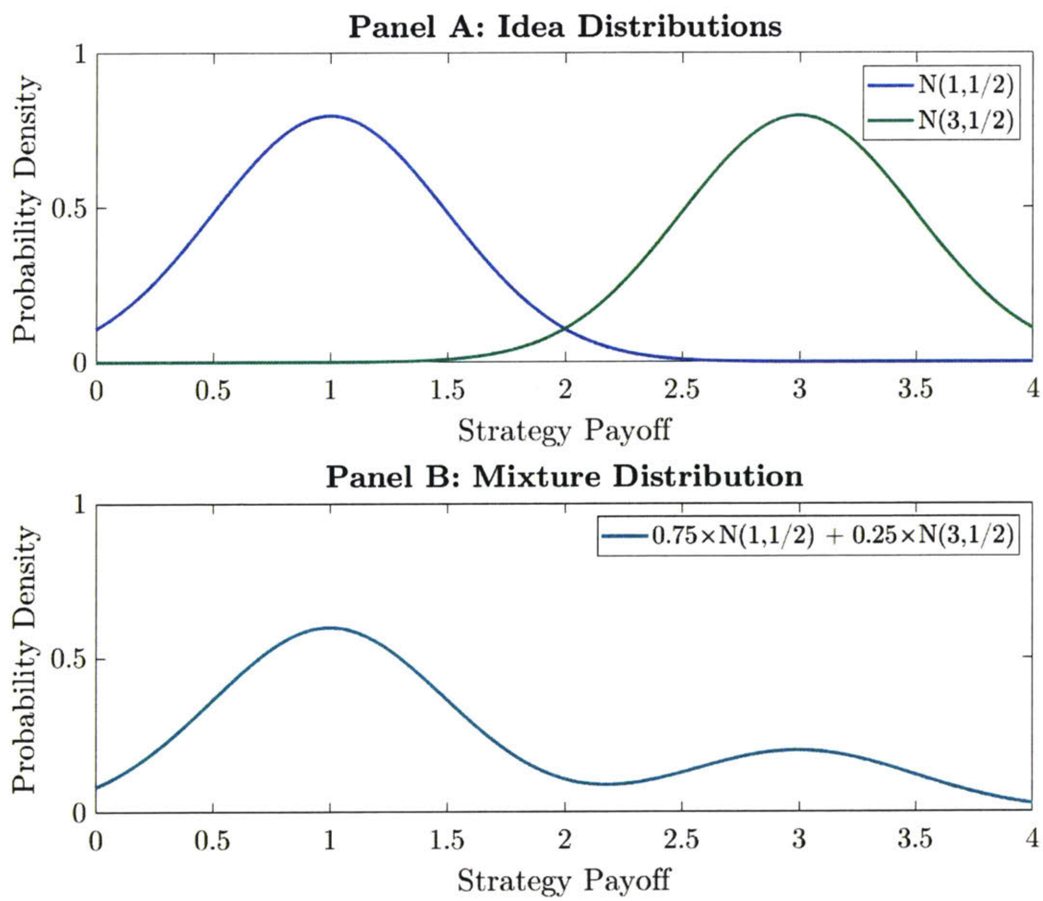
The second result has an important lesson for the lean startup approach to entrepreneurial strategy: sometimes it's necessary to un-pivot. Switching to a new strategy could involve disparaging the potential of the old strategy to investors, customers and employees. However, this could harm the startup's ability to credibility revert to the old strategy should the new strategy prove a disappointment. Pivoting could leave the startup worse off.

The last two results muddy the assumption that screening by venture capitalist has no causal effect on innovation or growth (Bernstein, Giroud and Townsend, 2016). Screening, the selection of high quality startups, is traditionally thought to imply venture capital provides no added value; the VC selected startups would have been successful (or not) anyway. But resolving the uncertainty about the potential of an idea has real value. If VC's prior experiences makes them better judges of an idea's underlying distribution than entrepreneurs, their selection process could correct the beliefs of entrepreneurs, optimally extending the search for new strategies in some cases. Selection itself could have a positive causal effect on innovation and growth.

The type of uncertainty modeled in this paper is not far from the traditional notions of a defined probability distribution; we simply relax the assumption that the true distribution of outcomes is known. Yet with this minor change recommendations for search behavior are radically different, suggesting a better understanding of decision making under uncertainty could be vital to enhancing entrepreneurship's role in innovation and economic growth.

