Online Appendices

Gender Inequality and Product Choice: Evidence from the Weinstein Scandal and #MeToo

A Appendix Tables and Figures

Table A.1: Summary statistics, unmatched sample

	Weinstein association = 0			Weinstein association = 1			
	Obs	Mean	SD	Obs	Mean	SD	(p-value)
Female protagonists	2295	0.38	0.48	1750	0.38	0.49	(0.73)
Feminine	2295	0.51	0.50	1750	0.51	0.50	(0.99)
Post-shock	2295	0.25	0.43	1750	0.27	0.44	(0.26)
Include female writers	2295	0.26	0.44	1750	0.25	0.43	(0.50)
Include female producers	2295	0.49	0.50	1750	0.54	0.50	(0.00)
Producer experience	2295	6.29	7.44	1750	11.87	9.35	(0.00)
Producer prior awards	2295	0.30	0.94	1750	1.23	2.05	(0.00)
Producer exp. w/ major studios	2295	0.49	0.43	1750	0.66	0.35	(0.00)
Producer exp. w/ top agencies	2295	0.55	0.39	1750	0.76	0.25	(0.00)
Producer team size	2295	2.79	1.70	1750	3.86	2.09	(0.00)
Writer experience	2295	0.78	1.51	1750	1.02	1.70	(0.00)
Writer team size	2295	1.27	0.50	1750	1.24	0.49	(0.10)
Top four agencies	2295	0.47	0.50	1750	0.59	0.49	(0.00)
Original	2295	0.64	0.48	1750	0.64	0.48	(0.93)
Complete script	2295	0.58	0.49	1750	0.55	0.50	(0.04)
Talent attached	2295	0.54	0.50	1750	0.59	0.49	(0.00)

Note: This table summarizes the variables by association with Weinstein using the unmatched sample. The sample is based on the unmatched sample used in Luo and Zhang (2022), dropping 143 observations that do not have the logline information that we use to generate the dependent variables on which this paper focuses. See Notes in Table 1 in the paper for the definitions of the variables.

Table A.2: Robustness: whether the project features female protagonists

Dependent variable	Female protagonists only		Female protagonists	Female protagonists only	Female protagonists		
Sample	All	Exclude both genders	Exclude cases with predicted gender		All	All	Single writer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
W Association × Post-shock × Includes women	-0.295***	-0.311***	-0.240**	-0.304***			-0.332***
	(0.073)	(0.083)	(0.100)	(0.096)			(0.092)
W Association × Post-shock	0.146*	0.146	0.107	0.171**		0.032	0.204**
	(0.080)	(0.087)	(0.084)	(0.071)		(0.076)	(0.091)
Includes women × Post-shock	0.103*	0.113*	0.045	0.074	0.007		0.142**
	(0.055)	(0.062)	(0.088)	(0.061)	(0.061)		(0.065)
Includes women × W Association	0.096***	0.111***	0.052	0.065			0.092**
	(0.027)	(0.029)	(0.032)	(0.037)			(0.039)
W Association	-0.070**	-0.070**	-0.019	-0.061***		-0.018	-0.051
	(0.033)	(0.032)	(0.026)	(0.017)		(0.027)	(0.033)
Includes women	-0.122**	-0.095*	-0.018	-0.114	-0.008		-0.071
	(0.049)	(0.049)	(0.079)	(0.067)	(0.049)		(0.065)
W Association (total) × Post-shock × Includes women					-0.039***		
					(0.013)		
W Association (total) × Post-shock					0.025*		
					(0.014)		
Includes women × W Association (total)					0.010		
					(0.007)		
W Association × Post-shock × Percentage of women						-0.361**	
						(0.169)	
Percentage of women × W Association						0.163**	
						(0.063)	
Post-shock × Percentage of women						0.103	
						(0.162)	
All controls	Y	Y	Y	Y	Y	Y	Y
Observations	1977	1845	1480	1480	1977	1976	1526
R^2	0.223	0.244	0.239	0.257	0.210	0.220	0.226

