APPENDIX

Description of Custom Estimator

In the context of this study, project flexibility takes the form of piloting, with the alternative, i.e., commitment, being straight-to-series development. Testing H3a and H3b requires understanding the relationship between project flexibility and project outcomes, essentially measuring the treatment effect of piloting a television program. Ideally, calculating the treatment effect would involve comparing outcomes between television programs that were randomly assigned to either piloting or straight-to-series production. However, in practice this does not occur, for two reasons: television network executives select which programs are piloted or ordered straight-to-series, and the creators of programs select between piloted and straight-to-series offers. This could potentially invalidate any measure of treatment effect, as any observed difference between piloting and straight-to-series programs may be caused by this selection rather than the decision to pilot per se.

Although interview and anecdotal evidence suggests this selection is based on the bargaining power of a program's creators, it still might be the case that this bargaining power is positively correlated with a program's underlying quality. Hence, an econometric strategy must be applied that can account for potential selection bias. Therefore, I employ a custom estimator, which takes advantage of Netflix's market entry as a natural experiment that generated variation in the propensity of incumbent networks to pilot programs.

Netflix was distinct from other firms in that it exclusively employed a straight-to-series development process, in contrast to the piloting process common among incumbent firms. Initially, Netflix may have done so for several reasons. First, no one would have wanted to make a pilot for a new market entrant that could potentially quit television production before ordering a single full program. Second, the lack of internal resources for piloting also may have played a role since Netflix had assigned only one person to new program development (Weisman, 2014). Finally, as a streaming service, Netflix lacked the primetime slot capacity constraint of the incumbent networks. For an incumbent network, skipping the pilot increased the risk of replacing a mediocre program with one of even lower quality. For Netflix, any new offering could be added to their inventory with some long tail benefit to some of its subscribers and without the potential downside of replacing an existing program.

Note that Netflix's data capability (Smith and Telang, 2016) does not seem to have been a factor in its decision to skip piloting. Netflix's Chief Content Officer explained that "the data just tells you what happened in the past. It doesn't tell you anything that will happen in the future;" in other words, it did not allow them to pick hits (Adalian, 2013). Instead, interview evidence suggests Netflix successfully used data to match customers with content, predicting which customers might be most interested in a developed program. Using data in this way was arguably already practiced by the traditional networks, as was the case for example when NBC renewed *The Office* based on the detailed viewership data available from its distribution through iTunes (Ryan, 2006). In contrast, Netflix's main streaming competitor, Amazon, took a data-driven approach to selecting programs. However, it later abandoned this approach and adopted traditional industry practices because it was "getting chewed up," as stated by an Amazon executive (Fritz and Flint, 2017).

Netflix began funding new television program production in 2011. Many had tried and failed to break into this market, so Netflix's funding of original content did not constitute a major industry event. For example, in 2001 A&E networks commissioned two scripted dramas that it cancelled by the end of 2002 when the network retreated from scripted content production. In 2013, Netflix surprised the industry when two of its programs, *Orange Is the New Black* and *House of Cards*, garnered Golden Globe and Emmy nominations. Subsequently, as depicted in Figure A1, incumbent networks began skipping pilots for their new programs, both because Netflix's success suggested commitment could pay off, as well as the fear of losing programs to Netflix since program creators preferred straight-to-series orders (Adalian, 2013; Brembilla, 2014). This shift in incumbent use of piloting generates variation that is useful in measuring the treatment effect of piloting.

INSERT FIGURE A1 HERE

To illustrate, first consider how selection affects the estimate of piloting's treatment effect. Directly comparing piloted and straight-to-series program outcomes could lead to a biased estimate if the factors used to select programs into piloted or straight-to-series production correlate with outcomes. To illustrate, assume each program i is associated with some ex-ante quality π_i. Programs are all above a lower bound quality $\underline{\pi}$, representing the outside option for the network, typically a rerun of a previously broadcast program. Assume selection on π_i determines whether new programs are piloted or ordered straight-to-series, with some cutoff π' such that piloting occurs when $\underline{\pi} < \pi_i < \pi'$, and a straight-to-series order occurs when $\pi' \leq \pi_i$. A naïve treatment effect of piloting $T_i = 1$ on outcome Y_i would then estimate the following:

$$\hat{\beta} = E\left[Y_i \middle| T_i = 1, \underline{\pi} < \pi_i < \pi'\right] - E[Y_i | T_i = 0, \pi' \leq \pi_i]. \quad \text{(A1)}$$

The selection bias manifests in the conditional terms involving π_i when π_i correlates with Y_i: selection of programs into straight-to-series can cause straight-to-series programs to be better than piloted programs, invalidating any measure of the treatment effect of piloting.