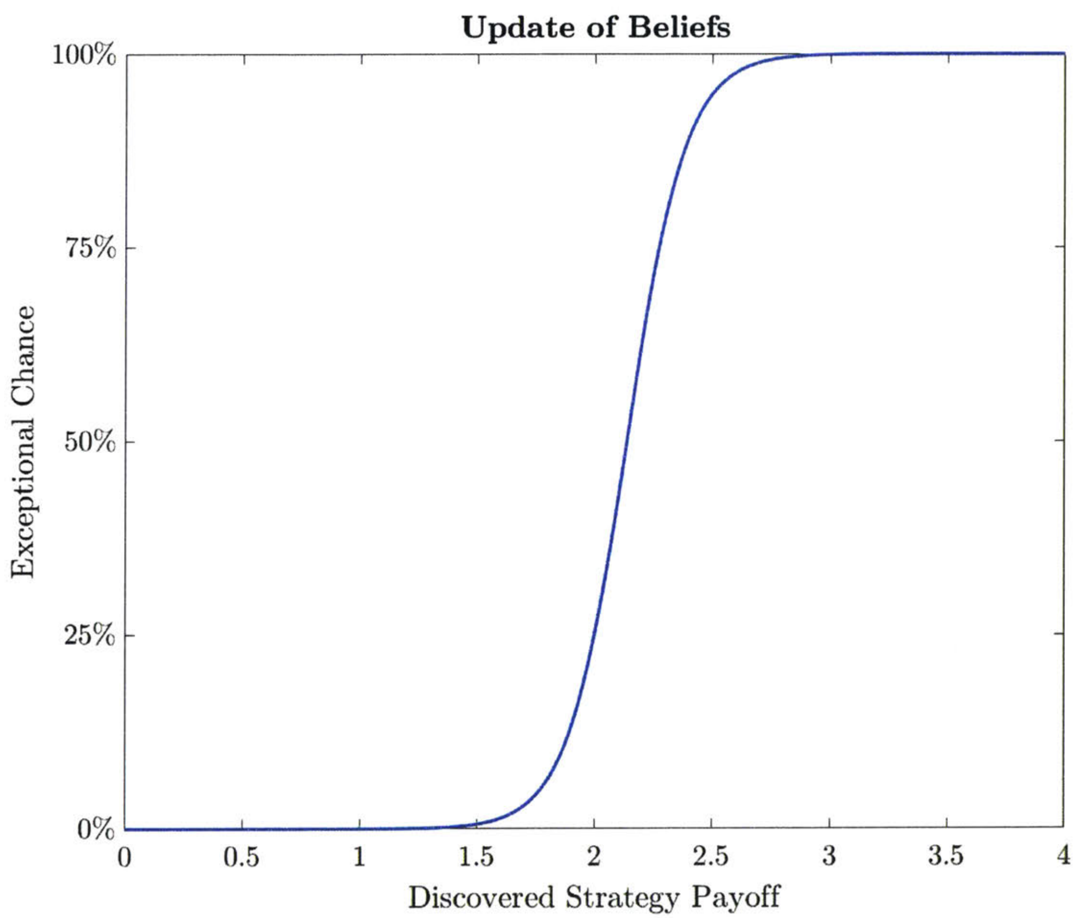
4.6 Figures

Figure 1



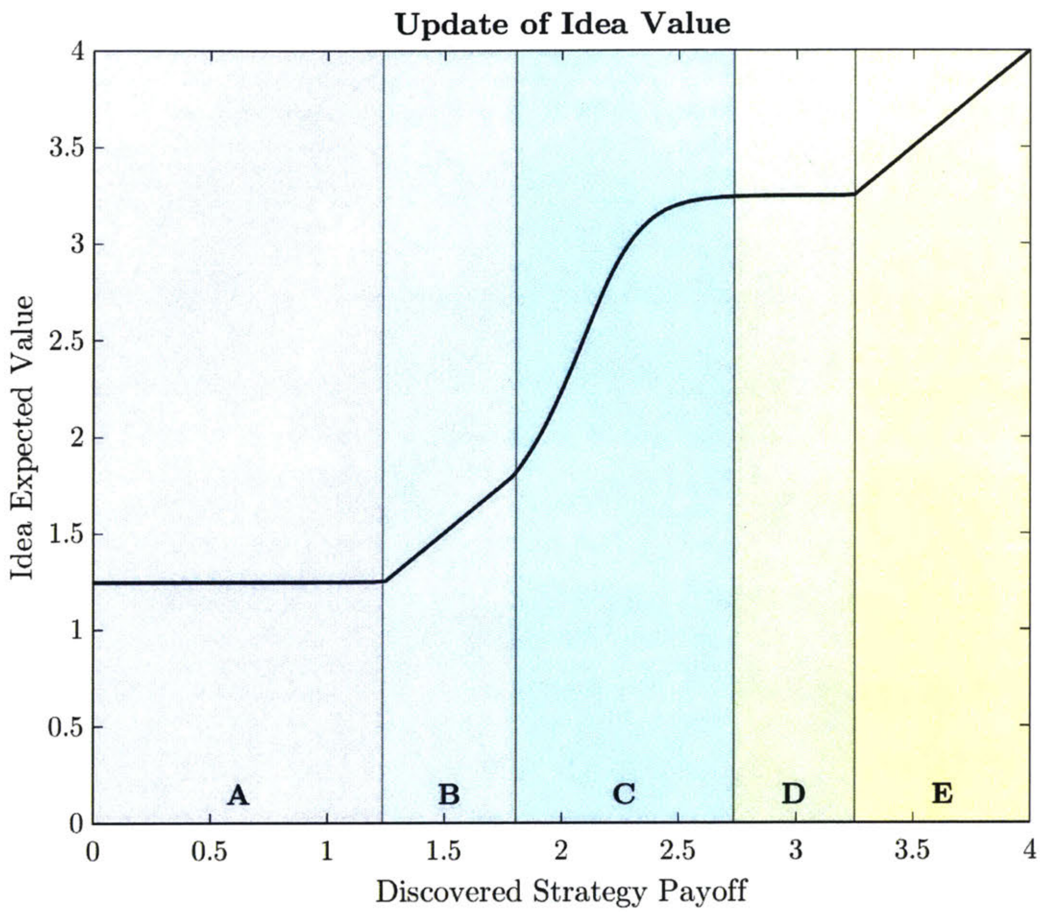
Panel A plots the probability distributions for strategy payoffs under two different types of ideas: regular ideas whose strategies have payoffs distributed $\mathcal{N}(1,1/2)$ and exceptional ideas whose strategy payoffs are distributed $\mathcal{N}(3,1/4)$. Panel B plots the mixture distribution when an entrepreneur believes there is a 25% chance her idea is exceptional.