Note: OLS regressions. Column 1 uses an indicator for stories featuring female protagonists only (those featuring protagonists of both genders are coded as zero). Column 2 drops cases with protagonists of both genders. Columns 3 and 4 drop projects for which the gender of the protagonists is predicted (25% of the matched sample) rather than manually coded based on the pronouns and names associated with the protagonists. Columns 5 uses a continuous measure of association with Weinstein, W Association (total), which is the total number of past released movies on which all the producers in the team had collaborated with Weinstein before the focal project was set up. Column 6 uses a continuous measure of female representation, which is the percentage of female writers and producers in the entire team. Column 7 uses projects with single writers. All regressions use the same set of controls as in Column 1 in Table 2. Standard errors (in parentheses) are clustered at the studio level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Joint decision over writer gender and protagonist gender

	(1)	(2)
Sample	All-male producers	Include female producers
Baseline choice: All-male writers & male protagonists		
Alternative choice 1: All-male writers & female protagonists		
W Association × Post Shock	1.190**	-0.754
	(0.600)	(0.504)
W Association	-0.589**	0.140
	(0.260)	(0.211)
Alternative choice 2: Include female writers & male protagonists		
W Association × Post Shock	1.474***	1.142***
	(0.515)	(0.355)
W Association	-0.538***	-0.020
	(0.177)	(0.146)
Alternative choice 3: Include female writers & female protagonists		
W Association × Post Shock	0.331	0.401
	(0.789)	(0.404)
W Association	-0.295	0.015
	(0.247)	(0.153)
All controls	Y	Y
Observations	949	1028

Note: Multinomial logit regressions (matched sample). The dependent variable indicates four options that vary by the gender of the writers and the gender of the protagonists: the baseline option of "all-male writers and male protagonists" and three alternative options, which are "all-male writers and female protagonists," "include female writers and male protagonists," and "include female writers and female protagonists." Column 1 uses the subsample of projects set up by production teams consisting only of male producers, while Column 2 uses projects developed by production teams that have at least one female producer. The controls are the same as our baseline analysis, including quarter, studio, and genre fixed effects. Standard errors clustered at the studio level (in parentheses). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Whether the project features female protagonists, by producer experience

Dependent variable	Female protagonists			
Sample	Teams with all men		Teams with women	
	Producer experience below median	Producer experience above median	Producer experience below median	Producer experience above median
	(1)	(2)	(3)	(4)
W Association × Post-shock	0.259***	0.026	-0.175**	-0.152**
	(0.057)	(0.098)	(0.072)	(0.066)
W Association	-0.137***	-0.011	0.027	0.073***
	(0.027)	(0.044)	(0.018)	(0.023)
Other controls	Y	Y	Y	Y
Genre FE	Y	Y	Y	Y
Studio FE	Y	Y	Y	Y
Quarter FE	Y	Y	Y	Y
Observations	365	455	605	552
R^2	0.232	0.135	0.243	0.318

Note: OLS regressions. The dependent variable of all columns is whether or not the project features female protagonists, with the omitted group including projects featuring only male protagonists. The four columns split the sample in four ways: first by whether the team consists of all men or includes at least one woman; and then by whether the overall experience of the producers is above or below the sample median. All regressions use the same set of controls as Column 1 in Table 2. Standard errors (in parentheses) are clustered at the studio level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Script characteristics

(a) Whether the project is based on completed scripts

Dependent variable	Completed script			
Sample	Teams with all men		Teams with women	
	Female protagonists	Male protagonists	Female protagonists	Male protagonists
	(1)	(2)	(3)	(4)
W Association × Post-shock	-0.288**	-0.116	0.102	-0.087
	(0.113)	(0.094)	(0.065)	(0.114)
W Association	-0.143**	-0.032	0.011	-0.008
	(0.058)	(0.041)	(0.056)	(0.037)
Other controls	Y	Y	Y	Y
Genre FE	Y	Y	Y	Y
Studio FE	Y	Y	Y	Y
Quarter FE	Y	Y	Y	Y
Observations	195	625	540	617
R^2	0.358	0.123	0.231	0.215

(b) Whether the project is based on completed and original scripts

Dependent variable	Completed and original script			
Sample	Teams with all men		Teams with women	
	Female protagonists	Male protagonists	Female protagonists	Male protagonists
	(1)	(2)	(3)	(4)
W Association × Post-shock	-0.315*	-0.120**	0.181***	0.058
	(0.157)	(0.049)	(0.061)	(0.105)
W Association	-0.070	-0.014	-0.050	-0.022
	(0.056)	(0.033)	(0.050)	(0.034)
Other controls	Y	Y	Y	Y
Genre FE	Y	Y	Y	Y
Studio FE	Y	Y	Y	Y
Quarter FE	Y	Y	Y	Y
Observations	195	625	540	617
R^2	0.308	0.170	0.248	0.196

