DRAFT

The Downside of Real Options: Evidence from Netflix's Entry into Television

Ankur Chavda
Massachusetts Institute of Technology

July 2018

Abstract

Is experimentation worthwhile? Real options theory predicts experimentation improves outcomes by filtering out bad projects: a small investment in an experiment yields a signal informative about whether a later, larger investment is worthwhile. However, the theory of commitment makes the opposite prediction: forgoing experimentation by ceding the right to shut down a project will improve outcomes. Until now, researchers have struggled to provide empirical evidence about when experimentation is worthwhile given the theoretical tension between real options and commitment. To provide this evidence, I use a setting where there is variation in the use of experimentation across a number similar projects: television show production. Using Netflix's entry into the industry as shock for identification, I find directionally opposite results depending on the efficacy of the experiment signal. When the signal is informative, any benefit from commitment is outweighed by the filtering effect of experimentation. When the signal is not informative, the benefit of commitment dominates and experimentation worsens outcomes. I additionally provide evidence on how commitment benefits a project; experimentation seems to shifts effort away from improving outcomes towards improving the experimental signal. The lack of commitment causes an incentive distortion that is the downside of real options.

1. Introduction

Uncertainty underlies many of the challenges firms face (Cyert and March, 1963), ranging from changes in demand (Siggelkow, 2001) to new innovations (Nelson, 1961; Sorensen and Stuart, 2000; Henderson and Clark, 1990; Christensen, 1997). Uncertainty has therefore long been a core topic for strategy researchers (Ghemawat, 2002), resulting in a number of approaches to dealing with uncertainty, including experimentation. Often the decision to experiment is modeled as a real option: a small investment which, depending on what is learned over time, can optionally lead to a larger, follow-on investment (Bowman and Hurry, 1993). In the strategy literature, this real options approach has been used to recommend experimentation for project selection (Wernerfelt and Karnani, 1987), natural resource exploration (Paddock, Siegel and Smith, 1988), R&D investments (McGrath and Nerkar, 2004), technology investments (Hurry, Miller and Bowman, 1992), joint ventures (Kogut, 1991) and well as product development (Huchzermeier and Loch, 1996).

However, experimentation in strategy has understudied side effects missing from the original financial contexts where real options theory was developed (Adner and Levinthal, 2004). These side effects derive from the close connection the firm often has with the asset underlying the real option. Consider an example based on the recent literature on experiments in entrepreneurship: Nanda and Rhodes-Kropf's (2016) early stage investments as experiments for venture capitalists and Ries's (2011) "Minimal Viable Products" as an experiment for entrepreneurs. A startup could target two potential markets, a business to business market that is lucrative but suffers from a long sales cycle or a consumer market that holds less long run promise but can quickly lead to sales. Due to the uncertainty involved with building a new product, the startup receives just a few months of venture capital funding for prototype to show evidence of demand. The need to provide short term results causes the startup to target the less valuable consumer market (Stein, 1989), making both the venture capitalist and the startup worse off in the long term (Baker, 1992). The ability of the startup to influence the real option's underlying asset, in this case the new product, results in a side effect unintended by the venture capitalist. Although real options allows for an experiment to come at a cost, it does not allow the experiment to influence outcomes, in much the same way that the purchase of a stock option will involve a transaction fee but should not affect the underlying stock's price.

The startup's suboptimal decision in the example above stems from the tension between real options and another important concept in the strategy literature: commitment (Ghemawat, 1991). Commitment theory asserts that locking into to a future course of action rather than retaining the option to change can result in benefits such as preventing competitive entry (Schmalensee, 1978), fostering core capabilities (Leonard-Barton, 1992) or aligning incentives (Lerner and Wulf, 2007; Azoulay, Graff Zivin and Manso, 2011). This begs the empirical question: when do these benefits predicted by commitment outweigh the gains from experimentation as predicted by real options? Most empirical papers on real options and commitment have focused on whether firm behavior is consistent with theory (Hurry, Miller and Bowman, 1992; Ghemawat, 1993; McGrath and Nerkar, 2004). Few papers address when a specific project's outcomes would be so adversely affected by an experiment that experimentation should be avoided. MacCormack, Verganti and

Iansiti (2001) provide evidence that experimentation can improve product quality but rely on a small dataset of products as dissimilar as a web browser (Netscape 3.0) and an internet start page (My Yahoo!). Thomke (1998) restricts his sample so the projects are comparable but has variation in underlying technology rather than experimentation. The empirical challenge in asking what is best for a specific project has been in constructing a data set that is large enough for statistical analysis, contains projects that are comparable to each other, exhibits variation in experimentation and allows for identification.

To address these shortcomings, I examine the development of new television shows. The process of creating new television shows has used experimentation long before strategy researchers advocated their use in uncertain environments. Each year in the fall, writers looking to create their own show (creators) approach US television networks with ideas. Before a new show is fully funded, the networks commission the first episode, termed the pilot, akin to an experiment on that show idea. Network executives would use the pilot to solicit a quality signal, either based on their own opinions or by showing the pilot to limited test audiences (Bunn, 2002). Conditional on a positive signal, the networks provide feedback on the pilot to creators, fund the rest of the show's first season episodes and publicly broadcast those episodes, much like a firm taking a new a product to market after a favorable outcome from a trial experiment.

The incumbent US television networks on average produce 300 to 400 such pilots a year, enough to construct a dataset large enough for statistical analysis. Television shows are arguably similar to each other: they are all classifiable into either comedies or dramas, run for either a half hour or a full hour and often based on similar themes such as the anti-hero present in both *The Sopranos* and *Breaking Bad*. Not all shows have a pilot; a share of shows skip the pilot and are ordered straight to series, providing the necessary variation in experimentation. For identification I use the 2013 entry of Netflix into television show production which precipitated an increase in straight to series orders by the incumbent networks, see Figure 1. I argue my point estimates are unbiased because they are consistent before and after the shock.

[Insert Figure 1 about here]

My primary outcome variable is a show's rating in the Internet Movie Database (IMDb), which I use as a proxy for quality. Theory suggests variation in the informativeness of the experimental signal should affect the value of experimentation; I use whether any of a show's creators have a previous award-winning show as a source of such variation. My results show experimentation is valuable only when the informativeness of the signal is high, otherwise experimentation worsens outcomes. I utilize the episode structure of television shows to test what mechanisms could be behind this loss in value. I find the higher show quality derives primarily from higher quality episodes later in a show's season, a pattern that consistent with mechanisms such as misalignment of incentives (Baker, 1992) but not learning from experimental feedback (Eisenhardt and Tabrizi, 1995) or planning (Delmar and Shane, 2003).

The rest of this paper proceeds as follows. Section 2 describes how the production of television shows has characteristics necessary to study the tension between the theories of real options and commitment. Section 3 models experimentation to provide comparative statics used to interpret my empirical results. Section 4 outlines the data used for this paper. Section 5 lays out my empirical strategy to measure the value of experimentation. Section 6 provides empirical results. Section 7 discusses the paper's results and concludes.

2. Studying Television Show Production

Real options theory does not necessary apply just because an experiment is used, nor does commitment theory apply whenever an experiment is not used. For television show production to have promise as a setting to study the tension between real options and commitment for experiments, the basic primitives of both theories must be present. In addition, the decision to pilot a show must contain a random element, otherwise no causal claim that could apply to experimentation in general can be made.

2.1 Real Options

Although a definition of real options a small investment which, depending on what is learned over time, can optionally lead to a larger, follow-on investment is accurate, it may also be overly broad since it includes objects such as financial options which are not the primary focus of strategy researchers. Bowman and Moskowitz (2001) survey the empirical work on real options in strategy and induce an operational definition based on how the concept has been applied by strategy researchers:

> *A common theme in all of these decisions is that they entail the use of a two stage process: In the first stage, a small investment is made that gives the company the right to participate in the project (i.e., the company purchases the option). Some time later, when more information is known, the second stage occurs when the company faces a choice about making a larger investment in the project (i.e., the company exercises the option).*

They suggest there are couple key elements to how strategy researchers have applied real options. First, there is a project being considered by a firm, ranging from investments in other firms (Kogut, 1991; Hurry, Miller and Bowman, 1992) to product development (Thomke, 1998; MacCormack, Verganti and Iansiti, 2001). Second, two distinct investment opportunities exist, an initial one that creates an option on the project and a second one that funds the project. Third, between the first and second investments, information is learned about the project which influences the decision to invest in the second stage. Empirical settings that display these characteristics will at least be consistent with how strategy researchers have operationalized real options in the past.

The production of television shows is one such setting. First, new television shows are projects under consideration for investment by firms. The incumbent television networks have historically providing the bulk of funding for new television shows by paying for the right to first broadcast the show (Blumenthal and Goodenough, 2006, p. 203). Each fall, the incumbent television networks decide which ideas for shows receive the investment necessary to produce the show. The outcome of these investments, television shows, are products in the sense of a good for sale, albeit intangible ones. Television audiences usually "pay" for watching them by consuming advertising in a bundle, but it is also common to pay for the product more directly by for example subscribing to a cable channel or purchasing a show's DVD. Television shows are also products in the differentiated sense in contrast to commodities: although both *My Little Pony* and *Bojack Horseman* are animations about equines, a typical consumer is likely to prefer one over the other.

Second, the investment in television shows is staged. Television show production follows a yearly schedule that starts with ideas and ends with a season of show episodes. The first decision point on the part of a network is whether to take an "option" on a show idea. In return for an upfront investment, an option gives the purchasing network the right of first refusal for the script based on a show idea. After the script is created, the network can make a second investment by commissioning a pilot, a show's first episode, based on the script. Finally, the network can order the show's first season for broadcast. For more details about the stages that scripts pass through, see the Appendix.

Lastly, there is uncertainty in television production that gets resolved at each stage as information about the show is learned. Ideas for shows initially are presented to the networks as treatments: a description of the show's concept usually less than five pages long. The next stage of a show's development, a script, typically requires an investment by the network on the order of one hundred thousand dollars and results in both the text for the first episode, running 30 to 50 pages, as well as a "bible", a full description of all aspects of the show's universe. The increased content gives the network more detail about the potential for the show idea. Following the script, some shows are piloted, requiring an investment of around two million dollars and resulting in the first episode of the show's season. Beyond the network being able to now see the show idea brought to life by actors, the network is able to "test" the pilot against a live audience in a controlled environment and gain a consumer signal of the show's quality. The networks use this additional information to inform the decision of whether to "greenlight" the show: produce the rest of the show's first season for broadcast.