Figure 2



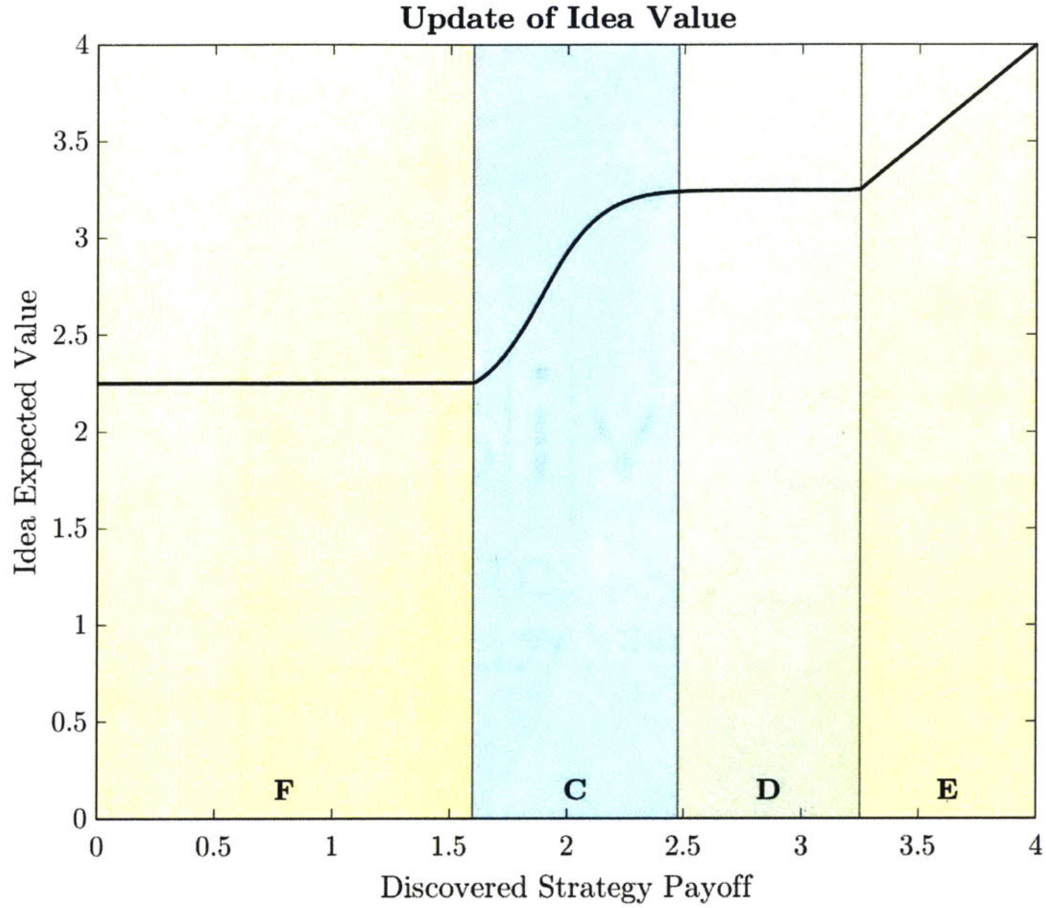
Given an initial 25% belief an idea is exceptional, Figure 2 plots how a discovered strategy changes that belief depending on the strategy's payoff. Regular ideas follow a $\mathcal{N}(1,1/2)$ distribution while exceptional ideas have a $\mathcal{N}(3,1/2)$ distribution.