Note: OLS regressions. The dependent variable of all columns is whether or not the project features female protagonists, with the omitted group including projects featuring only male protagonists. The four columns split the sample in four ways: first by whether the team consists of all men or includes at least one woman; and then by whether the overall experience of the producers is above or below the sample median. All regressions use the same set of controls as Column 1 in Table 2. Standard errors (in parentheses) are clustered at the studio level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Release outcomes, by gender stereotypes of project content

Dependent variable	Released							
Sample	Teams with all men				Teams with women			
	Female protagonists		Male protagonists		Female protagonists		Male protagonists	
	Relatively feminine	Relatively masculine	Relatively feminine	Relatively masculine	Relatively feminine	Relatively masculine	Relatively feminine	Relatively masculine
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
W Association × Post-shock	-0.616***	-0.395	-0.093	-0.040	0.129***	-0.083	0.181*	0.062
	(0.145)	(0.267)	(0.077)	(0.076)	(0.035)	(0.090)	(0.093)	(0.120)
W Association	0.349***	0.013	-0.031	-0.028	-0.025	-0.011	-0.011	0.001
	(0.054)	(0.102)	(0.048)	(0.027)	(0.030)	(0.041)	(0.030)	(0.033)
Other controls	Y	Y	Y	Y	Y	Y	Y	Y
Observations	117	78	225	400	376	164	270	347
R^2	0.528	0.698	0.236	0.156	0.172	0.364	0.273	0.167

Note: OLS regressions. This table examines a project's outcome by the gender stereotypes of project content. The first four columns on projects by all-male teams. It splits this subsample in four ways by: (1) whether the project features female protagonists, and (2) whether the gender-role measure is above or below the sample median (i.e., relatively more or less feminine). The last four columns replicate the first four columns using projects by teams with at least one woman. All regressions use the same set of controls as Column 1 in Table 2. Standard errors (in parentheses) are clustered at the studio level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: By proportion of female producers in the production team

Dependent variable	Female protagonists		
Sample	female producers ≤ 25%	25% < female producers ≤ 50%	female producers > 50%
	(1)	(2)	(3)
W Association × Post-shock	0.038	-0.229**	-0.293**
	(0.135)	(0.094)	(0.129)
W Association	0.046	0.028	0.115*
	(0.108)	(0.029)	(0.064)
All controls	Y	Y	Y
Observations	168	621	239
R^2	0.342	0.293	0.270

Note: OLS regressions using only observations with at least one female producer. The dependent variable of all columns is whether or not the project features female protagonists, with the omitted group including projects featuring only male protagonists. Columns 1-3 use the subset for which the proportion of female producers is less than 25%, 25%-50%, and >50%, respectively. The controls are the same as our baseline analysis, including quarter, studio, and genre fixed effects. Standard errors clustered at the studio level (in parentheses). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Writers' and producers' past tendency to work on female-protagonist projects

(a) Writers

Dependent variable	Past tendency to write female-protagonist stories		
Sample	All	Teams with all men	Teams with women
	(1)	(2)	(3)
W Association × Post-shock	-0.000	-0.016	0.025
	(0.029)	(0.026)	(0.042)
W Association	0.009	-0.007	0.017
	(0.013)	(0.024)	(0.015)
Other controls	Y	Y	Y
Genre FE	Y	Y	Y
Studio FE	Y	Y	Y
Quarter FE	Y	Y	Y
Observations	1977	820	1157
R^2	0.288	0.279	0.342

(b) Producers

Dependent variable	Past tendency to produce female-protagonist stories		
Sample	All	Teams with all men	Teams with women
	(1)	(2)	(3)
W Association × Post-shock	0.007	0.024	-0.014
	(0.016)	(0.034)	(0.020)
W Association	-0.002	-0.023	0.005
	(0.011)	(0.016)	(0.009)
Other controls	Y	Y	Y
Genre FE	Y	Y	Y
Studio FE	Y	Y	Y
Quarter FE	Y	Y	Y
Observations	1977	820	1157
R^2	0.041	0.201	0.193