Equation (A1) also applies when television program creators select between offers of piloting and straight-to-series production. Based on interviews with program creators and articles written in the industry press, a preference exists for straight-to-series production, since creators run the risk of having nothing to show for their effort under piloting (Bunn, 2002). Therefore, program creators may select straight-to-series offers from networks over piloting offers when given a choice. This selection by creators could result in a biased estimator in the following way. Suppose π_i represents the bargaining power of program i's creator and T_i whether that creator was able to command a straight-to-series order for the program. Creators above the cutoff π' have the bargaining power to receive straight-to-series orders. If creator bargaining power is correlated with a creator's ability to develop high-quality television programing, the treatment effect of piloting measured in Equation (A1) is similarly biased by the selection of network offers by creators.

To address this bias, I assume Netflix's market entry changed the piloting cutoff, either by altering bargaining power in favor of the creators who preferred straight-to-series production or by lowering network expectations about the value of piloting. Suppose π' drops from π_0' before Netflix's entry to $\pi_1{}'$ afterwards, as illustrated in Figure A2. Then in the period before Netflix's entry, piloted programs are observed with qualities across both ranges $\underline{\pi} < \pi_i < \pi_1{}'$ and $\pi_1{}' < \pi_i < \pi_0{}'$. After Netflix's entry, straight-to-series programs are observed across ranges $\pi_1' < \pi_i < \pi_0{}'$ and $\pi_0' < \pi_i$.

INSERT FIGURE A2 HERE

An estimator which calculates the propensity that a program is piloted could take advantage of this shift by lowering the weight on piloted observations far below $\pi_0{}'$ prior to Netflix's entry and lowering the weight on straight-to-series observations far above π_1' after Netflix's entry (Hirano, Imbens, and Ridder, 2003). Let $r = 0$ represent the regime before Netflix's entry and $r = 1$ the period afterwards. Then this method to test H3a (and similarly H3b) estimates the following equation, which measures a local average treatment effect for programs with ex-ante quality $\pi_1' < \pi_i < \pi_0{}'$:

$$\hat{\beta} = E[Y_{it}|T_i = 1, r = 0, \pi_1{}' < \pi_i < \pi_0{}'] - E[Y_{it}|T_i = 0, r = 1, \pi_1{}' < \pi_i < \pi_0{}']. \quad \text{(A2)}$$

In contrast to equation (A1), in equation (A2) I avoid selection bias by comparing programs of similar quality before and after Netflix's entry.

The estimator I use for (A2) is weighted as follows. Let $f(X_i) = \Pr(T_i = 1|X_i)$ be the probability of piloting based on observed predictor covariates X_i. For equation (5), straight-to-series observations in the post-Netflix period are inversely weighted by $1 - f(X_i)$. This weighting places more emphasis on observations close to $\pi_1{}'$ and increasingly less emphasis on observations far above $\pi_1{}'$. Since $\pi_1' < \pi_0'$, observations above the previous cutoff of π_0' receive little emphasis. Similarly, piloted observations in the pre-Netflix period are inversely weighted by $f(X_i)$, causing piloted observations in the range $\pi_1' < \pi_i < \pi_0'$ to receive greater weight relative to those below π_1'. The region of common support is defined so that differences between observables are minimized across piloted and straight-to-series programs.

Column 1 of Table A1 lists the results of estimating the probability of straight-to-series production, $1 - f(X_i)$. The strongest predictor for straight-to-series production is having outside funding, or in other words not being *Major Studio Funded*, as the point estimate is -0.993. Table A2 provides the covariate balance results between piloted and straight-to-series programs for the predictor variables (i.e., X_i), including a dummy *Filmed in LA* to reflect the location fixed effect. Following equation (A2), piloted (straight-to-series) programs are restricted to the period before (after) 2013. Panel A shows the difference in variable means between piloted and straight-to-series programs for all new programs on incumbent networks. All the predictor variable means are significantly different between piloted and straight-to-series programs for these observations. Panel B of Table A2 shows the results for the unweighted region of common support using the probit predictor of straight-to-series production. In this subsample, none of the predictor variables show significant differences between the piloted and straight-to-series observation means. Finally, Panel C of Table A2 further restricts the sample by excluding observations that lack *IMDb Rating*, the outcome variable used in the main results. The removed observations are mostly pilots that were not ordered to series, but also includes programs present in the FilmLA dataset that simply failed to correspond to a unique program in the IMDb dataset.