Figure 3



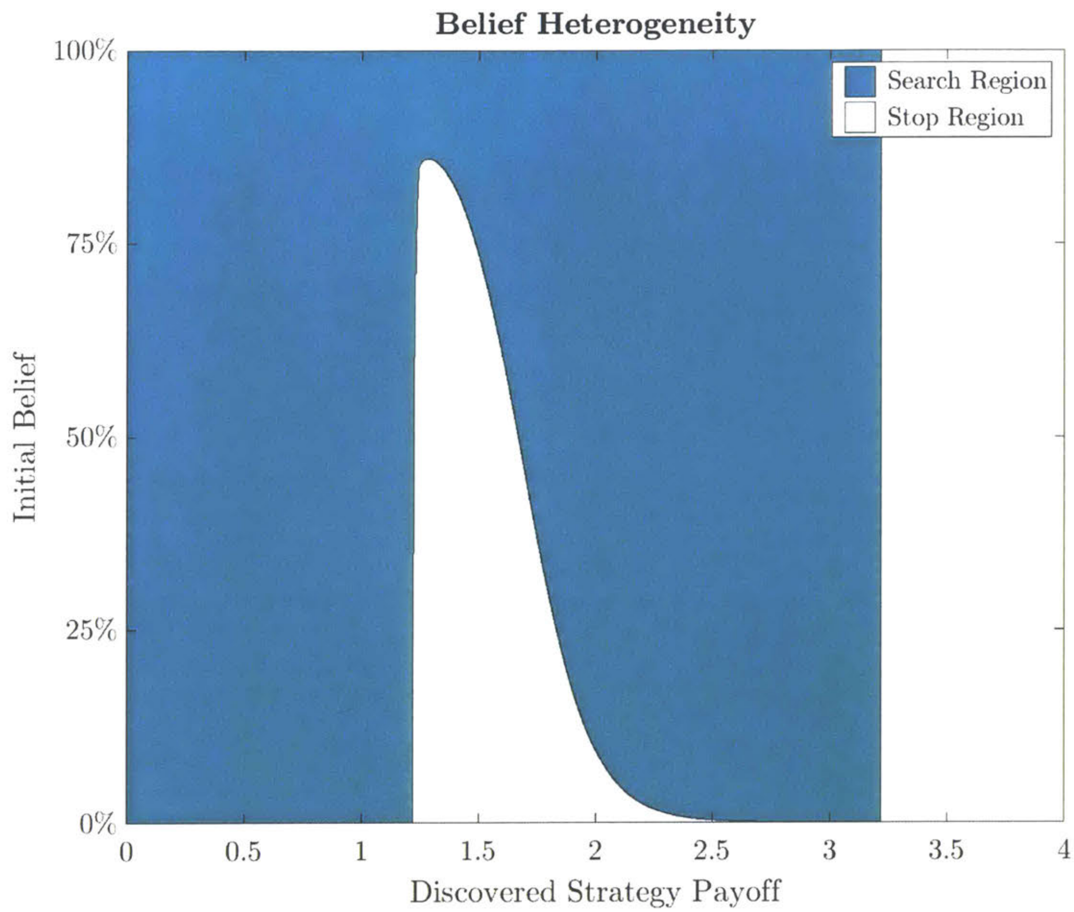
Entrepreneur initially has an outside option with payoff zero and a 25% belief that her idea is exceptional. Figure 3 plots the entrepreneur's updated value on her idea after the first strategy is discovered. Regular ideas follow a $\mathcal{N}(1,1/2)$ distribution while exceptional ideas have a $\mathcal{N}(3,1/2)$ distribution.

Figure 4



Entrepreneur initially has an outside option with payoff zero with a 25% belief that her idea is exceptional and then discovers a strategy with payoff 2.25. Figure 4 plots the entrepreneur's updated value on her idea after the second strategy is discovered. Regular ideas follow a $\mathcal{N}(1,1/2)$ distribution while exceptional ideas have a $\mathcal{N}(3,1/2)$ distribution.

Figure 5



For entrepreneurs with an initial outside option of zero, Figure 5 plots the search behavior after the first strategy is discovered as a function of the entrepreneur's initial belief the idea is exceptional and that first strategy's payoff. Regular ideas follow a $\mathcal{N}(1,1/2)$ distribution while exceptional ideas have a $\mathcal{N}(3,1/2)$ distribution.