Note: OLS regressions. The dependent variable of Panel (a) is the tendency of the writers on a team to write female-protagonist stories before the focal project. It is the percentage of female-protagonist stories out of all the television series and films the writer obtained writing credits for before the focal project. To aggregate this measure to the project level, we take the mean of this variable among all the writers of the same project if the project includes more than one writer. Column 1 uses the entire sample, and Columns 2 and 3 split the sample by whether there are women on the team. Apart from the usual controls, we also include a dummy indicating that the dependent variable is not defined (because none of the writers on the project had any past writing credits). Panel (b) replicates Panel (a) but uses a similar measure on the past tendency to produce female-protagonist stories for the producers on the team. Standard errors (in parentheses) are clustered at the studio level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Construction of the Outcome Variables

We use the logline of a script, a short summary typically consisting of one to three sentences, that is recorded in the Done Deal Pro database to construct the three variables that we use to characterize content: 1) the gender of the protagonists; 2) the appeal of the movie to a typical man and a typical woman in the U.S.; and 3) how the protagonist is portrayed relative to traditional perceptions of gender stereotypes.

B.1 Female protagonists

B.1.1 Mechanical Turk (MTurk) survey question

We ask MTurk workers to help us classify the gender of the protagonists. In particular, we ask the following question:

> Based on the following logline:
>
> *Pierre, a quietly resourceful bartender, returns to his hometown after the death of his parents. When he falls in love with the enigmatic Stella, he is unwittingly drawn into a circle of fate pitting him against the volatile criminal, Shane.*
>
> I think the gender of the protagonist is:
>
> ○ Female ○ Male ○ Both female and male (if multiple protagonists) ○ Unclear

We provide the following instructions that help MTurkers decide:

> Please help us classify the gender of the protagonist (lead character) of a movie, based on its logline (1-2 sentence summary). The following are a few notes that might help you decide:
>
> - A movie typically has a single central protagonist, so please select the third option only if you are sure that there are multiple central protagonists.
> - If both a male and a female character show up in a logline, typically, the one that shows up first is the protagonist (e.g., "Pierre, a quietly resourceful bartender, returns to his hometown after the death of his parents. When he falls in love with the enigmatic Stella, he is unwittingly drawn into a circle of fate pitting him against the volatile criminal, Shane." In this example, Pierre is the protagonist.)
> - Sometimes, you have to read to the end of the logline before you realize the gender of the protagonist (e.g., "Aiman, a young correctional officer, befriends older colleague Koon, an executioner at the prison, and Aiman must grapple with the possibility that he may have to take over the older man's job." In this example, "he" is used to refer to Aiman, the lead character, and this gender pronoun shows up at the end.)
> - Apart from gender pronouns, you can sometimes make inferences from the name of the lead character (e.g., if the protagonist's name is Emily, the protagonist is clearly female). Please indicate "unclear" if you are not sure about the gender of the name of the protagonist.
> - Please do not classify the gender purely based on the person's profession, unless you are absolutely sure (for example, a U.S. sports figure could be a male or a female. If this is the only information available, please indicate "unclear.")

- Please click “unclear” if there is not enough information to tell whether the protagonist is a man or a woman or if the logline does not refer to a man or a woman.

The payment per assignment, which asks the above question about a single logline, is $0.04. We restrict the respondents to workers who have approval rates greater than 95%, have the number of Human Intelligence Tasks (HITs) approved to be greater than 500, and are located in the United States.

B.1.2 Construction of the variable *female protagonists*

We construct the variable, *female protagonists*, via a two-step process.

In Step 1, we construct a variable to capture the gender of the protagonist based on MTurk workers' answers to the above question through an iterative process. We ask two different MTurk workers to classify the gender of the protagonist for each of the 4,045 loglines in the unmatched sample. If their answers coincide, we code the gender of the protagonist accordingly. For loglines for which the two answers differ (1011 in total, and 25% of the entire sample), we ask a third MTurk worker the same question independently to break the tie. For these loglines, if two out of the three answers are consistent with each other, we code the gender of the protagonist accordingly. For a small number of loglines for which all three answers differ (242 in total, and six% of the entire sample), we ask a fourth MTurk worker independently. For these loglines, if two out of the four answers are the same, we code the gender of the protagonist accordingly. For loglines for which all four answers differ (28 observations in total), we (the authors) code up the final answer.

The variable *protagonist gender* generated via the above process categorizes 26.58% of the loglines as featuring female protagonists; 43.26% as featuring male protagonists; 6.85% as featuring both female and male protagonists; and 23.31% as unclear.