INSERT TABLE A1 HERE

INSERT TABLE A2 HERE

Figure A3 provides a visualization of the calculated propensity values. Panel A plots the predicted probability of a straight-to-series order for all programs, while Panel B plots only the subset of programs that were in the region of common support—those programs within the 5th and 95th percentiles of all programs regarding their probability of being straight-to-series. Due to this region of common support, my estimates represent the difference in outcomes between programs with a low probability of straight-to-series order. If straight-to-series orders strongly correlate with inherent ex-ante program quality, this study essentially estimates the local average treatment effect of piloting middle-quality programs rather than the best or worst programs on the network.

INSERT FIGURE A3 HERE

However, an estimator based on equation (A2) alone would compare programs across time periods, which may introduce non-stationary bias: the types of programs on incumbent networks may have changed after Netflix's entry, which would be the case if, for example, consumer expectations for program quality shifted and this was reflected in the IMDb ratings or if Netflix attracted the industry's strongest talent after market entry, lowering the quality of new programs on incumbent networks. Assuming selection bias is accounted for, equation (A2) can be rewritten in the following way for observations in the range $\pi_1' < \pi_i < \pi_0$:

$$\hat{\beta} = \underbrace{E[Y_{it}|T_i = 1, r = 1] - E[Y_{it}|T_i = 0, r = 1]}_{\textbf{Treatment Effect}} + \underbrace{E[Y_{it}|T_i = 1, r = 0] - E[Y_{it}|T_i = 1, r = 1]}_{\textbf{Non-stationary Bias}}. \tag{A3}$$

The first difference is the treatment effect of interest as it constitutes the difference between piloted and straight-to-series observations in the period after Netflix's entry. The second difference is a bias term, which represents how outcomes for piloted programs differed before and after Netflix's entry. An estimate of the treatment effect using equation (A2) would be biased upward if piloted programs worsened between the first and second period and biased downward if they improved. Again, a similar estimator is employed to test for such bias by estimating the second difference in equation (A3).

Let $f(X_i) = \Pr(r_i = 1|X_i, T_i = 1)$; now weights are a function of the probability a piloted program appeared in the post-Netflix period. Column 2 of Table A1 presents the result of estimating $f(X_i)$. Piloted programs observations in the pre-Netflix period are inversely weighted by $1 - f(X_i)$. In effect, the piloted programs in this period that were most like the piloted programs in the post-Netflix period receive higher weights. In contrast, piloted programs in the post-Netflix period are inversely weighted by $f(X_i)$, placing higher weights on pilots most like those in the pre-Netflix period. Observations used in this estimator are restricted to those within a region of common support.

To finally calculate the treatment effect of piloting presented in Table 7, I adjust the selection-bias-corrected measurements provided by the estimator for A2 with the measurement of non-stationary bias calculated above. Hence, Table 7's estimates should be devoid of both types of bias. The two measurements, one of equation (A2) and another of the bias factor of equation (A3), are presented in Table A3. Columns (1) to (3) provide the results, unadjusted for non-stationary bias. The *Piloted* coefficient in Column (3), for example, represents the difference in IMDb ratings between piloted half-hour programs in the pre-Netflix era and straight-to-series half-hour programs in the post-Netflix era. The *Post 2013* coefficient in Column (6) represents the difference between piloted half-hour programs in the post-Netflix period and piloted half-hour programs in the pre-Netflix period. By adding these two coefficients, I mechanically calculate the difference between piloted half-hour programs in the post-Netflix period and straight-to-series half-hour programs in the post-Netflix period. This calculation is reflected in the point estimates presented in Table 7 of the main results and is also done for *Piloted* × *Hour Long*, with *Post 2013* × *Hour Long* as the corresponding non-stationary bias adjustment.

INSERT TABLE A3 HERE

One challenge in this approach is calculating standard errors. To my knowledge, no commonly accepted method exists to generate standard errors for an estimator calculated by comparing regression results across differently weighted sub-samples, where some observations are in both regressions and others are only in a single regression. As a result, for expositional simplicity, the *p*-values presented in Table 7 have been tabulated by adding the variance of each coefficient. More specifically, the *p*-values of *Piloted* presented in Columns (4), (5), and (6) of Table 7 were the result of adding the variance of the selection-bias-adjusted estimate of *Piloted* and the variance of the estimate of non-stationary bias, *Post 2013*. Similarly, the *p*-value of *Piloted* × *Hour Long* presented in Column (6) of Table 7 was the result of adding the variance of the selection-bias adjusted estimate of *Piloted* × *Hour Long* and the variance of the corresponding estimate of non-stationary bias, *Post 2013* × *Hour Long*. Unfortunately, as I could not calculate a co-variance matrix containing both coefficients, I could not further adjust the p-values for co-variance between these coefficients. This may result in the p-vlaues being too small if there is positive co-variance between the pairs of estimates.