Television show production therefore is consistent with other settings where strategy scholars have studied real options. The show it itself is a project, piloting is a separate investment stage from a full season, and piloting reveals information about whether the full season investment is worthwhile.

2.2 Commitment

Ghemawat (1991) defines commitment as earlier choices that constrain later ones. He calls out three characteristics projects must have for investments in them to be considered commitments. First, they must be durable in the sense of being long lasting and affecting future decisions. Second, the must be specialized in that not all future courses of action are enabled by them. Third, they must be untradeable; significant frictions must prevent a firm from offloading a committed investment.

A network that skips a stage of the television production process is committing to the project in a way consistent with Ghemawat's definition of commitment. Durability manifests in television when for example ordering a show straight to series by directly greenlighting a script means one less funding slot available for the shows that are piloted. This could lead to a piloted show not getting funding even if post-pilot it appears better than the previously committed show. Shows are specialized in that a committed show that attracts a small audience will not enable the network to generate as much revenue as a show attracting a large audience. Show commitments are also not tradable between networks since the content needs and audience demographics vary even between the major national networks; a diverse comedy like ABC's *Modern Family* would be an unlikely fit for CBS due to CBS's more culturally conservative audience (Canfield, 2016).

Skipping the pilot phase by ordering a show straight to series is therefore consistent with the theory of commitment. It affects a network's later development decisions, restricts the set of actions available to the network and cannot be trivially undone.

2.3 Netflix

Since experimenting with a pilot represents a real options approach to uncertainty and not experimenting exhibits the qualities of commitment, the difference in outcomes between piloted and straight to series shows could be used to test the tension between real options and commitment. However, the fact that the network decides whether to experiment with a pilot opens the possibility of selection bias; if for example straight to series show outcomes are observed to be better than piloted show outcomes, the networks could simply be picking shows with better ex-ante properties for straight to series production. Therefore, to move beyond correlation and assert causation I need evidence that at least some shows shifted between piloted and straight to series unrelated to the ex-ante potential outcome of the show.

To that end I focus on Netflix's entry into television show production. 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: Netflix was initially only attracting show ideas that everyone else had passed on (Dawn, 2013).

The low expectations for Netflix is related 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. Another reason is related to the technological change Netflix represented as a streaming service. 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 signal. Finally, the lack of internal resources may have played a role: 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*, won both Golden Globe and Emmy awards. Afterwards, Netflix turned into a desirable place for creators to bring new show ideas (Adalian, 2013). In addition, the 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). In effect, 2013 was the year Netflix successfully broke into the prestige television industry: for the rest of this paper I will refer to the award winnings of 2013 as Netflix's true entry date rather than its initial funding of shows in 2011. This entry constituted an unexpected shock to the decision making of networks, providing my setting with the variation in experimentation necessary to make causal claims.

3. The Network's Decision to Pilot

Why do networks pilot some shows and order others straight to series? First, piloting may not be worth it. In 2008, Jeff Zucker, NBC's CEO at the time, outlined a plan to reduce piloting because the cost of pilots was not worth the amount of information they generated (Nordyke, 2008), a trade-off that suggests the network's decision making was focused on the payoffs from experimentation. Second, the network may not have the choice to pilot. After being kicked off *Two and a Half Men* for substance abuse, Charlie Sheen secured enough independent funding to fully develop his next show, *Anger Management*. This independent funding meant networks could only order the full first season or pass on the idea, an example of when supplier bargaining power could force a network to order a show straight to series (Andreeva, 2011; Dillon, 2012).

These two possible drivers of a show's production decision, payoffs and bargaining, were also cited in how the networks reacted to Netflix's entry (Adalian, 2013). Netflix's entry could have raised beliefs about the payoff from ordering a show straight to series, reducing the rate of piloting. Alternatively, Netflix's entry could have improved supplier bargaining power, again reducing the rate of piloting.

Models of payoffs and bargaining can help interpret the observed outcomes in my data, indicating when the tradeoff between the benefits of a real options approach and commitment can be identified.

3.1 Experimentation

The existing strategy literature is missing a formal model of the decision to experiment that incorporates both theories of real options and commitment. 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 accounting for the endogenous decision to experiment 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.

Let V be the final value of any successful product in a particular market and 0 be the value of an unsuccessful product. A product i is successful with probability θ_i. Creating a product has fixed cost c associated with it, regardless of success. The decision maker can choose to use real options to make a small investment on the product with incremental cost e. With probability p_i, this small investment results in a positive information signal, indicating that product will be successful. By Bayesian updating, the product has probability ${\theta_i}/{p_i}$ of being successful conditional on a positive signal. Let $\lambda_i = 1$ represent the decision to experiment on the product. The decision maker therefore maximizes:

$$\max_{\lambda_i \in \{0,1\}} \lambda_i (p_i (\frac{\theta_i}{p_i} V - c) - e) + (1 - \lambda_i)(\theta_i V - c) \tag{1}$$

The Appendix provides a generalized version of Equation 1 and proofs of comparative statics.

How might Equation 1 apply to a network's decision to pilot a particular television show idea i? Each idea likely varies in probability of success θ_i and information contained in the pilot signal p_i: a creator with a long history of successful shows may have a higher probability of success θ_i and higher chance of positive pilot signal p_i (Nussbaum, 2018). Conditional on a positive pilot signal received with probability p_i, the network will order a season of episodes to broadcast. I assume the value of a successful product for a network V is not going to vary by i: V is determined more by the network's audience size and the uncertainty lies in whether show i will be a "hit" with probability 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 payed to the creator of a show for the right to broadcast the show. I also assume show specific characteristics other than the main scripted genre categories of drama and sitcoms do not affect

e or c. 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).

Equation 1's comparative statics suggest networks are more likely make a straight to series order for show ideas of higher ex-ante quality in terms of higher θ_i or p_i. However, the observed quality of piloted shows is not necessarily worse than of straight to series shows since only piloted shows with a positive signal have outcomes in my dataset.

How would this model capture the incumbent networks' reaction to Netflix's entry? Concern about the informativeness of the pilot signal has been used by network executives in the past to justify a limited number of straight to series orders (Nordyke, 2008); perhaps for any given show beliefs about p_i increased, decreasing the usefulness of experimentation since experimentation cost e is no longer worth the lower probability of avoiding development cost c. This change in beliefs would have lowered the observed quality of straight to series shows since there would be less filtering of low quality show ideas by pilots.

However, the usefulness of the pilot signal is not the only reason given by television executives to justify straight to series. Some argue that the underlying quality of the television show is improved by commitment: that the first episode becomes better integrated with the rest of the season (Hawley, 2014), high skilled labor becomes easier to recruit ((Anonymous Emmy Nominated Producer, 2017) or the entire season benefits from an extended planning period (Adalian, 2013). Underlying these propositions is the belief that somehow the expected payoff conditional upon success changes depending on whether or not an experiment is taken: in contrast to the assumptions of real options theory, the experiment does not just generate a signal, it also influences the distribution of possible outcomes.

In the strategy literature, this concern about the real options approach to experimentation being insufficient was expressed by Adner and Levinthal (2004), who criticized real options models in strategy that assumed the option holder and underlying asset holder are two distinct entities. They suggest this separation ignores challenges to executing real options that firms face, specifically the difficultly in terminating a firm's own projects after a negative experimentation signal. But project termination is not the only challenge when the holder of the option and asset is the same. Experimentation could for example affect product values due to mis-aligned incentives (Baker, 1992): if the project team's effort to increase a positive experimentation signal detracts from the effort required to increase the value of the final product, experimentation on a project could result in lower final product quality. Another mechanism could be uncompensated worker preferences (Bonhomme and Jolivet, 2009): high quality team members may avoid projects under experimentation since experimentation reduces job security. Also, a larger investment in planning could be justified under commitment but not under real options (Milgrom and Roberts, 1990), leading to more favorable outcomes (Delmar and Shane, 2003) under commitment.

These theoretical mechanisms of course mirror the benefits of commitment expressed by television professionals and therefore may be present in my setting. Consider a version of the network's decision function of that includes a fixed, non-stochastic term V_c, an upside of commitment when V_c is positive:

$$\max_{\lambda_i \in \{0,1\}} \lambda_i (p_i (\frac{\theta_i}{p_i} V - c) - e) + (1 - \lambda_i)(\theta_i (V + V_C) - c) \tag{2}$$

Similar to Equation 1, observed differences in outcomes could be attributed to the selection of higher quality, θ_i, shows rather than any benefit to commitment, $V_C > 0$. However unlike in Equation 1, Equation 2 provides a second channel through which Netflix's entry could have increased straight to series orders. If network executives realized for example that incentives were more optimal under commitment and increased their beliefs about V_C, some shows whose value of θ_i and p_i were previous too low for a straight to series order would now be given one. This change in beliefs would have therefore again lowered the observed quality of straight to series shows if the true value of V_C was invariant.

Regardless of reason behind a shift in straight to series orders, an experimental model where the networks are selecting shows between piloting and straight to series predicts a drop in observed straight to series show quality after Netflix's entry.

3.2 Bargaining

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.

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.

More formal arguments are provided in the appendix, but 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.

3.3 Implications for Empirics

The two models above provide a set of testable hypotheses.

Hypothesis 1: Under either model, Netflix's entry would have increased the amount of straight to series production by incumbents.

Hypothesis 2A: Under the experimentation model, Netflix's entry would have reduced the observed quality of straight to series shows.

Hypothesis 2B: Under the bargaining model, Netflix's entry would have an indeterminate effect on the quality of straight to series shows.

4. Data and Measures

To empirically test the above propositions, I pool data from three sources: Gracenote, Film L.A., and the Internet Movie Database (IMDB). Gracenote, a subsidiary of Neilson 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. 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. 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.

I match shows between Gracenote and Film L.A. using a combination of show title, alternative titles, network and year. As a publicly curated dataset, IMDB often has more variation in the titles associated with each show; bigram matching is used to add the IMDB data. Table 1 provides summary statistics for the combined dataset.