For a validity check, the left panel of Table B.1 summarizes the percentage of loglines in each category that are written by female writers. It is intuitive that the percentage of female writers is substantially higher for loglines with female protagonists compared to loglines that are categorized as featuring male protagonists (54% vs. 13%, p-value is = 0.0000). Seventeen percent of the loglines that are categorized as unclear are written by female writers, which suggests that the majority of these “unclear” cases are likely to feature male protagonists, even though the short summary does not have clear gender pronouns or protagonist names that are clearly female or male. This is consistent with our manual examination of a randomly selected subset of “unclear” cases: for the majority of these loglines, the contextual information strongly suggests that the protagonist is male.

In Step 2, to better utilize observations in the ‘unclear’ category, which comprise 23.31% of the sample, we employ machine-learning techniques to predict whether a given logline is likely to feature female or male protagonists based on the contextual information. In particular, we use the subset of loglines that are clearly identified as featuring female and male protagonists as the input into a supervised machine-learning model.[45] A random 80% of the clearly labeled loglines are allocated to the training set and the other 20% to the testing set.

To implement the procedure, we first encode the loglines into vectors. Recent advancements in natural language processing allow us to encode sentences or words in more meaningful forms than simply counting the

[45] We chose not to include cases with both male and female protagonists, mainly because they made up of a very small percentage of the total sample, and it is possible that the coders were less stringent about finding a single central character in these cases.

Table B.1: Percentage of female writers, by the gender of the protagonist

Step 1			Step 2		
Protagonist gender	N	Percentage of female writers	Protagonist gender	N	Percentage of female writers
Female	1075	53.8%			
Male	1750	13.0%			
Both	277	27.0%			
Unclear	943	17.3%	Female (predicted)	180	33.3%
			Male (predicted)	763	13.6%
Total	4045	26.0%			

Note: In Step 1, MTurk workers classify the gender of the protagonist based on the pronouns and names associated with the protagonist. For 23.31% of the loglines for which the gender of the protagonist is not immediately clear, we use supervised machine-learning method, based on observations that are clearly coded as featuring male or female protagonists in Step 1, to predict the gender of the protagonist.

frequency of the actual words used in the texts, taking into consideration the semantic and syntactic information of words.[46] The word2vec techniques and Bidirectional Encoder Representations from Transformers (BERT) language models are two important ones.[47] We use BERT models, created and published in 2018 by Jacob Devlin and his colleagues at Google (Devlin et al., 2018), in our paper. A key advantage of BERT models relative to word2vec models is that the former takes into consideration the context of each word. For example, bank as a financial institute is different from bank of the river. BERT models will yield different vectors for the word 'bank' in these two different contexts, whereas word2vec will produce the same vector. In addition, BERT models will transform an entire paragraph into a vector in a single step, whereas with word2vec, one needs to transform each word into a vector first and then aggregate all the word vectors included in a logline to a single vector.

We use a pre-trained BERT model to convert each logline into a 768-dimension vector and then run a logistic regression to train the supervised learning classification. The label (female and male protagonist) is the dichotomous dependent variable, and the logline vectors are the predictors. The trained model achieves an accuracy rate of 98.04% for the training set and 94.86% for the testing set.

We then use the trained model to predict the gender of the protagonists for loglines that are categorized as 'unclear' in Step 1. Among the 943 loglines, 180 (19%) are predicted to feature female protagonists, and the other 763 (81%) are predicted to feature male protagonists. Such a split is consistent with the above statistic and our manual inspection that a vast majority of the unclear loglines likely feature male protagonists. To provide a reality check of this prediction, the right panel of Table B.1 tabulates the share of female writers for each predicted category: this percentage is 13.6% for loglines that are predicted to feature male protagonists, which is virtually the same as the percentage of female writers for loglines that are clearly identified as featuring male protagonists in Step 1. The percentage of female writers is 33.3% for loglines that are predicted to feature female protagonists, which is closer to the same percentage for loglines categorized as 'both' in Step 1 than to that for loglines that are clearly identified as featuring female protagonists in Step 1. The difference in the percentage of female writers for the two predicted categories among all the "unclear" loglines (33.3% vs.

[46] These vectors are known to capture analogical reasoning—for example, king − men + women = queen—or capture meaningful relationships with other vectors (e.g., the vector of dog will be quite similar to the vector of cat or domestic).