An alternate way to test H3a and H3b relies just on Table A3, which provides the separate, properly calculated p-values using robust standard errors for the selection-bias-adjusted and non-stationary coefficients. First, as a baseline, from the *Pilot* coefficient of Column (3), we can infer that half-hour piloted programs prior to 2013 fared better than similar straight-to-series programs after 2013 (p-value 0.002). As the coefficient on the non-stationary bias estimate of *Post 2013* was positive as presented in Column (6), this means the half-hour straight-to-series programs after 2013 were in fact better than comparable ones before 2013, so the treatment effect of piloting half-hour programs is likely underestimated in Column (3). Next, H3a predicts the treatment effect of hour-long series will be lower. This is supported by the *Piloted* × *Hour Long* coefficient in Column (3), which is strongly negative (p-value 0.002) and compares hour-long pilots before 2013 to similar hour-long straight-to-series programs after 2013. Additionally, the non-stationary bias estimate for hour-long programs, *Post 2013* × *Hour Long* presented in Column (6), is also negative, which suggests the corresponding hour-long straight-to-series programs before 2013 were in fact better than those after 2013, which would imply the *Piloted* × *Hour Long* coefficient in Column (3) overestimates the treatment effect: that in fact the piloting had an even worse effect on IMDb ratings for hour-long programs. Hence the non-stationary bias works overall in the favor of H3a's prediction. For H3b, using just the results presented in Table A3 leads to the same conclusion as in Table 7, that H3b's prediction is not supported. The differences between *Piloted* and *Piloted* × *Hour Long* coefficients listed in Column (3) are not statistically significant and applying the non-stationary bias adjustment coefficients from Column (6) does not meaningfully change that evaluation.

Robustness Checks

Alternate Outcome Variables

As an alternative test of H3b, the *count* of IMDb ratings—called "votes" in the IMDb database—can be used as an outcome variable in lieu of IMDb ratings. Although IMDb ratings capture a sense of a program's quality and, as I argue above, correlate with numerous other outcome variables, this does not directly represent the commercial success of a program. The count of a program's IMDb ratings, on the other hand, is perhaps a better proxy of a program's audience size as only programs seen by a wide audience are likely to garner a high number of votes. Therefore, I test hypotheses H3a and H3b, which pertain to investment outcomes, using IMDb votes as a robustness check on the main findings, which use IMDb ratings. The results of this robustness check are presented in Columns (1) to (3) of Table A4.

INSERT TABLE A4 HERE

Although not as statistically significant, the estimated coefficients using IMDb votes as an outcome variable are consistent with the estimates generated with IMDb ratings as an outcome variable. This is perhaps unsurprising as ratings and votes are highly correlated.

In addition, Columns (4) to (6) of Table A4 run a similar analysis with a program's renewal as the dependent variable. Here we do not see any evidence in support of H3a and H3b. There are several reasons why this might be the case. First, the dataset may be underpowered to measure any effect on renewals, especially given the clustering used in my analysis. The binary variable *Renewed* has a standard deviation of 0.50, larger relative to the potential effect size of 0.10, or a 10% change in renewal rate. This contrasts with the measured effect sizes on both ratings and votes, which were larger than a standard deviation for their respective variables. Second, the renewal decision itself has become less of an indicator for a program's success in recent years. Consider the 2017 limited series *Twin Peaks: The Return*, a continuation of the original program from 1990. There was no intention for the program to run for more than a single season and this, therefore, was a failure under a renewal metric. Yet it was viewed by its network as a significant success, driving unprecedented levels of new subscriptions to Showtime (Thaxton, 2017). So even the lack of a relationship between the use of piloting and renewal rates could not rule out piloting having some effect on the underlying quality of television programs.

Actor Casting Timeline

In my discussion of the results on veteran actors, which test the hypotheses pertaining to resource allocation, I assumed that actor casting occurs after the pilot or straight-to-series order decision: resource allocation happens after the decision to retain flexibility in a project. Under this assumption, the results suggest that veteran actors are less likely to attach themselves to programs that are piloted versus those that are ordered straight-to-series. I interpret this finding as supporting H2a: the risk of project termination during the experimental phase hampers a project's ability to attract necessary resources. However, if actor casting occurs prior to the piloting decision, an alternate explanation exists. Veteran actors may simply attach themselves to better scripts, and better scripts are the ones that are ordered straight-to-series. This selection of scripts, therefore, would result in fewer veteran actors in piloted programs than straight-to-series programs, without the piloting decision per se being the cause.