[Insert Table 1 about here]

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 smaller networks (CW, Freeform, TNT, Lifetime, SyFy, Starz, A&E and TBS). When Netflix won its own Emmy and Golden Globe awards, it essentially successfully entered the business of high quality, scripted television production. By restricting my dataset to these seventeen 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 dropped. 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 21 out of 704 or 3% of shows were ordered straight to series, increasing to 14% after Netflix's entry. The share of piloted shows getting ordered to series was more stable, at 54% before Netflix's entry and 52% afterwards.

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 wone a major Emmy or Golden Globe award. This variable separates shows into two approximately even groups, those with creators having a track record of critically acclaimed success and those without.

The next three of summarized variables deal with the financing behind shows. The firms that run the networks typically also have a studio which provides financing for show development, for example NBCUniversal Inc. runs both the NBC network and NBC studios. Often, a show that is ordered by a network will receive financing from the network's parent company studio but not always, as a counterexample ABC's *Modern Family* was financed by 20th Century Fox. The share of shows funded by a network's corporate studio has been rising over time, from 48% of shows in the pre-period to 57% in the post-period.

ABC, NBC, Fox, CBS, Sony and Time Warner are considered the "big six" studios. These studios act as gatekeepers for television production; less than 20% of shows make it on the air without one of them backing a show, I mark such shows as independently funded in my dataset. An example would be *Anger Management*, a sitcom funded by the independent studio Lionsgate Television as a vehicle for Charlie Sheen to return to television after he was kicked off *Two and a Half Men* for substance abuse (Andreeva, 2011; Dillon, 2012).

Although the Los Angeles area continues to be most common location for television production, its share of total production has dropped, mostly due to competition from locations such as Georgia and Vancouver. These alternate locations are attractive either because of tax incentives or lower labor and set costs.

The adaptation, genre and length variables provide some indication of the stability of show types over my period of interest. Although there is a slight drop in comedies, the rest of these observables remain consistent, suggesting the kind of television being produced did not change due to Netflix's entry.

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. Figure 2 provides a histogram of this first season IMDB rating. Since the distribution of ratings is skewed due to the mean being much closer to the upper cutoff of 10 than the lower cutoff of 1, heteroskedastic errors are used in all this paper's regressions.

[Insert Figure 2 about here]

Table 2 shows correlation between IMDB ratings and two other measures of show performance: the renewal decision made by the networks and winning a major Emmy or Golden Globe award. 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. Part of the observed fall in award rate can be attributed to the lack of award data for shows broadcast in 2017. However, the overall trend is downward as the number of new shows each year has gone from 117 from 2008-13 to 140 from 2014-17 while the number of awards has stayed constant. IMDB ratings have a much stronger relationship with renewal than awards, suggesting it may be a better measure of a show's financial rather than critical success.

[Insert Table 2 about here]

5. Empirical Framework

Determining whether there is a downside to experimentation in television missing from real options theory could be as simple as comparing the mean outcomes for piloted versus straight to series shows. However, several biases could be present in that estimate. Perhaps most benign is attrition bias; not all piloted show get broadcast, so outcomes will be missing for those shows. This is likely to improve the average outcome of observed piloted shows as lower quality pilots will not be funded to series and therefore understate any potential downside to real options. Selection bias is of greater concern. The television networks could be picking better shows for straight to series production so any observed downside of piloting could be generated by the selection of lower quality shows for pilots. Below I outline my approach to attaining an unbiased estimate of the downside of experimentation.

5.1 OLS estimate

To start, an OLS approach could be used to estimate the relationship between piloting and outcomes.

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

In Equation 3, $FirstSeasonRating_{int}$ is the average IMDB rating of a show's first season episodes while α_n and δ_t are network and year fixed effects. X_i are additional control variables such as the length of a show, genre and track record of the show's executives. β is the coefficient of interest that indicates the improvement in show quality from funding a show straight to series rather than piloting.

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

$$\beta = E\left[\theta_i(V + V_C)\middle|\theta_i > \frac{(1-p_i)c - e}{V_c}, Z_{int}\right] - E\left[\frac{\theta_i}{p_i}V\middle|\theta_i < \frac{(1-p_i)c - e}{V_c}, Z_{int}\right] \tag{4}$$

with Z_{int} representing the controls a_n, δ_t and X_i. The selection bias manifests itself in the conditional terms $\theta_i \gtrless \frac{(1-p_i)c-e}{V_c}$; ex-ante higher θ_i shows would be produced straight to series resulting in an inability to even sign the benefit of commitment, V_c. However, if the controls Z_{int} are sufficient for the conditional independence assumption to hold, Equation 3 then estimates

$$\beta = E\left[\theta_i(V_c - \frac{1-p_i}{p_i}V)\middle| Z_{int}\right] \tag{5}$$

In Equation 5, the term $-\frac{1-p_i}{p_i}V$ is negative by construction, representing the attrition bias from the lack of observed outcome for low quality pilots not getting funded to series. Therefore, if the estimate of β is positive then necessarily V_c must also be positive. Rather than directly measure V_c, β measures the average treatment effect in changing production: how much on average are broadcast show outcomes improved by switching from piloting to a straight to series production model? This encapsulates the attrition of dropping bad shows after piloting as well as any benefit from ordering a show straight to series. When β is positive, commitment's improvement in outcomes outweighs the screening out of bad outcomes predicted by real options theory.

However, Equation 5 predicts a heterogeneous effect depending on p_i. When the chance a show will pass the pilot phase is high, the term $\frac{1-p_i}{p_i}V$ shrinks leading to a positive β whenever $V_c > 0$. Conversely when the chance a show will pass the pilot phase is low, $\frac{1-p_i}{p_i}V$ increases leading to a negative β, regardless of the sign of V_c. The tradeoff between commitment and real options changes as the value of the experimentation signal changes.

[Insert Figure 3 about here]

Figure 3 suggests one approach to explore how β might vary: 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. The data is restricted to the period prior to Netflix's entry when almost all shows were piloted. 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 p_i 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 p_i. Therefore, I use Equation 5 as an empirical test of whether variance in p_i affects the return to experimentation as predicted by Equation 1.

$$\begin{aligned} FirstSeasonRating_{int} = \alpha_n + \delta_t + \beta_S StraightToSeries_i + \beta_A AwardExec_i \\ + \beta_{SA} StraightToSeries_i \times AwardExec_i + \gamma X_i + \varepsilon_{int} \end{aligned} \tag{5}$$

The two coefficients of interest are β_S and β_{SA}. β_S provides an estimate of Equation 4 when p_i is low. The benefits from experimenting is higher because piloting will more likely weed out bad ideas so β_S should be lower than the average treatment effect measured in Equation 3. Conversely, β_{SA} estimates Equation 4 when p_i is high, implying real options is less useful in weeding out bad ideas so β_{SA} should be higher than Equation 3's β. It is possible that β_S is negative and β_{SA} is positive. Such heterogeneity in treatment effect would result in the practical recommendation that experiments are useful depending on uncertainty; when a manager is doubtful about the quality of a new product idea, experimentation is optimal. However, as uncertainty is reduced, a real options approach to experimentation could be detrimental. A manager executing on a product idea that is like other successful past ideas is better off skipping experimentation to instead capitalize on the benefits of commitment.

5.2 Matching estimate

Using Equation 5 to estimate β has two main drawbacks. First, it may not be reasonable to assume conditional independence to hold; there could be unobserved variables causing correlation between the observed variables and error term. 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 5's X_i term but not combinations of genres and that sci-fi dramas 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 drama rather than measuring the return to experimentation.

Matching estimators can mitigate this selection bias by looking at shows that are similar. 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 drama 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 using estimating how observables affect the probability a script will be piloted versus ordered straight to series

$$StraightToSeries_{int} = \alpha_n + \delta_t + \gamma X_i + \varepsilon_{int} \tag{6}$$

restricting the observations to those where the predicted $\widehat{StraightToSeries}_{int}$ lies in a region of common support, and then estimating Equation 5 within those observations, weighted by the inverse propensity score. The weights skews the estimate towards piloted shows that were predicted most similar to straight to series shows and straight to series shows most similar to piloted shows.

5.3 Netflix as a shock

The results from the matching estimator will still be biased if outcomes are not independent conditional on observables. Although for example I observe information about whether a show has an award-winning creator, I lack information about the relationship between a show's creators, something the network executives do know and could be material to the piloting decision. Such unobservables could mean that higher quality shows are being selected for straight to series orders due to factors I am unable to control for. An estimate of Equation 5's β_{SA} would then be overestimated and I would draw the wrong conclusion about relative benefits of experimentation.

To test for such selection bias, I use Netflix's entry into television as an exogenous shock to the network's decision process. Changes in estimates of my coefficients of interest before and after the shock can reveal whether selection bias is present. Under a decision model reflected in Equation 4 where selection is not fully controlled for by Z_{int}, Netflix's entry would have lowered the incumbent networks' beliefs about the informativeness of piloting (increase in p_i) or raised their beliefs about the benefits of commitment (increase in V_C). An increase in beliefs about p_i or V_C would make straight to series more attractive for any given θ_i, lowering the cut-off θ_i necessary for a show to bypass the pilot. If after Netflix's entry the distribution of θ_i remained the same conditional on observables Z_{int} and selection bias is present, the coefficients of interests in

Equation 5 estimated on observations prior to Netflix's entry should be higher than when estimated afterwards. If there is no selection bias, the estimate should remain unchanged.

No change in such an estimate could be explained by a bargaining model. Rather than considering the payoffs from piloting versus straight to series, a bargaining model would predict the networks ordered shows straight to series only when the show's executives had sufficiently high bargaining power. If prior to Netflix's entry, shows with high bargaining power were not the highest quality shows, a shock that increased the bargaining power of creators such as the entry of Netflix could have increased the share of straight to series orders without affecting observed differences in the average quality of piloted and straight to series shows.