[47] The original models are trained on the BooksCorpus with 800M words and a version of the English Wikipedia with 2,500M words. BERT has achieved state-of-the-art performance on a number of NPL tasks.

13.6%) is statistically significant (the p-value is 0.0000).

Finally, the main outcome variable that we use in our paper, *female protagonists*, equals one if the protagonist's gender is classified as 'female' or 'both' by MTurk workers in Step 1, or if the protagonist's gender is 'unclear' in Step 1 and is predicted to feature female protagonists in Step 2. In the paper, we show the robustness of our core results to an alternative definition of 'female protagonist only' and to excluding all the "unclear" cases from the analysis.

Conceptually, what we mean by the gender orientation of an idea is that it is more likely to appeal to an audience of a given gender. At the intuitive level, by placing a woman at the center of the story universe, female-protagonist stories are likely appeal to a female audience. To verify this concept, we also construct a variable that is intended to capture the movie's likely appeal to an audience of a given gender more directly. We explain the construction of this variable in detail in the next section and show that the gender of the protagonists is highly correlated with this demand-side measure. Finally, we conduct an analysis using *Rotten Tomatoes* critic and user review data, which are presented in Table B.2. The results show that female reviewers (both critics and users) are more likely than male reviewers to rate movies with a majority of female leads more favorably compared to movies with a majority of male leads.

Table B.2: Critic and user reviews, by the gender of the protagonist

DV	Critic rating	User rating
	(1)	(2)
Female critic × Majority female leads	0.016**	
	(0.007)	
Female users × Majority female leads		0.123**
		(0.062)
Movie fixed effects	Y	Y
Critic (user) fixed effects	Y	Y
Observations	109498	464200
R^2	0.366	0.315

Note: OLS regressions using critic-movie- and user-movie-level data of 761 movies released in theaters between January 2014 and the third quarter of 2019. We construct this dataset by starting with the 2,872 movies that were released in U.S. theaters between January 2014 and September 2019 listed by *the-numbers.com*. We are left with 1,450 movies after dropping people listed as cast and crew for whom the confidence level of the prediction of *genderize.io* is below 90%; and keeping movies with non-missing information on the leading cast, the writers, the director, and the producers (the dropped movies are mostly indie productions or foreign movies). We further drop another 257 movies because we cannot find a match on *www.rottentomatoes.com*, 375 movies because the earliest professional reviews listed on *www.rottentomatoes.com* were published more than 180 days before the U.S. theatrical dates (these movies are mostly re-releases of foreign movies), and 57 movies because there is no review information on *www.rottentomatoes.com*. The final sample includes 761 movies.

The dependent variable of Column 1 is a dummy variable indicating 'fresh' (positive) versus 'rotten' (negative), which is a binary rating scheme that *Rotten Tomatoes* uses to translate the reviews by professional critics. The dependent variable of Column 2 is the user rating that ranges from 0 to 5. The regressions include movie fixed effects and critic (or user) fixed effects. Standard errors (in parentheses) are clustered at the movie level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Gender appeal

As mentioned above, we construct a variable intended to capture the movie's likely appeal to an audience of a given gender. While this measure captures the gender orientation of an idea directly from the demand perspective, we rely on the gender of the protagonist as our primary measure in the paper because it is more precisely measured. Consumer appeal of a movie is notoriously difficult to predict for early-stage movie projects, let alone with the short descriptions that we have. Nonetheless, we show below that the gender of the protagonists is highly correlated with the demand-side measure.

B.2.1 MTurk survey design

MTurk workers are asked to classify the appeal of a logline separately for men and women at a 1-5 scale. Below is an example of a task:

> Below is a logline, which is a short summary of a movie. Using your best judgement, please help us rate the appeal of the movie for men and women:
>
> *An American civilian turned self-taught spy works with the FBI to bring down a Russian intelligence agent on American soil.*
>
> 1. Based on this logline, how much do you think a typical ***woman*** in the U.S. will like this movie?
>
> ◯ Really Dislike ◯ Dislike ◯ Neither Like or Dislike ◯ Like ◯ Really Like
>
> 2. Based on this logline, how much do you think a typical ***man*** in the U.S. will like this movie?
>
> ◯ Really Dislike ◯ Dislike ◯ Neither Like or Dislike ◯ Like ◯ Really Like
>
> 3. How do you identify your gender?
>
> ◯ Woman ◯ Man ◯ Transgender ◯ Non-conforming ◯ Prefer not to respond