Figure 6

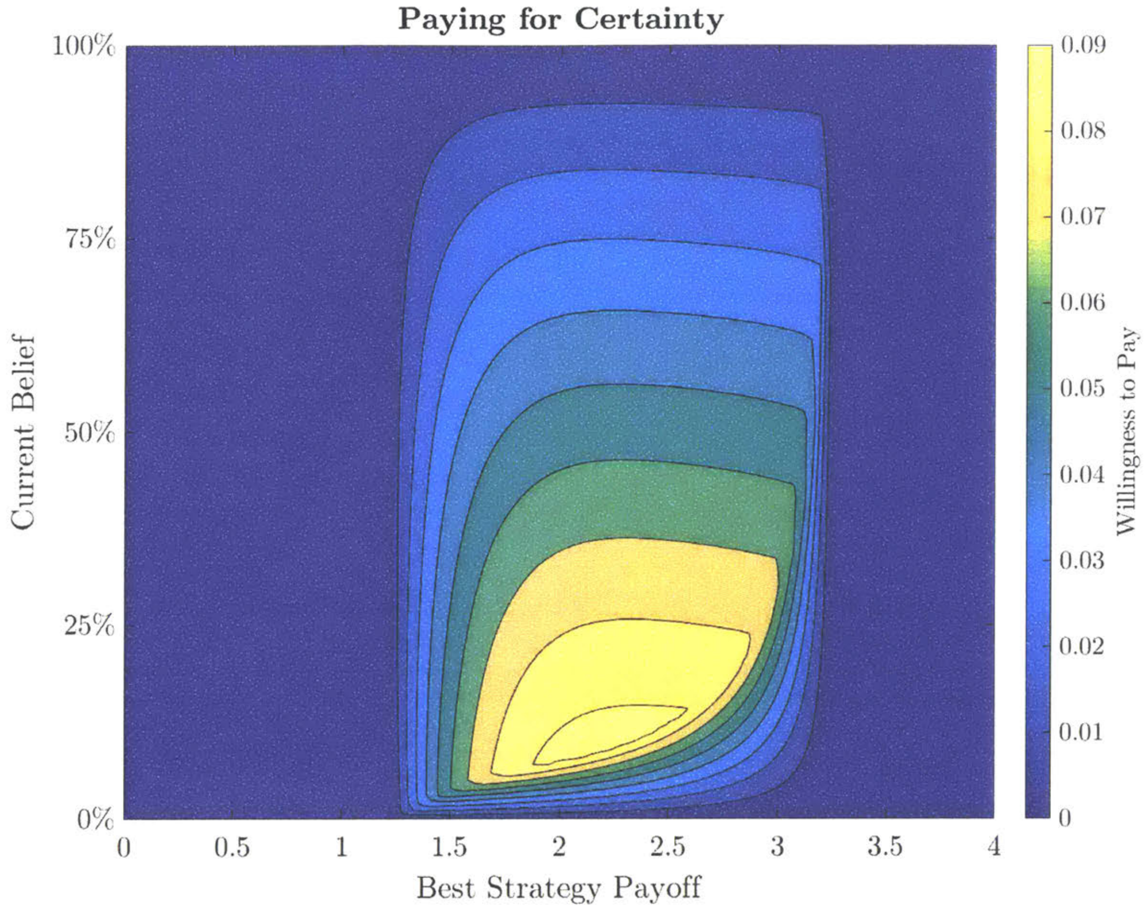


Figure 6 plots an entrepreneur's willingness to pay for knowledge of idea's true distribution as a function of the best strategy so far discovered and the entrepreneur's current belief the idea is exceptional. Regular ideas follow a $\mathcal{N}(1,1/2)$ distribution while exceptional ideas have a $\mathcal{N}(3,1/2)$ distribution.

4.7 Proofs

Theorem 1

First Part

The probability y is less than some constant u can be written as the intersection of all events where the n random variables were bounded by u plus the cost of search:

$$Pr(y \leq u) = Pr[x_1 < u + c, x_2 < u + 2c, \ldots]$$

Using marginal distributions this becomes:

$$Pr(y \leq u) = \prod_{n=1}^{\infty} Pr(x_n \leq u + nc | x_1 \leq u + c, \ldots, x_{n-1} < u + (n-1)c\,)$$

For each n we know based on A1

$$\Pr(\underline{x}_n \leq u + nc) \geq Pr(x_n \leq u + nc | x_1 \leq u + c, \ldots) \geq \Pr(\overline{x}_n \leq u + nc)$$

From Chow and Robbins (1963) Lemma 8, we know the stopping problem with independent identical distributions $\underline{X}$ and $\overline{X}$ having finite first movements has solutions $\underline{y} < \infty$ and $\overline{y} < \infty$ so can write:

$$Pr\left(\underline{y} \leq u\right) = \prod_{n=1}^{\infty} Pr\left(\underline{x}_n \leq u + nc\right) \geq Pr(y \leq u) \geq \prod_{n=1}^{\infty} Pr(\overline{x}_n \leq u + nc) = Pr(\overline{y} \leq u)$$

By the squeeze theorem this implies $Pr(y \leq u) = 1$.