I again use a propensity score matching to estimate Equation 7. If no selection bias is present, Equation 7's coefficients β_A and β_{SA} should be similar to the same coefficients in Equation 5. β_{SP} and β_{SAP} would be zero. If the networks were selecting higher quality shows for straight to series orders, β_{SP} and β_{SAP} would be negative, with higher estimates for β_A and β_{SA} in Equation 8 than in Equation 5.

$$\begin{aligned} FirstSeasonRating_{int} &= \beta_S StraightToSeries_i + \beta_A AwardExec_i \\ &+ \beta_{SA} StraightToSeries_i \times AwardExec_i \\ &+ \beta_{SP} StraightToSeries_i \times Post2013_t \\ &+ \beta_{SAP} StraightToSeries_i \times AwardExec_i \times Post2013_t \\ &+ \delta_t + a_n + \varepsilon_{int} \end{aligned} \tag{7}$$

5.4 Network-Year as an Observation

Another approach to measuring the extensive marginal benefit of moving a show from piloting to straight to series is to consider the average quality of a network's portfolio new shows. Each year, each network creates a set of new shows. Some of those shows are piloted, while others are ordered straight to series. Measuring the average show quality is akin to taking the expected value of equation 4. If the value of commitment V_C is non-existent (or negative), than an exogenous event like Netflix's success should not improve the portfolio of shows, regardless of whether the network is selecting better shows for straight to series. In fact, if experimentation properly stops bad ideas from being produced, the portfolio should get worse.

The diff-in-diff estimator of equation 10 would measure whether the portfolio of shows worsened or improved after being shocked by Netflix's success.

$$\begin{aligned} ShareLowFirstSeasonRating_{nt} &= \beta_S ShareStraightToSeries_{nt} + \beta_{SP} ShareStraightToSeries_{nt} \\ &\times Post2013_t + \delta_t + \alpha_n + \epsilon_{nt} \end{aligned} \tag{8}$$

In Equation 8 the outcome variable is changed from the previous $FirstSeasonRating_{int}$ to $ShareLowFirstSeasonRating_{nt}$. First, the observation is now of yearly network level outcomes.

Second, the outcome I focus on is the share of shows on the network that year that did badly, scoring less than a 6 IMDB rating. Equation 8 has far fewer observations that my previous regressions, minute differences in overall IMDB rating will be difficult to detect. The point of experimentation is to find failures early and the easiest failures to detect are likely to be shows whom are likely to score badly if ordered to series. Therefore, specifically coding for these types of shows can assist in determining the sign of the value of experimentation, at the cost of making the coefficient itself harder to interpret.

5.5 Mechanism behind value of commitment

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.

$$EpisodeRating_{eint} = \alpha_n + \delta_t + \lambda_e EpisodeNumber_e + \beta_e EpisodeNumber_e \times StraightToSeries_i + \gamma X_i + \varepsilon_{eint} \tag{9}$$

In Equation 9, 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.

The Appendix provides additional details but the trajectory of β_e would vary by underlying mechanism. If the possibility that a pilot would never be seen by an audience dampens effort from a show's cast and crew, I would observe the first episode to benefit the most from a straight to series order. If having additional time to plan out a season before production starts is important, all straight to series episodes would be of equal higher quality than piloted episodes. If the need to pass the pilot phase distorts incentives so that the pilot is improved at the expense of the rest of the season, straight to series shows should get relatively better to piloted shows after the first episode.

6. Empirical Results

Column 1 of Table 3 estimates Equation 3. Overall, we see there is no relationship between straight to series production and the first season's IMDB rating; there is no average treatment effect from experimentation across the entire population. However, Column 2 shows an underlying heterogenous effect based on Equation 5, which uses whether the show has an award-winning creator as a proxy for a higher chance of passing the pilot phase. Theory suggests that the benefit of straight to series over piloting should be higher for shows likely to pass the pilot phase. Shows with a high degree of uncertainty, i.e. those that do not have an award-winning creator, benefit from experimentation that identifies bad ideas. Shows with a low degree of uncertainty, i.e. those with an award-winning creator, are less likely to have bad ideas that experimentation can identifies. It is in these low uncertainty cases where a real options approach to decision making will suboptimally experiment. The OLS estimates are consistent with that theory. The point estimates are modest but significant in measuring the average treatment effect in the award-winning creator subpopulation: 0.325 IMDB ratings is around half a standard deviation. OLS estimates that commitment can be expected to make a show better, but not necessarily take a mediocre show and make it great.

[Insert Table 3 about here]

Columns 3 and 4 provide estimates of Equation 3 and 5 using a propensity score matching estimator. Table 4 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. Table 5 shows the main predicators of straight to series in the propensity score estimate are independent funding and a filming location outside of Los Angeles. Both these covariates could be measuring a similar mechanism; a show could for example be filmed in Vancouver to receive partial financial support from Canadian government. Either financial support from an independent studio or a government entity would enable a show to bypass the network's funding process of first piloting a show. 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.

[Insert Table 4 about here]

[Insert Table 5 about here]

[Insert Figure 4 about here]

Returning to the matching results of Table 3, Column 3 shows that in general straight to series is not correlated with outcomes as with the OLS estimator. Column 4 shows this lack of relationship masking a heterogenous effect depending on whether an award-winning creator is attached to the show. The negative point estimate on the *Straight to Series* variable can be interpreted as indicating the average outcome of shows that lack an award-winning creator are worsened when piloting is not used; the experiment is overall helpful in identifying bad ideas that should not be ordered to series. This learning from experimentation outweigh any potential benefit from commitment, hence such shows display worse outcomes when ordered straight to series. In contrast the positive coefficient on *Straight to Series * Creator w/ Award* is similar to the OLS findings that under low uncertainty, experimentation can be harmful to outcomes.

Table 6 checks whether the relationship between straight to series and having an award-executive creator changed after Netflix's entry. Column 2 uses the triple difference style Equation 7 to measure the effect's change with the matching estimator. The insignificant estimates of both *Straight to Series * Creator w/ Award * Post 2013* and *Straight to Series * Post 2013* suggest that as the networks sent more shows straight to series, the quality of straight to series shows did not drop. This is consistent with the estimates in Table 3 being unaffected by bias due to the networks selecting ex-ante better shows for straight to series production. Since there was a large increase in number of straight to series shows, from about 3% to 14%, any selection bias should have detectable; its reasonable to expect shifting from the selection of the top 3% of all shows to the top 14% of shows would have had a noticeable effect on observed quality. Such a large shift could happen without changing observed outcomes if bargaining power driving the decision to order a show straight to series and that bargaining power was largely uncorrelated with quality, at least within the set of creators with higher than average bargaining power.

[Insert Table 6 about here]

Table 7 is based on Equation 8, using a fixed effect estimator that drops time invariant terms such as the network fixed effect. In Column 1, no overall effect is seen in the quality of the network's portfolio of shows after Netflix's entry as the network's share of straight to series production increases. This is consistent with the results from Columns 1 and 3 in Table 3. When the dataset is split by whether shows had award-winning creators, the results are consistent with Columns 2 and 4 of Table 3. In Column 2 of Table 7, restricting to shows that lack an award-winning creator shows an increase in the share of shows that do badly. With the lack of a pilot to filter out these bad shows, the network's outcomes are worse. In Column 3 of Table 7, the portfolio of award-winning creator shows displays an improvement as more straight to series production is undertaken: there are fewer failures with very low quality shows on average. Since there around 80 shows ordered per year on the incumbent networks, the point estimates suggest the kind of 10% increase in straight to series precipitated by Netflix would about one more failure in shows that lacked an award-winning creator and prevented about one failure in shows with award-winning creators.

Finally Figure 5 implements Equation 9 by plotting the improvement of each episode of a commitment show's initial run relative to the first, pilot episode. The later episodes seem to be of higher quality than the first episode, suggesting the mechanism behind the upside of commitment may involve optimizing the effort; piloting incentivizes a show's cast and crew to focus on passing the pilot phase rather than make the highest quality show possible.

[Insert Figure 5 about here]

The trend depicted in Figure 5 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).

7. Conclusion

This paper provides evidence for three main points. First experimentation improves outcomes when there is uncertainty about the value of the underlying project as predicted by real options theory: it effectively avoids investment in poor projects. Second, when there is certainty about project value, the benefits of commitment outweigh any potential gains from the experimentation. Third, incentives that distort effort towards improving the experimental signal could be at the core of why experimentation can sometimes worsen outcomes.

These results can apply broadly to a number of different areas within strategy. Take venture capital investment as one example, where the experimentation manifests as the amount of funding provided to a startup. A venture capitalist should experiment on startup created by an unproven group of recent college grads by funding the startup for a short period of time; there is too much uncertainty about final outcomes. However, an experienced serial entrepreneur with a new company idea might warrant a larger upfront investment from venture capitalist that funds the startup for a longer period of time. Not only is it more efficient in a real options approach to avoid the extra costs associated with reviewing the startup’s progress when another round of funding is likely anyway, skipping the experiment could incentivize 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 real options 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 procuring a real option 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). Real options theory 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 real options: the lack of commitment.

8. Appendix

8.1 Television production primer

The production function of a television show can be broken into five stages.

Treatment

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 tangible step after someone has an idea is the drafting of 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. Since production companies and studios are deficit financed (they take a loss when a show is originally broadcast on TV hoping to recoup during syndication), there is an interest in selecting ideas that have the potential to generate the approximately 100 episodes necessary to warrant a syndication deal. High level talent has first look deals with production companies or studios. This means ideas (or if at later stage scripts) must be rejected by the production company or studio before it can go on the open market.

Script

Treatments are pitched to network executives. Agents act as brokers that match ideas to networks and gate keep access to network executives. Networks are looking for ideas that are a good match for their audience. Although there are some demographic target differences between ABC, NBC, and CBS, in general similar content would appeal to all the broadcast networks. On the other hand an idea that is suitable to Lifetime probably isn't going to be interesting to Syfy. Any treatment that is judged to have the potential be successful on the network is funded for pilot script as well as a “bible”, a full description of all aspects of the show’s universe. Scripts typically cost on the order of $100k and may have a penalty attached to them if the network opts not to shoot or air a pilot. Penalties are generally attained by producers or writers with a proven track record to incentivize the network to follow through on production rather than have the script die after purchase. Scripts (or pilots) that are abandoned by the network can be pitched to other networks after the network's option period has expired.