To check for this possibility, I consider the effect this selection would have on the distribution of veteran actors within piloted programs. If veteran actors can distinguish at the script level between the quality of a program that will be piloted versus ordered straight-to-series, they should also be able to distinguish the quality difference between piloted programs that are passed over versus ordered to series. Hence, we should see a difference in the propensity of veteran workers associated with failed pilots versus successful pilots.

Figure A4 plots the share of programs with veteran actors by whether it was a failed pilot, successful pilot, or straight-to-series order. First, I note that as expected by H2a, we see a gap between straight-to-series programs and pilots in general. Second, there is no material difference within pilots: pilots that were eventually ordered had the same likelihood of attracting a veteran actor as those that were not ordered to series. This runs counter to the alternative explanation of actor selection: we lack evidence that these veteran actors are selecting the better scripts that eventually become better pilots. Therefore, I view the alternative explanation as unlikely, and argue that the flexibility mechanism behind H2a is plausible.

INSERT FIGURE A4 HERE

TABLE A1 Propensity Estimators

	(1) Straight-to-Series	(2) Post 2013
Producer Count	0.032	0.223
	(0.711)	(0.000)
Sister Studio Funded	0.200	0.266
	(0.119)	(0.001)
Major Studio Funded	-0.993	-0.098
	(0.000)	(0.492)
Experienced Producer	-0.065	-0.133
	(0.661)	(0.194)
Network Producer	0.144	-0.143
	(0.398)	(0.215)
Award Producer	0.067	-0.041
	(0.679)	(0.713)
Constant	-1.097	-0.283
	(0.018)	(0.376)
Location FE	X	X
Network FE	X	X
Programs (*N*)	1453	1260
Pseudo R^2	0.362	0.071
Sample	Full	Pilots

Probit regression with *p*-values in parenthesis based on standard error terms.
Observations of programs from 2008 to 2017.

TABLE A2 Covariate Balance of Observables

Panel A: Pre-2013 Pilots and Post-2013 Straight-to-Series Observations

	Piloted		Straight-to-Series			
	N	Mean	*N*	Mean	Diff.	*P*-val
Producer Count	740	3.36	175	3.52	-0.16	0.00
Sister Studio Funded	740	0.49	175	0.36	0.13	0.00
Major Studio Funded	740	0.81	175	0.43	0.38	0.00
Experienced Producer	740	0.50	175	0.42	0.08	0.06
Network Producer	740	0.18	175	0.13	0.06	0.07
Award Producer	740	0.26	175	0.21	0.06	0.12
Filmed in LA	740	0.62	175	0.25	0.37	0.00

Panel B: Common Support

	Piloted		Straight-to-Series			
	N	Mean	*N*	Mean	Diff.	*P*-val
Producer Count	187	3.51	35	3.51	-0.00	0.99
Sister Studio Funded	187	0.55	35	0.51	0.03	0.74
Major Studio Funded	187	0.62	35	0.49	0.13	0.14
Experienced Producer	187	0.51	35	0.43	0.08	0.39
Network Producer	187	0.21	35	0.14	0.07	0.37
Award Producer	187	0.32	35	0.26	0.06	0.49
Filmed in LA	187	0.33	35	0.34	-0.01	0.90

Panel C: Common Support & IMDb Ratings Available

	Piloted		Straight-to-Series			
	N	Mean	*N*	Mean	Diff.	*P*-val
Producer Count	94	3.53	28	3.50	0.03	0.80
Sister Studio Funded	94	0.59	28	0.46	0.12	0.26
Major Studio Funded	94	0.64	28	0.46	0.17	0.10
Experienced Producer	94	0.64	28	0.50	0.14	0.19
Network Producer	94	0.36	28	0.18	0.18	0.07
Award Producer	94	0.40	28	0.29	0.12	0.26
Filmed in LA	94	0.33	28	0.39	-0.06	0.54

Pilot observations are restricted to programs from 2008 to 2013; straight-to-series observations are restricted to programs from 2014 to 2017. Common support cutoff range is 10^{th} to 90^{th} percentile.