Second Part

Since $\overline{X}$ dominates the marginal distributions generated in $\mathcal{P}$, we can upper bound $E[y]$ by the summation:

$$E[y] \leq \int_0^{\infty} Pr(y \geq t)dt \leq \int_0^{\infty} Pr(\overline{y} \geq t)dt \leq \sum_{t=0}^{\infty} Pr(\overline{y} \geq t)$$

Using the convergence of infinite products since $Pr(\overline{y} \geq \mathrm{x}) \in [0,1]$, taking the log, using the integral convergence test and applying integration by parts yields as an upper bound:

$$\sum_{t=0}^{\infty} Pr(\overline{y} \geq t) \leq \prod_{t=0}^{\infty}\prod_{n=1}^{\infty} F_{\overline{X}}\,(x < t + nc) \leq \sum_{t=0}^{\infty}\int_1^{\infty} \ln(F_{\overline{X}}(x < t + nc))dn \qquad \text{(L1.A)}$$

Integration by parts on the integral gives:

$$\int_1^{\infty} \ln\left(F_{\overline{X}}(t + nc)\right)dn = n\ln\left(F_{\overline{X}}(t + nc)\right)\Big|_{n=1}^{\infty} - \int_1^{\infty} \frac{nc}{F_{\overline{X}}(t + nc)} f(t + nc)dn \qquad \text{(L1.B)}$$

First consider the first term in L1.B as $n \to \infty$. Since $F_{\overline{X}} \in [0,1]$ and $n > 0$, $n \ln\left(F_{x_n}(u+nc)\right) \leq n\left(1 - F_{x_n}(u+nc)\right)$. For any distribution with a finite first moment and some non-negative number n we can write:

$$E[x] = \int_{-\infty}^{n} xf(x)dx + \int_{n}^{\infty} xf(x)dx \geq \int_{-\infty}^{n} xf(x)dx + \int_{n}^{\infty} nf(x)dx$$

Which implies

$$E[x] - \int_{-\infty}^{n} xf(x)dx \geq n[1 - F(n)]$$

As $n \to \infty$, the left-hand side goes to zero, so the right-hand side is bounded from above by zero. Since by A2 $F_{\overline{X}}$ has a finite first moment and $n > 0$, $n \ln\left(F_{\overline{X}}(t+nc)\right)$ must therefore be bounded from above by zero as $n \to \infty$.

Similarly, since $F_{\overline{X}} \in [0,1]$ and $n > 0$, $n \ln\left(F_{\overline{X}}(t+nc)\right) \geq \mathrm{n}F_{\overline{X}}(t+nc)$ and

$$E[x] = \int_{-\infty}^{n} xf(x)dx + \int_{n}^{\infty} xf(x)dx \leq \int_{-\infty}^{n} nf(x)dx + \int_{n}^{\infty} xf(x)dx$$

Which implies

$$E[x] - \int_{-\infty}^{n} xf(x)dx \leq nF(n)$$

Whose left-hand side lower bound goes to zero as $n \to \infty$ so the right-hand side is bounded from below by zero. Again since by A2 $F_{\overline{X}}$ has a finite first moment and $n > 0$, $n \ln\left(F_{\overline{X}}(t+nc)\right)$ must therefore be bounded from below by zero as $n \to \infty$, implying by the squeeze theorem that $n \ln\left(F_{\overline{X}}(t+nc)\right)$ goes to zero as $n \to \infty$.

Next for the integral term of L1.B, using a change in variables and the increasing property of $F_{\overline{X}}$, we can bound the integral by:

$$-\int_{1}^{\infty} \frac{nc}{F_{\overline{X}}(t+nc)} f(t+nc)dn \leq -\frac{1}{c}\int_{t+c}^{\infty} (s-t)\, f(s)ds$$

Therefore L1.A is bounded from above by the expression

$$\sum_{t=0}^{\infty} [-\ln\left(F_{\overline{X}}(t+c)\right) - \frac{1}{c}\int_{t+c}^{\infty} (s-t)\, f_{\overline{X}}(s)ds]$$

The first term is finite when $\bar{X}$ has a finite first moment while the second term is bounded from above by the second moment of $\bar{X}$

$$\int_0^\infty \int_{t+c}^\infty (s-t)\, f_{\bar{X}}(s) ds dt = \int_c^\infty \int_0^{s-c} (s-t)\, dt f_{\bar{X}}(s) ds = \frac{1}{2}\int_c^\infty (s^2 - c^2) f_{\bar{X}}(s) ds$$

Therefore L1.A is bounded from above whenever the second moment of $\bar{X}$ is finite. Since L1.A bounds $E[y]$ from above, $E[y]$ is therefore bounded from above as well. ∎

Corollary 1

Add a random variable X_0 which always has its outcomes mapped to the constant outside option z. Let stopping happen at stage $n = 0$ if the outside option is taken without any draws. Since X_0 is independent of the other draws, we can write

$$Pr(y \leq u) = Pr[x_1 \leq u + c, x_2 \leq u + 2c, \dots] \Pr[z \leq u]$$

If $z > u$, we can always pick another finite $u' > z$ since z is finite, making $\Pr[z \leq u] = 1$ and allowing us to follow the rest of Lemma 1 proof that $y < \infty$.