It's possible in a vertically integrated firm that the seller of the script and the network are part of the same firm. Decision making here is usually made separately by each side. The studio selling the script to the network made its own decision on what treatments to pitch. The network independently decides what of the studio's treatments to finance into scripts. The studio within a vertically integrated firm has the advantage of potentially having insider information about what kinds of scripts are desired by the network. The vertically integrated network can be thought of as having the right of first refusal on treatments by the “sister” studio.

Pilot

Based on the success of shows in the current season, the network has some idea of how many slot will be open for new shows the following season. For each slot, the network picks 2 to 3 scripts from its pool of optioned scripts to pilot. Pilot involves funding the production company to make the first episode of the series. Pilot is the first stage of decision making where the needs of a vertically integrated studio and network are jointly optimized by an executive in charge of both.

First Season

The network now decides which pilots to green light for production. On the extensive margin, some existing shows that executives hoped to cut could say on the air if there is enough uncertainty about whether a piloted show could do better. A non-vertically integrated network would simply rank the pilots by a quality measure and select the shows with the highest potential. A vertically integrated firm would green light pilots that clearly had strong potential regardless of the backing studio. However for pilots with less consensus on success the firm would bias towards the sister studio if having to choose between a pilots owned inside or outside the firm.

Scheduling is an important aspect of a new show's success. Networks believe placing a new show after an existing hit creates a readymade audience for the new show that greatly boosts the chances of the new show's success. A similar but muted effect occurs when a new show is placed before an existing show.

Usually the costs of shows in the network's consideration set are close enough that the relative cost of a show is immaterial to the green lighting decision. If costs are a concern, the network is more likely to reduce the number of pilots green lit rather than alter the mix of pilots green lit. Reducing the number of scripted pilots may lead the network to pick up more lower cost reality TV shows for their upcoming season.

Subsequent Season

If show has high viewership ratings there's a high likelihood of renewal. There has been cases when for example CBS had a very strong lineup so a new show was cut even though the show's ratings would have been more than enough to be renewed at another network. Contract negotiations can also cause renewal to fail; the original actor's contract is limited to 7 years by California law so talent can turn a highly profitable show like Friends into a low performer if their bargaining position is too strong. Talent can also choose to end the show after a number of years as in Seinfeld.

It's also possible for a vertically integrated firm to nudge owned shows on the margin towards renewal with the intent of recouping costs during syndication. The downside is losing a slot for a new show that might have done better on the network.

8.2 Mathematical proofs

Experimentation versus Commitment [NEW]

Using Athey (2002) to try and make a more generalized version of the next section. Still needs work.

Let $\theta_i \in \mathcal{R}$ represent the decision maker's type knowledge of project i and ω the decision marker's knowledge about the state of the world. Let $g(p;\theta_i,\omega)$ be the family of probability distribution of signals $p \in \mathcal{R}$ which can vary based on type θ_i. Let $f(v;p,\theta_i,\omega)$ be the family of distributions representing potential payoffs to the project with $v \in \mathcal{R}$. Let any upside or downside of commitment be a shift in the mean of f by $u \in \mathcal{R}$ and therefore additive to v. Let the cost of taking a real option or experimenting take the form $h(\theta_i) \in \mathcal{R} \to \mathcal{R}^+$. Let $x_i \in \{0,1\}$ represent the decision makers choice of whether to experiment, $x_i = 0$, or commit, $x_i = 1$, to project i and $x_i^*(\theta_i)$ represent the optimal choice of x_i conditional on the project's type θ_i. If

1) θ_i and ω are both ordered so that increasing either is "better" for outcomes (need to work on using ω to model a shift in the distribution of p)
2) $g(p;\theta_i,\omega)$ have the MLRP in p (θ_i and ω are independent, need to figure out how ordering family would work)
3) $f(v;p,\theta_i,\omega)$ have the MLRP in v (θ_i and p have joint distribution, need to figure out how to characterise they jointly affect f and ordering)
4) $h(\theta_i)$ is non-decreasing in θ_i

Then

1) $x_i^*(\theta_i,u,\omega)$ is weakly increasing in θ_i, u and ω

Proof:

- Commitment investment decision

Decision maker selects between invest $y_i = 1$ or not invest $y_i = 0$. Because no signal is received, we can fold the marginal distribution g into f and write the optimization problem as

$$y_i^*(\theta_i,u,\omega) = argmax_{y_i \in \{0,1\}} \int_{\mathcal{R}} y_i(v+u)f(v;\theta_i,\omega)dv$$

Since the integrand is log-supermodular, $y_i^*(\theta_i, u, \omega)$ is weakly increasing in its parameters. Conditional on u and w, this means outside of corner cases there is a cutoff θ_i^y above which investment will occur and below which investment will not occur.

- Experiment investment decision

Decision maker selects between invest $z_i = 1$ or not invest $z_i = 0$, after observing p.

$$z_i^*(p)(\theta_i, \omega) = argmax_{z_i(p) \in \mathcal{R} \to \{0,1\}} \iint_{\mathcal{R}^2} z_i(p) v f(v; \theta_i, p, \omega) g(p; \theta_i, \omega) dp dv - h(\theta_i)$$

Since the payoff part of the integrand $vf(v; \theta_i, p, \omega)g(p; \theta_i, \omega)$ is increasing in p, any potentially optimal decision function $z_i(p)$ must also increase in p, or the decision maker would be better off either not investing with the lower p or investing with the higher p. As long as there is some p for which investing is optimal and another p for which investment is not optimal (otherwise why bother with the experiment), we must have single crossing: there is some p_i^* conditional on θ_i, ω where investing happens at $p \geq p_i^*$ but not at $p < p_i^*$. So we can rewrite the optimization problem as

$$p_i^*(\theta_i, \omega) = argmax_{q \in \mathcal{R}} \iint_{\mathcal{R}^2} 1_{\geq q}(p) v f(v; \theta_i, p, \omega) g(p; \theta_i, \omega) dp dv - h(\theta_i)$$

I believe $p_i^*(\theta_i, \omega)$ is decreasing in θ_i, ω but not sure what distributional assumptions are needed to prove.

- Tradeoff decision

Using the cutoff $\theta_i^y(u, w)$ we can write the x_c decision as

$$x_i^*(\theta_i, u, \omega) = argmax_{x_i \in \{0,1\}} x_i \left[\int_{\mathcal{R}} y_i^*(\theta_i, u, \omega)(v + u) f(v; \theta_i, \omega) dv \right] + (1 - x_i) \left[\iint_{\mathcal{R}^2} z_i^*(p)(\theta_i, \omega) v f(v; \theta_i, p, \omega) g(p; \theta_i, \omega) dp dv - h(\theta_i) \right]$$

Commitment is optimal as long as

$$\iint_{\mathcal{R}^2} [y_i^*(\theta_i, u, \omega)(v + u) - z_i^*(p)(\theta_i, \omega) v] f(v; \theta_i, p, \omega) g(p; \theta_i, \omega) dp dv + h(\theta_i) > 0$$

Ignoring the u term for a moment, the bracketed term measures the difference between how commitment results in the integrand's payoff as long as $\theta_i \geq \theta_i^c$ while experimentation allows the decision maker to drop values of v that are expected to be negative conditional on p. The higher the θ_i, the lower the cut-off p for when the expected value of v will be negative, which means more states of p where there will be no difference between experimentation and commitment. When θ_i is low, many states of p have a negative bracket value and as θ_i increases, more states of p will have a zero bracket value, so the function overall is increasing in θ_i. It's also increasing in u.

Not sure about ω, there's probably a better way to model how beliefs about p changed after Netflix's entry.

Experimentation versus Commitment

Showing that piloting is submodular in p, θ and V_C when $V_C > 0$.

$$\max_{\lambda \in \{0,1\}} \lambda(p(\frac{\theta}{p}V - c) - e) + (1-\lambda)(\theta(V + V_C) - c)$$

For θ, the set of values for $\lambda \in \{0,1\}$ and $\theta \in [0,1]$ is trivially a lattice and poset. The set of values for λ are independent of and therefore weakly increasing in θ. The network's objective function f has increasing differences in $(\lambda, -\theta)$ since for $\theta' < \theta$

$$-\theta' V_C > -\theta V_C$$

$$\theta' V - pc - e - \theta' V - \theta' V_C + c > \theta V - pc - e - \theta V - \theta V_C + c$$

$$f(\lambda = 1, \theta') - f(\lambda = 0, \theta') > f(\lambda = 1, \theta) - f(\lambda = 0, \theta)$$

Therefore, by the Monotone Maxim Theorem λ is decreasing in θ.

For p, we must have the set of possible values to be $p \in [\theta, 1]$ for the post-experiment probability of success to be bounded from above by 1. These values still constitute a poset independent of λ. We have λ decreasing in p since for $p' < p$

$$-p'c > -pc$$

$$p'\left(\frac{\theta}{p'}V - c\right) > p\left(\frac{\theta}{p}V - c\right)$$

$$p'\left(\frac{\theta}{p'}V - c\right) - e - \theta(V + V_C) + c > p\left(\frac{\theta}{p}V - c\right) - e - \theta(V + V_C) + c$$

$$f(\lambda = 1, p') - f(\lambda = 0, p') > f(\lambda = 1, p) - f(\lambda = 0, p)$$

For V_C, the set of values for $\lambda \in \{0,1\}$ and $V_c \in \mathcal{R}^+$ is again a lattice and poset. The set of values for λ are independent of and therefore weakly increasing in V_C. The network's objective function f has increasing differences in $(\lambda, -V_C)$ since for $V_c' < V_c$

$$-\theta V_C' > -\theta V_C$$

$$\theta V - pc - e - \theta V - \theta V_C' + c > \theta V - pc - e - \theta V - \theta V_C + c$$

$$f(\lambda = 1, V_C') - f(\lambda = 0, V_C') > f(\lambda = 1, V_C) - f(\lambda = 0, V_C)$$

Therefore, by the Monotone Maxim Theorem λ is decreasing in V_C.