TABLE A3 Non-Stationary Bias Adjustment

	(1) IMDb Rating	(2) IMDb Rating	(3) IMDb Rating	(4) IMDb Rating	(5) IMDb Rating	(6) IMDb Rating
Producer Count	-0.425	-0.293	-0.235	0.048	0.030	0.036
	(0.033)	(0.135)	(0.148)	(0.647)	(0.783)	(0.741)
Sister Studio Funded	-0.015	-0.038	-0.140	-0.122	-0.050	-0.055
	(0.960)	(0.900)	(0.609)	(0.435)	(0.738)	(0.719)
Major Studio Funded	0.375	0.997	1.477	-0.080	-0.119	-0.114
	(0.535)	(0.206)	(0.065)	(0.570)	(0.455)	(0.470)
Experienced Producer	-0.073	0.073	0.230	0.084	0.099	0.089
	(0.761)	(0.783)	(0.360)	(0.621)	(0.540)	(0.581)
Network Producer	0.399	0.183	-0.062	0.072	0.036	0.036
	(0.147)	(0.546)	(0.839)	(0.607)	(0.798)	(0.799)
Award Producer	0.298	0.158	0.107	-0.041	-0.070	-0.065
	(0.279)	(0.604)	(0.703)	(0.743)	(0.549)	(0.584)
Piloted	0.322	0.372	1.282			
	(0.128)	(0.070)	(0.002)			
Hour Long		0.879	1.442		0.441	0.504
		(0.032)	(0.000)		(0.000)	(0.002)
Piloted × Hour Long			-1.358			
			(0.002)			
Post 2013				-0.062	-0.040	0.032
				(0.536)	(0.673)	(0.855)
Post 2013 × Hour Long						-0.123
						(0.552)
Constant	7.425	5.770	4.848	6.211	5.986	5.945
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Programs (*N*)	122	122	122	327	327	327
Adjusted R^2	0.461	0.536	0.592	0.188	0.225	0.223
Sample	Overlap	Overlap	Overlap	Bias	Bias	Bias

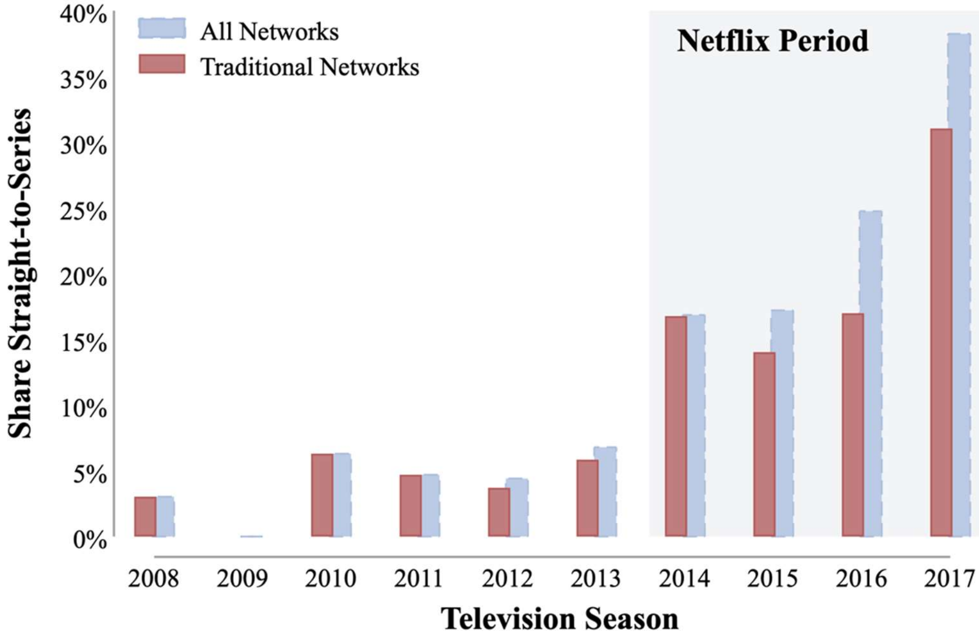
OLS estimation with *p*-values in parentheses based on error terms clustered by *Network* × *Year*.
Includes *Network* and *Location* fixed effects as well as controls for *Producer Count*, *Sister Studio Funded*, *Major Studio Funded*, *Experienced Producer*, *Network Producer*, and *Award Producer*.
Uses inverse-propensity-score-based weighting over a region of 10^{th} to 90^{th} percentile common support.