For $E[y] < \infty$ portion of Lemma 1 $\Pr(\bar{y} > t) = 1$ for the set of $t \leq z$. This modifies L1.A's middle term to:

$$\prod_{t=z}^{\infty} \prod_{n=1}^{\infty} F_{\bar{X}}(x < t + nc)$$

The change in product region does not affect the resulting calculations showing the sufficiency of a second moment on $\bar{X}$.

Lemma 1

Smith and McCardle (2002) provide in Proposition 5 a framework to determine the comparative statics for dynamic programming problems such as the value function of Equation 1. Our setup can be described in their framework as:

1) Each decision-making stage is index by k with $k = 0$ being the last stage
2) Possible actions at each stage a_k are either stop $a_k = 1$ or continue $a_k = 0$
3) A state vector x_k containing four elements: whether stopping has ever occurred $\lambda_k \in \{1,0\}$, the current best option z_k, beliefs α_k about the current distribution, and cost of continuing search c_k.

4) A reward function $r_k(a_k, x_k)$ that equals zero if stopping has ever occurred in the past, z_k when $a_k = 1$ and $-c_k$ when $a_k = 0$.
5) A transition function for the state vector $\tilde{x}_{k-1}(a_k, x_k)$ that accounts for a random element, in our case the discovered item from search. The stopping state element transitions to true whenever $a_k = 1$, the current best option is the maximum of the previous z_k and the value of the discovered item, the belief state is updated using the function $\pi(\alpha_k, x)$, and cost remains constant.
6) *Increasing* as the closed convex cone property under consideration

Conditional on action a_k, the reward function $r_k(a_k, x_k)$ is weakly increasing in λ_k, z_k, α_k and -c_k, satisfying clause (a) of their Proposition 5. By A4 and A6 the transition functions are weakly increasing in the random element as well as distribution order, and therefore satisfy clause (b) whenever A1 holds. Since both clauses have been met, our Equation 1 should be increasing in z, $-c$ and α if the value function of Equation 1 itself exists, i.e. has a solution, which was shown in Corollary 1 under A4-A5.

Corollary 2

Let α^* be the true distribution of strategies. When $\alpha' \leq \alpha^*$ we have $\pi(\alpha', x) \leq \pi(\alpha^*, x)$ from A6 which means $v(max\{z, x\}, \pi(\alpha', x)) \leq v(max\{z, x\}, \pi(\alpha^*, x))$ from Lemma 1. Therefore

$$\int v(max\{z, x\}, \pi(\alpha', x)) \cdot f(x; \alpha^*)dx \leq \int v(max\{z, x\}, \pi(\alpha^*, x)) \cdot f(x; \alpha^*)dx$$

A4 then gives us that

$$\int v(max\{z, x\}, \pi(\alpha', x)) \cdot f(x; \alpha')dx \leq \int v(max\{z, x\}, \pi(\alpha', x)) \cdot f(x; \alpha^*)dx$$

which is sufficient to show $v(z, \alpha') \leq v(z, \alpha^*)$. A similar argument holds when $\alpha'' \geq \alpha^*$.

Theorem 2

Let the current best option value z bet set to y''. For any draw $y \leq z$, stopping will occur in the following period if:

$$z > \int v(\max\{z, x\}, \pi(\pi(\alpha, y), x)) g(x; \pi(\alpha, y))dx - c$$

First note the integral is increasing in y by the properties of A1, A2 and L1.

Define

$$c^* \equiv \int v\big(\max\{z,x\}, \pi(\pi(\alpha,y''),x)\big)g\big(x;\pi(\alpha,y'')\big)dx - z$$

so the decision maker is indifferent between stopping and continuing after drawing y'', making continuing search optimal.