Submodular requires the cross terms to show increasing differences. For $V_C' > V_C$ and $\theta' > \theta$

$$f(V_C{}',\theta') - f(V_C,\theta') > f(V_C{}',\theta) - f(V_C,\theta)$$

$$\left[\lambda(\theta' V) + (1-\lambda)\big(\theta'(V + V_C{}')\big)\right] - \left[\lambda(\theta' V) + (1-\lambda)\big(\theta'(V + V_C)\big)\right]$$
$$> \left[\lambda(\theta V) + (1-\lambda)\big(\theta(V + V_C{}')\big)\right] - \left[\lambda(\theta V) + (1-\lambda)\big(\theta(V + V_C)\big)\right]$$

$$\theta'(V_C' - V_C) > \theta(V_C' - V_C)$$

For $V_C' > V_C$ and $p' > p$ there is no interaction term so

$$f(V_C{}',p') - f(V_C,p') \geq f(V_C{}',p) - f(V_C,p)$$

$$(1-\lambda)\theta(V_C' - V_C) \geq (1-\lambda)\theta(V_C' - V_C)$$

For $\theta' > \theta$ and $p' > p$ again there is no interaction term so

$$f(\theta',p') - f(\theta,p') \geq f(\theta',p) - f(\theta,p)$$

$$\left[\lambda\left(p'\left(\frac{\theta'}{p'}V\right)\right) + (1-\lambda)\big(\theta'(V + V_C)\big)\right] - \left[\lambda\left(p'\left(\frac{\theta}{p'}V\right)\right) + (1-\lambda)\big(\theta(V + V_C)\big)\right]$$
$$\geq \left[\lambda\left(p\left(\frac{\theta'}{p}V\right)\right) + (1-\lambda)\big(\theta'(V + V_C)\big)\right] - \left[\lambda\left(p\left(\frac{\theta}{p}V\right)\right) + (1-\lambda)\big(\theta(V + V_C)\big)\right]$$

$$V(\theta' - \theta) \geq V(\theta' - \theta)$$

Bargaining

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 $\hat{V}_s^i$.

An 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.

Determining the mechanism behind benefit of commitment

To start, consider a basic multitasking model with two actions that represent the level of effort placed in the first episode a_1 and subsequent episodes a_2. The payoff to the show's creator is the sum of effort $a_1 + a_2$ minus cost function $a_1^3 + a_2^3$. A cubic cost function is used rather than the traditional quadratic in order 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 both effort levels set at ${}^{1}/_{\sqrt{3}}$. Suppose rather than receiving payoff $a_1 + a_2$, the show creator only gets that payoff with probability $p \in (0,1)$; the network can decide not to broadcast the show.

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

Now the level effort drops to $\sqrt{{}^{p}/_{3}}$; the show creator puts less effort into the show if there is a chance the show will not be broadcast.

But the show creator can influence the probability of a show being broadcast. Suppose the network evaluates the entire show to determine the whether the show will air and $p = \frac{1}{2}a_1 + \frac{1}{2}a_2$.

$$\max_{a_1, a_2 \in \mathcal{R}^+} (\frac{1}{2}a_1 + \frac{1}{2}a_2)(a_1 + a_2) - (a_1^3 + a_2^3)$$

Now effort level rises to ${}^{2}/_{3}$. Reducing the creator's payoff by only conditionally broadcasting a show will not necessarily reduce effort if the creator can influence the probability of broadcasting the show. This model cannot predict that the overall quality of a show will be higher or lower based on if a show skips the pilot.

Networks however do not evaluate the entire show when making the straight to series order, just the pilot. Working on the rest of the series can still help the quality of the pilot by hinting at where the rest of the series will take the first episode's story, but effort put into the pilot itself becomes more impactful to the probability a show gets ordered to series. Let $\gamma \in [0,1]$ represent the relative importance of pilot effort to effort into the rest of the episodes in determining whether a show gets greenlit.

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

This leads to an asymmetric solution where the optimum effort a_1^* is increasing in γ while a_2^* is decreasing in γ.

$$a_1^* = \frac{1}{6}\left(1 + 2\gamma + \sqrt{5 - 4\gamma(1-\gamma)}\right), a_2^* = \frac{1}{6}\left(1 + 2(1-\gamma) + \sqrt{5 - 4\gamma(1-\gamma)}\right)$$

Therefore this model does predict that piloting will have a different distribution of effort across episodes than skipped pilots. Piloting implies $\gamma > 1/2$ which results in a shift in effort towards the pilot episode relative to a show that skipped the pilot. This shift does not happen with shows that skipped the pilot. Therefore as the season goes on, piloted shows should get worse relative to shows where the pilot was skipped.

More generally the same results occur with a decision model of

$$\max_{a_1, a_2 \in \mathcal{R}^+} p(a_1, a_2, \gamma) v(a_1, a_2) - c(a_1, a_2)$$

whenever payoff v is weakly increasing in its arguments, cost c is increasing fast enough so the optimum a_1^* and a_2^* are finite and the probability p of reaching the payoff show increasing differences in a_1, γ as well as $a_2, -\gamma$.

In a planning model, one can imagine planning time a and an execution time b as a shares 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 commitment benefit to all episodes would occur 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, as in the previous section's bargaining model.

The task model would 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. 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.

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

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 as long as the distributions are non-degenerate of course 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 are 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.

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.

Andreeva, N., 2011. FIRED! Charlie Sheen Axed From 'Two And A Half Men', He Fires Back & Vows To Sue. *Deadline Hollywood*, March 7th.

Angwin, J., and Vranica, S., 2006. Marketer's Tactic Signals Big Shift in a TV Ad Ritual. *The Wall Street Journal*, January 4th.

Athey, S., 2002. Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, *117*(1), pp.187-223.

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.

Blumenthal, H.J. and Goodenough, O.R., 2006. *This business of television*. Billboard Books.

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.

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.

Canfield, David. "How CBS's 2016 Lineup Catered to Trump's America." *Slate Magazine*, Slate, 29 Nov. 2016,

Christensen, C.M., 1997. The Innovator's Dilemma. Harvard Business School Press. *Boston, MA*.

Cyert, R. M., & March, J. G. (1963). A behavioral theory of the firm. Englewood Cliffs, NJ, 2. Chapters 6

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.

Dillon, N., 2012. #Won! Charlie Sheen's comeback show nears pickup, *New York Daily News*. August 10th

Duane, Daniel 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.

Eisenhardt, K.M. and Tabrizi, B.N., 1995. Accelerating adaptive processes: Product innovation in the global computer industry. *Administrative science quarterly*, pp.84-110.

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.

Ghemawat, P., 1993. Commitment to a Process Innovation: Nucor, USX, and Thin-Slab Casting. *Journal of Economics & Management Strategy*, *2*(1), pp.135-161.

Ghemawat, P., 2002. Competition and business strategy in historical perspective. *Business history review*, *76*(1), pp.37-74.

Gibbons, R. and Henderson, R., 2012. Relational contracts and organizational capabilities. *Organization Science*, *23*(5), pp.1350-1364.

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

Henderson, R.M. and Clark, K.B., 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative science quarterly*, pp.9-30.

Hirano, K., Imbens, G.W. and Ridder, G., 2003. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, *71*(4), pp.1161-1189.

Iacus, S.M., King, G. and Porro, G., 2012. Causal inference without balance checking: Coarsened exact matching. *Political analysis*, *20*(1), pp.1-24.

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

Leonard-Barton, D. 1992. Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal*, 13, 111-125.

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

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.

McGrath, R. G. and Nerkar, A., 2004. Real options reasoning and a new look at the R&D investment strategies of pharmaceutical firms. *Strategic Management Journal*, *25*(1), pp.1-21.

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.

Nelson, R. R. (1961). "Uncertainty, Learning, and the Economics of Parallel Research and Development Efforts," *Review of Economics and Statistics*, vol. 43, pp.351-64.

Nussbaum, Emily, 2018. "How Ryan Murphy Became the Most Powerful Man in TV." *The New Yorker*, The New Yorker, 8 May

Nordyke, Kimberly, 2008. "NBC's Zucker Outlines New Approach to TV Programming." *Reuters*, 30 Jan.

Paddock, J.L., Siegel, D.R. and Smith, J.L., 1988. Option valuation of claims on real assets: The case of offshore petroleum leases. *The Quarterly Journal of Economics*, *103*(3), pp.479-508.

Ries, E., 2011. *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses.* Crown Books.

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

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

Siggelkow, Nicolaj. 2001. "Change in the presence of fit: The rise, the fall, and the renaissance of Liz Claiborne." *Academy of Management Journal* 44: 838-857.

Sorensen, Jesper B., and Toby E. Stuart. 2000. "Aging, obsolescence, and organizational innovation." Administrative Science Quarterly 45: 81-112

Stein, J.C., 1989. Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *The Quarterly Journal of Economics*, *104*(4), pp.655-669.

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

Vogel, H.L., 2015. Entertainment industry economics. *Cambridge Books*.

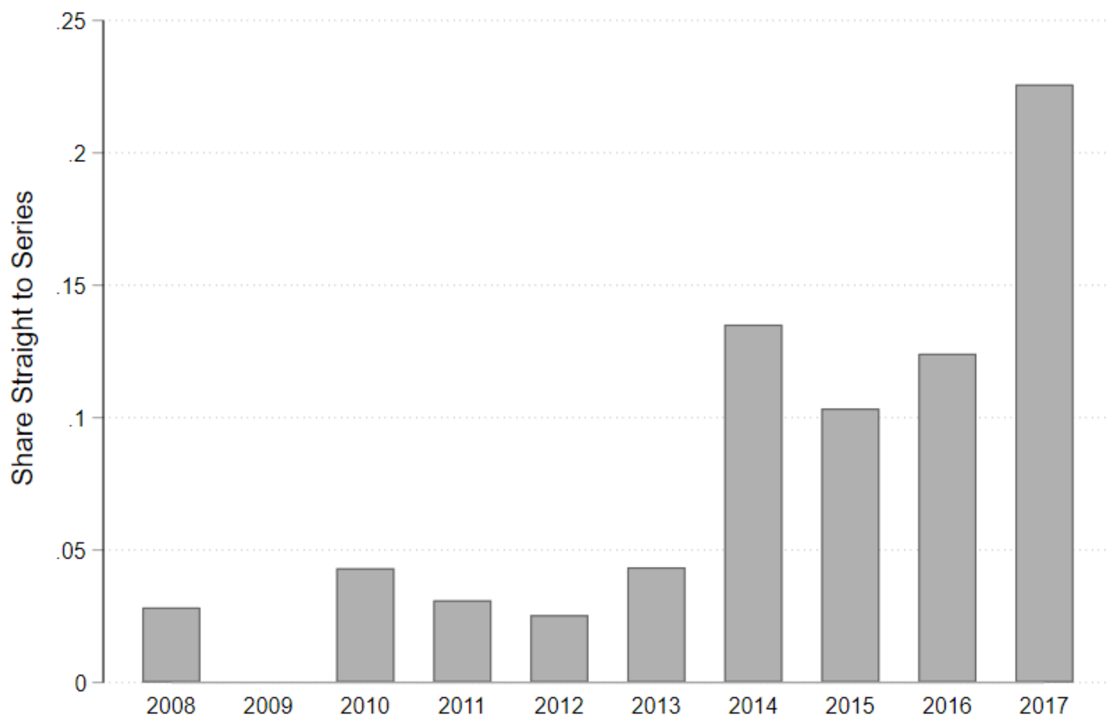
Weisman, A., 2014. The 3 Reasons Netflix Decided On Its Straight-To-Series Viewing Model. *Business Insider*. June 25th.