TABLE A4 Outcome Robustness Check

	(1) IMDb Log Votes	(2) IMDb Log Votes	(3) IMDb Log Votes	(4) Renewed	(5) Renewed	(6) Renewed
Producer Count	-0.183	-0.000	0.063	0.004	0.006	-0.001
	(0.637)	(0.999)	(0.839)	(0.975)	(0.958)	(0.995)
Sister Studio Funded	0.083	0.051	-0.061	0.319	0.319	0.331
	(0.879)	(0.919)	(0.899)	(0.067)	(0.068)	(0.061)
Major Studio Funded	-0.269	0.594	1.124	-0.445	-0.435	-0.492
	(0.838)	(0.686)	(0.424)	(0.367)	(0.366)	(0.319)
Experienced Producer	-0.150	0.052	0.225	0.128	0.130	0.112
	(0.784)	(0.923)	(0.675)	(0.463)	(0.451)	(0.521)
Network Producer	0.735	0.435	0.165	-0.118	-0.121	-0.092
	(0.244)	(0.496)	(0.798)	(0.514)	(0.484)	(0.608)
Award Producer	0.868	0.674	0.617	0.190	0.188	0.194
	(0.106)	(0.207)	(0.209)	(0.314)	(0.320)	(0.312)
Piloted	0.079	0.188	1.220	0.219	0.222	0.131
	(0.860)	(0.666)	(0.085)	(0.075)	(0.073)	(0.669)
Hour Long		1.220	1.841		0.014	-0.053
		(0.056)	(0.005)		(0.956)	(0.887)
Piloted × Hour Long			-1.545			0.132
			(0.077)			(0.732)
Constant	8.558	6.262	5.245	0.162	0.135	0.245
	(0.000)	(0.001)	(0.002)	(0.784)	(0.796)	(0.687)
Programs (*N*)	122	122	122	122	122	122
Adjusted R^2	0.562	0.596	0.610	0.260	0.251	0.246
Sample	Overlap	Overlap	Overlap	Overlap	Overlap	Overlap

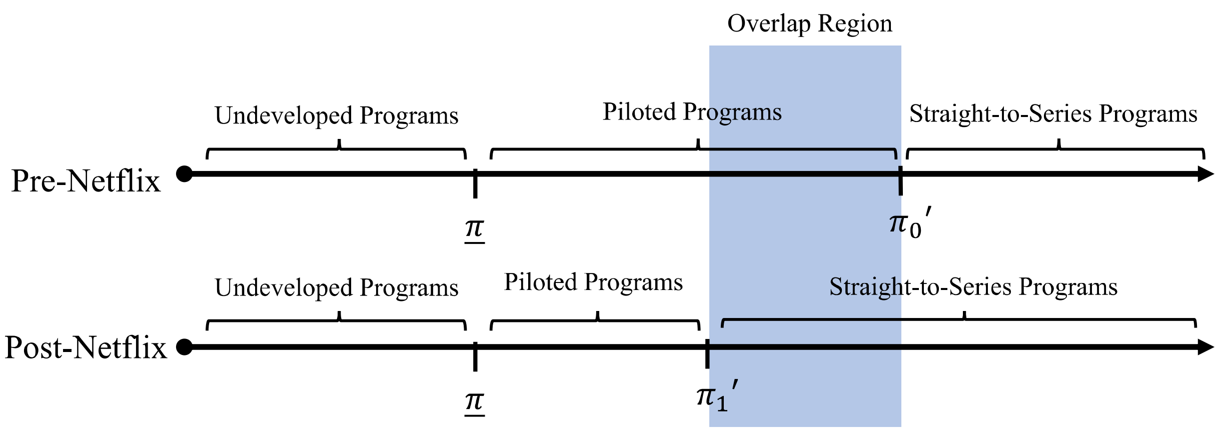
OLS estimation with *p*-values in parentheses based on error terms clustered by *Network* × *Year*.
Includes *Network* and *Location* fixed effects as well as controls for *Producer Count*, *Sister Studio Funded*, *Major Studio Funded*, *Experienced Producer*, *Network Producer*, and *Award Producer*.
Uses inverse-propensity-score-based weighting over a region of 10^{th} to 90^{th} percentile common support and applies a bias adjustment to point estimates and standard errors, see Appendix for details.