Under A3 and the decreasing nature of the integral there must exists a $y' < y''$ such that:

$$c^* > \int v\big(\max\{z,x\}, \pi(\pi(\alpha,y'),x)\big)g\big(x;\pi(\alpha,y')\big)dx - z$$

So that stopping occurs at y'. Therefore stopping optimally occurs at the lower y' yet search optimally continues at the higher y''.■

Theorem 3

Since by A4 the $\bar{\alpha}$ distribution has a first moment, the cutoff optimal stopping rule applies to the value function that lacks learning, see Example 5 in Robbins (1970):

$$v(z,\bar{\alpha}) = \max\{z, \int v(max\{z,x\},\bar{\alpha}) \cdot f(x;\bar{\alpha})dx - c\}$$

This function provides an upper bound to the value function of Equation 1, therefore there is an item value $\bar{z}$ above which stopping will always occur, regardless of distribution belief α. As in Theorem 2, given an existing z, let $\bar{y} = \max\{z,\bar{z}\}$. Then we have $y' < y'' < \bar{y}$ with stopping occurring at y' and above $\bar{y}$ but not at y'', therefore the decision to search is non-monotonic in discovered item. ■

Theorem 4

Define

$$v(z,\theta) = \max\left\{z, \int v(max\{z,x\},\theta) \cdot g(x;\theta)dx - c\right\}$$

$$v(z,\alpha) = \max\left\{z, \int v(max\{z,x\},\pi(\alpha,x)) \int g(x;\theta)\, h(\theta;\alpha)d\theta dx - c\right\}$$

Paying for certainty about the underlying distribution would be worth doing if for a given belief α:

$$\int v(z,\theta)\, h(\theta;\alpha)d\theta > v(z,\alpha)$$

The LHS optimizes two cases relative to the RHS. One where termination happens when $\int g(x;\theta)\,h(\theta;\alpha)d\theta dx$ generates a distribution of x that predicts a lower continuation value than under the true θ and another where continuation happens when $\int g(x;\theta)\,h(\theta;\alpha)d\theta dx$ generates a distribution of x that predicts a higher continuation value than under the true θ. ■

4.8 References

Adam, K., 2001. "Learning while searching for the best alternative." *Journal of Economic Theory,* 101(1), pp.252-280.

Bernstein, S., Giroud, X. and Townsend, R.R., 2016. "The impact of venture capital monitoring." *The Journal of Finance,* 71(4), pp.1591-1622.

Chow, Y.S. and Robbins, H., 1963. "On optimal stopping rules." *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete,* 2(1), pp.33-49.

Feldman, D., 1962. "Contributions to the 'two-armed bandit' problem." *The Annals of Mathematical Statistics,* 33(3), pp.847-856.

Fiet, J.O. and Patel, P.C., 2008. "Entrepreneurial discovery as constrained, systematic search." *Small Business Economics,* 30(3), pp.215-229.

Fiet, J.O., 2007. "A prescriptive analysis of search and discovery." *Journal of Management Studies,* 44(4), pp.592-611.

Fiet, J.O., Piskounov, A. and Patel, P.C., 2005. "Still Searching (Systematically) 1 for Entrepreneurial Discoveries." *Small business economics,* 25(5), pp.489-504.

Gaglio, C.M. and Katz, J.A., 2001. "The psychological basis of opportunity identification: Entrepreneurial alertness." *Small business economics,* 16(2), pp.95-111.

Gans, J.S., Stern, S. and Wu, J., 2016. "Foundations of entrepreneurial strategy." *Strategic Management Journal.*

Kirzner, I 1997. "Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach" *Journal of Economic Literature,* 35:1, 60-85.

Knight, F.H., 1921. *Risk, uncertainty and profit.*

Kohn, M.G. and Shavell, S., 1974. "The theory of search." *Journal of Economic Theory,* 9(2), pp.93-123.

McCall, J.J., 1970. "Economics of information and job search." *The Quarterly Journal of Economics,* pp.113-126.

Robbins, H., 1952. "Some aspects of the sequential design of experiments". *Bulletin of the American Mathematical Society.* 58 (5)

Robbins, H., 1970. "Optimal stopping." *The American Mathematical Monthly,* 77(4), pp.333-343.

Rothschild, M., 1974. "A two-armed bandit theory of market pricing." *Journal of Economic Theory,* 9(2), pp.185-202.

Schumpeter, J., 1942. "Creative destruction." *Capitalism, socialism and democracy,* 825, pp.82-85.

Shaked, M. and Shanthikumar, J.G., 2007. *Stochastic orders.* Springer Science & Business Media.

Shane, S. 2000. "Prior Knowledge and the Discovery of Entrepreneurial Opportunities." *Organization Science,* 11(4): 448-469.

Smith, J.E. and McCardle, K.F., 2002. "Structural properties of stochastic dynamic programs." *Operations Research,* 50(5), pp.796-809.

Stigler, G.J., 1962. "Information in the labor market." *Journal of political economy,* 70(5, Part 2), pp.94-105.

Talmain, G., 1992. "Search from an unknown distribution an explicit solution." *Journal of Economic Theory,* 57(1), pp.141-157.

Wald, Abraham. "Sequential tests of statistical hypotheses." *The Annals of Mathematical Statistics*, 16.2 (1945): 117-186.

Wallis, W.A., 1980. "The statistical research group, 1942–1945." *Journal of the American Statistical Association,* 75(370), pp.320-330.