Wernerfelt, B. and Karnani, A., 1987. Competitive strategy under uncertainty. *Strategic Management Journal*, *8*(2), pp.187-194.

10. Figures and Tables

10.1 Figures

Figure 1. Share of Shows Ordered that Skipped the Pilot Phase



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

Figure 2. Distribution of First Season IMDB Ratings

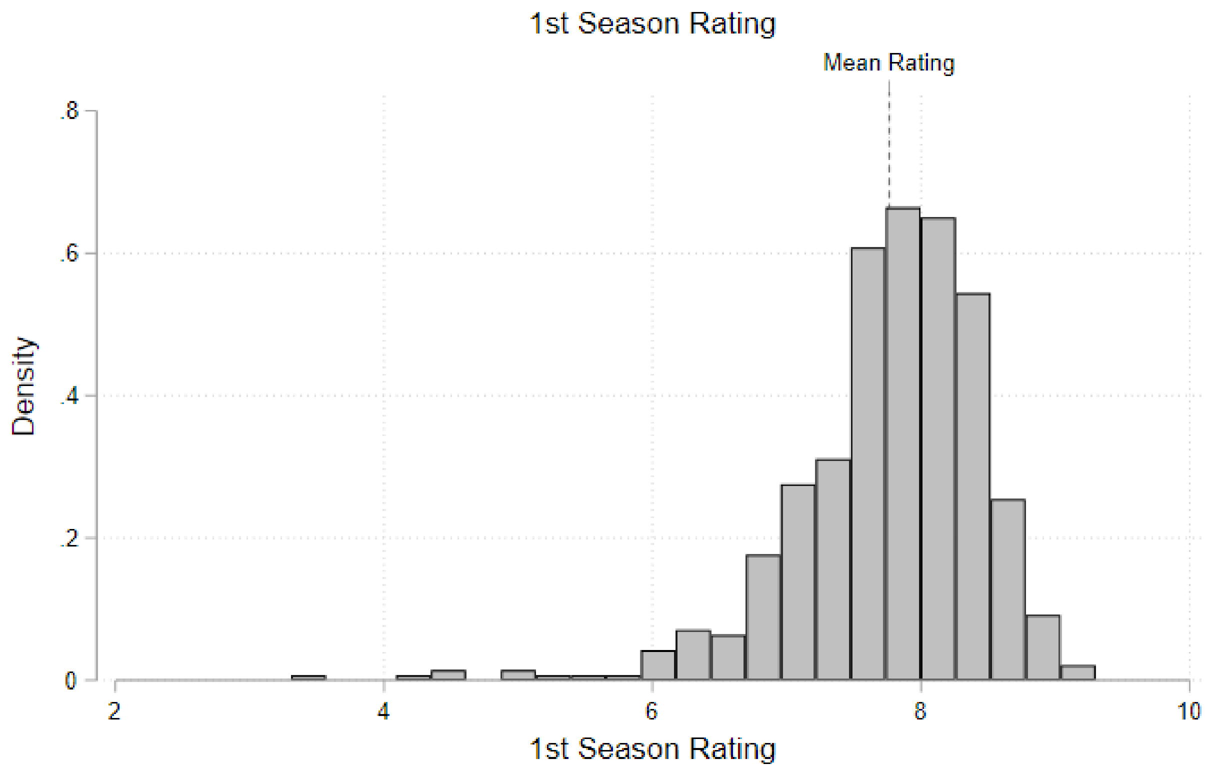
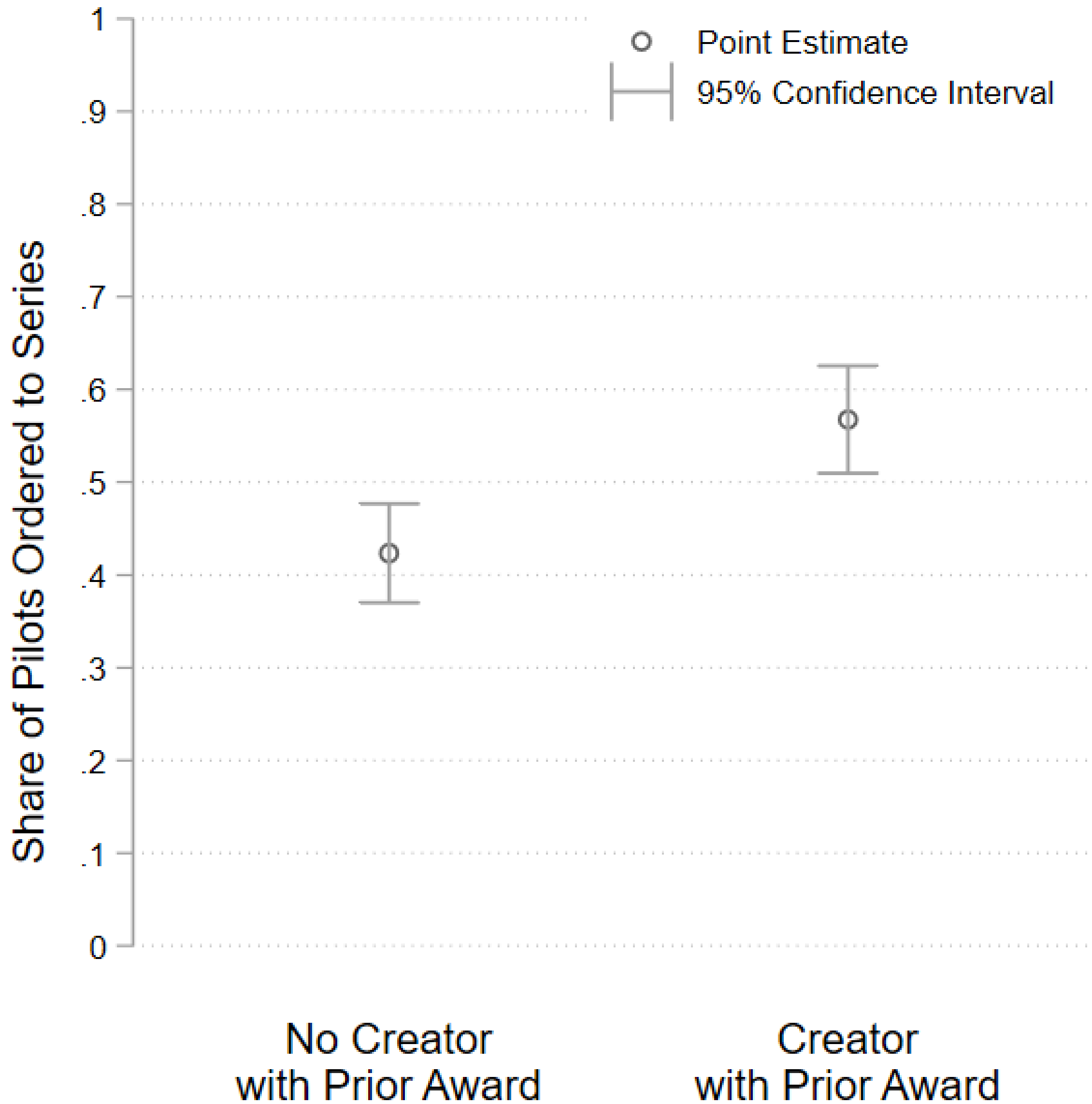


Figure 3. Rate of Shows Passing Pilot Phase



Point estimates and confidence intervals for the share of piloted scripted shows ordered to series on US incumbent networks for 2008 to 2013 by whether script was associated with executives that had a prior Emmy or Golden Globe winning show.