FIGURE A1 Share of Programs Ordered Straight-to-Series



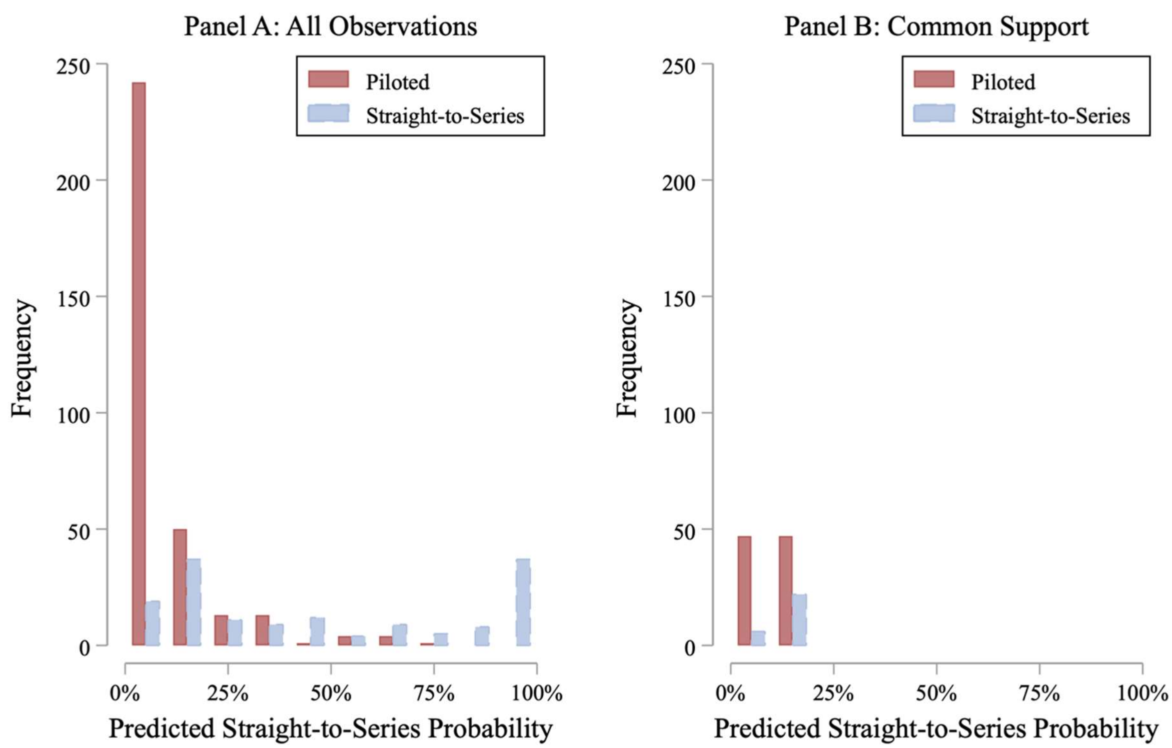
Scripted programs on networks that were either piloted or ordered directly to series. Traditional Networks excludes the streaming networks Netflix, Amazon, and Hulu.

FIGURE A2 Overlap in Program Ex-Ante Quality



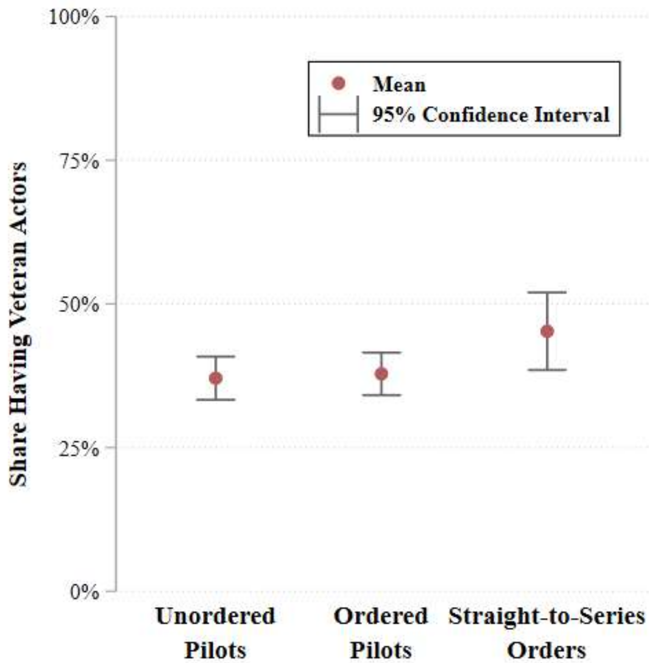
Overlap region representing theoretical shift in ex-ante quality for Straight-to-Series programs on incumbent networks 2008–2013 versus 2013–2017.

FIGURE A3 Estimator Region of Common Support



Predicted probability of a straight-to-series order based on pre-order program characteristics, with region of common support limited to programs within the 10th to 90th percentile probability of straight-to-series order.

FIGURE A4 Share of Programs with Veteran Actors



Plot of *Veteran Actor* mean and 95% confidence interval for all observations in dataset by type of order.