Figure 4. Matching Estimator Distribution of Straight to Series Predictor

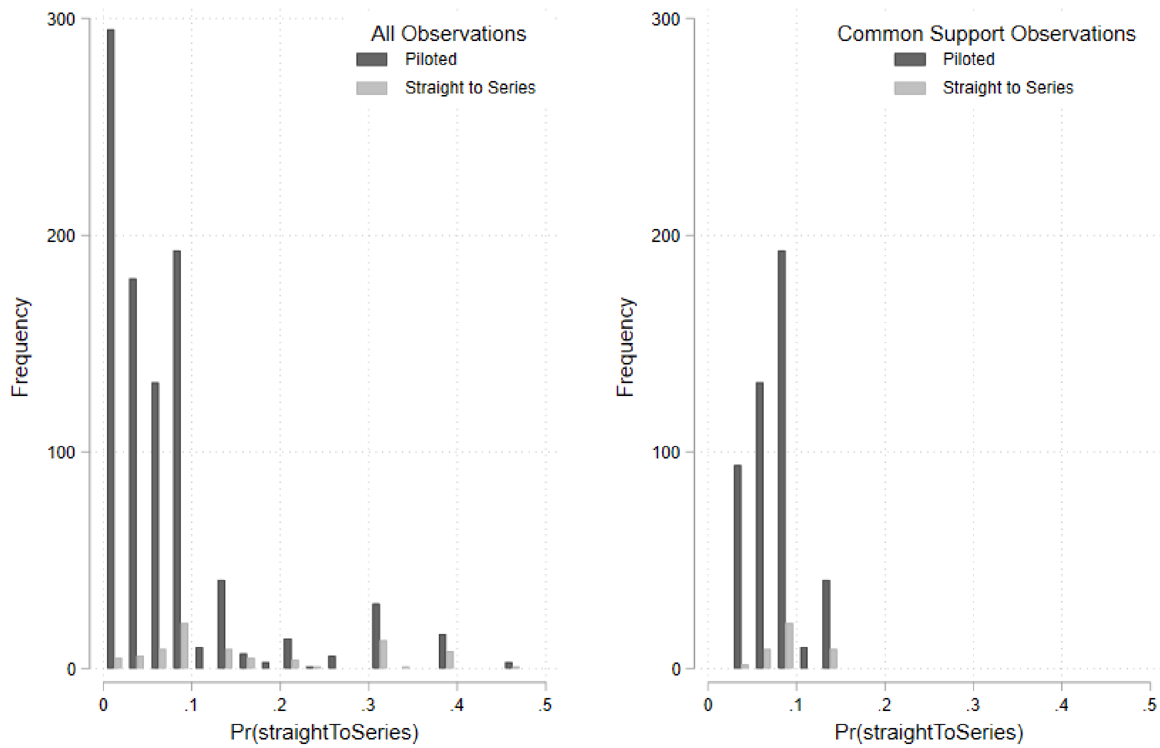
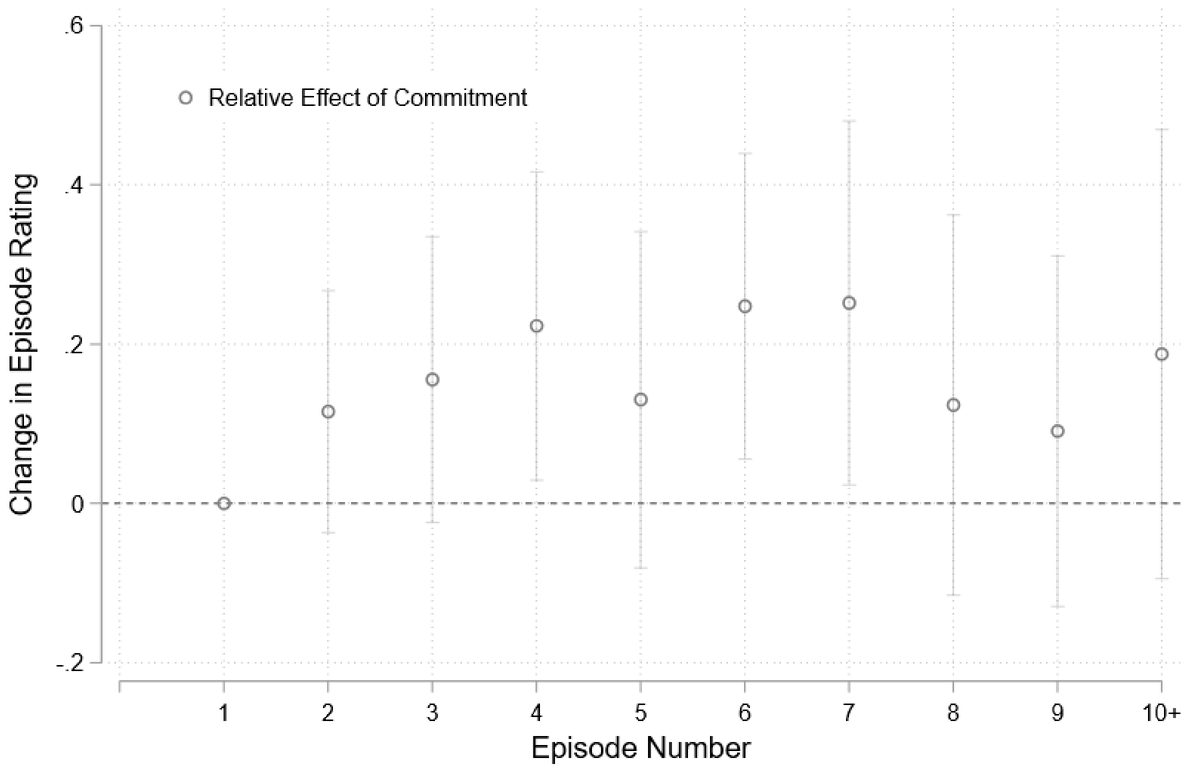


Figure 5. Episode Ratings of Straight to Series versus Piloted Shows with Award-winning Creators



10.2 Tables

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

	2008-13	2014-17	Overall
Piloted or ordered shows	704	562	1266
Piloted shows	683	481	1164
Piloted shows ordered to series	369	249	618
Straight to series ordered shows	21	81	102
Has creator with prior award	44.7%	49.3%	46.8%
Funded by broadcasting network	47.9%	57.3%	52.1%
Funded independently of major studios	16.9%	19.4%	18.0%
Filmed outside Los Angeles	42.3%	57.1%	48.9%
Adapted from other media	18.7%	18.3%	18.5%
Is a comedy	47.5%	42.3%	45.2%
Is a drama	56.7%	57.5%	57.1%
Hour long show	58.8%	59.0%	58.8%
Ordered shows	390	330	720
Mean IMDB first season rating	7.72	7.84	7.77
Standard deviation in IMDB first season rating	0.716	0.684	0.704
Renewed past initial order	63.9%	56.5%	60.6%
Winning major Emmy or Golden Globe award	4.3%	1.5%	3.1%

Table 2. Correlation between Outcome Variables

Correlations Between Show Outcome Measures

	1st Season Rating	Show Won Award	Renewed or Extended
1st Season Rating	1		
Show Won Award	0.0641	1	
Renewed or Extended	0.282***	0.161***	1

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 3. Relationship between Straight to Series and IMDB Rating

	OLS		Matching	
	(1)	(2)	(3)	(4)
	1st Season Rating	1st Season Rating	1st Season Rating	1st Season Rating
Straight to Series	0.0904	-0.0877	-0.00446	-0.310**
	[0.0796]	[0.0980]	[0.0911]	[0.148]
Creator w/ Award	0.120*	0.0717	0.251**	0.0132
	[0.0625]	[0.0694]	[0.119]	[0.0849]
Straight to Series * Creator w/ Award		0.325**		0.497**
		[0.131]		[0.206]
Shows (N)	577	577	295	295
Deg. of Freedom	144	144	112	112
Adj. R-Squared	0.102	0.107	0.237	0.264

Standard errors in brackets
Include show length, year and network FE. Clustered by network * year. Data is restricted to shows on incumbent networks from 2008 to 2017. Matching uses an inverse propensity score based weighting over a region of common support.
* p<0.10, ** p<0.05, *** p<0.01

Table 4. Covariate Balance in Full and Matched Sample

	Full		Matched	
	Piloted	Straight to Series	Piloted	Straight to Series
Funded by broadcasting network	52.4%	48.0%	63.2%	70.7%
Funded independently of major studios	15.7%	44.1%	7.7%	9.8%
Adapted from other media	17.9%	26.5%	22.1%	31.7%
Is a comedy	47.0%	24.8%	16.2%	17.1%
Is a drama	55.3%	76.2%	86.2%	85.4%
Filmed outside Los Angeles	46.7%	73.5%	71.5%	82.9%
Hour long show	56.8%	80.4%	87.2%	85.4%

Table 5. Logistic Estimation of Straight to Series from Covariates

	(1) Straight to Series
Straight to Series	
Funded by broadcasting network	0.320
	[0.245]
Funded independently of major studios	1.796***
	[0.265]
Adapted from other media	0.351
	[0.280]
Is comedy	-0.402
	[0.746]
Is drama	0.342
	[0.995]
Filmed outside Los Angeles	0.893***
	[0.324]
Hour long show	-0.152
	[0.732]
Constant	-3.669***
	[0.829]
Shows (N)	1014
Deg. of Freedom	
Adj. R-Squared	

Standard errors in brackets

Logistic estimation of straight to series used in IPW matching. No fixed effects are included. Data is restricted to shows on incumbent networks from 2008 to 2017.

* p<0.10, ** p<0.05, *** p<0.01

Table 6. Netflix Induced Change in Estimate of Straight to Series Cross Award Exec

	OLS (1) 1st Season Rating	Matching (2) 1st Season Rating
Straight to Series * Creator w/ Award * Post 2013	-0.0304	-0.359
	[0.296]	[0.333]
Straight to Series * Creator w/ Award	0.359	0.913***
	[0.238]	[0.258]
Straight to Series * Post 2013	0.184	0.264
	[0.211]	[0.288]
Creator w/ Award * Post 2013	-0.0856	-0.386**
	[0.146]	[0.175]
Straight to Series	-0.219	-0.610**
	[0.172]	[0.252]
Creator w/ Award	0.104	0.156
	[0.0795]	[0.0972]
Constant	7.707***	7.820***
	[0.0847]	[0.150]
Shows (N)	577	295
Deg. of Freedom	144	112
Adj. R-Squared	0.105	0.284

Standard errors in brackets
Includes network, show length and year fixed effects. Matching estimator uses inverse probability weights. Data is restricted to shows on incumbent networks from 2008 to 2017.
* p<0.10, ** p<0.05, *** p<0.01

Table 7. Effect of Straight to Series Production on Network's Portfolio

	All	No Creator w/ Award	Creator w/ Award
	(1)	(2)	(3)
	Share Low Rating	Share Low Rating	Share Low Rating
Share Straight to Series	0.00958	0.000248	0.0291
	[0.0448]	[0.0525]	[0.0542]
Share Straight to Series * Post 2013	0.0176	0.150**	-0.151**
	[0.0517]	[0.0644]	[0.0588]
Constant	0.122***	0.150***	0.116***
	[0.0143]	[0.0371]	[0.0302]
Shows (N)	145	120	107
Deg. of Freedom	16	16	16
Adj. R-Squared	0.132	0.128	0.126

Standard errors in brackets
Includes year fixed effects. Fixed effect estimator used clustered at network level. Data is restricted to shows on incumbent networks from 2008 to 2017.
* p<0.10, ** p<0.05, *** p<0.01