We ask the respondents to rate a movie's appeal based on their beliefs of how much a typical man or woman may like the movie, rather than on their own preferences because we want to capture the social perception of a movie's gender appeal. This is also likely to correspond with what the producers may perceive as a movie's appeal to consumers of a given gender. This method also does not require us to specify the respondents' gender ex-ante, which would have significantly raised the cost of soliciting respondents on MTurk. Wühr et al. (2017) conduct two separate studies that compare men's and women's own preferences with their evaluations of the preference of a typical man or a typical woman for 17 movie genres. The results show that for the majority of the genres, the perception of others' preferences is consistent in direction with one's own preferences (e.g., both surveys show that women prefer romance more than men), even though it tends to overestimate the actual gender differences.

The payment per assignment, which asks the above questions about a single logline, is \$0.07. We restrict the workers to those who have approval rates greater than 95%, have the number of Human Intelligence Tasks (HITs) approved to be greater than 500, and are located in the United States to minimize differences in perceptions of appeal across cultures.

Each logline is classified 15 times by unique MTurk workers. After dropping classifications with very low work times (i.e., completion time less than ten seconds), we are left with a total of 50,163 usable classifications for 4,045 loglines. About 37% of the classifications were completed by women and 61% by men. Compared to male MTurk workers, female MTurk workers tend to rate a logline as having slightly higher female appeal (by 0.03) and lower male appeal (by 0.07). We do not think this systematic difference between female and male MTurkers is concerning. Because we randomize the order of the loglines, there are no reasons to expect that the share of female MTurkers rating a given logline will differ systematically by the key variables we use in the paper, such as the gender of the protagonists, the writers, or the producers. Moreover, the share of female MTurkers is not significantly correlated with the gender appeal measure we construct based on these answers. On average, loglines are rated as having higher appeal to men than to women (by 0.06), which is not surprising given that a greater proportion of the loglines are developed primarily for a male audience (e.g., having a male protagonist).

B.2.2 Construction of the variable *gender appeal*

For each logline, we conduct paired t-tests comparing the answers to the above two gender-appeal questions by the same MTurker.[48] We define a *gender appeal* variable that equals -1 if a logline is rated significantly more appealing to men than to women at the 5% level based on one-sided p-values; 1 if a logline is rated significantly more appealing to women than to men at the 5% level based on one-sided p-values; 0 for the remaining loglines.

Among the 4,045 projects, 22% are rated as significantly more appealing to women than to men; 26% are rated as significantly more appealing to men than to women; and the remaining 51% are likely to attract women and men to a (statistically) similar degree. Table B.3 tabulates the percentage of writers within each category of gender appeal that are female. Also consistent with what we expect, this percentage is increasing monotonically as we move from projects that are rated as more appealing to men to more appealing to women.

Table B.3: Percentage of female writers, by the appeal to a given gender

Gender appeal	N	Percentage of female writers
More appealing to men	1,079	10%
Similarly appealing to both genders	2,073	22%
More appealing to women	893	51%
Total	4,045	26%

The correlation between *gender appeal* and *female protagonists* is 0.47 (p-value = 0.000). Table B.4 tabulates the two measures of gender orientation. 48.13% of movies featuring female protagonists are rated as significantly more appealing to women, whereas this percentage is only 8.36 for movies featuring male protagonists). Similarly, 38.96% of movies featuring male protagonists are rated as significantly more appealing to men, whereas this percentage is only 4.06 for movies featuring female protagonists).

B.3 Gender-role measure: feminine-masculine characteristics

B.3.1 The construction of the measure

We aim to construct a measure that captures how a protagonist is portrayed relative to the traditional perception of gender stereotypes. To achieve this goal, we use the set of feminine and masculine characteristics in the Bem Sex-Role Inventory (BSRI) as the benchmark (Bem, 1974). We then create a measure that calculates a focal logline's distance to the feminine keywords relative to the masculine keywords in the vector space.

The Bem Sex-Role Inventory (BSRI) inventory is considered the gold standard of gender-role evaluation

[48] Recall that our survey asks the same MTurk worker to rate the appeal of a given logline to both men and women. This allows us to conduct paired t-tests that help remove the baseline differences among the respondents.

Table B.4: Tabulation by gender appeal and by the gender of the protagonist

		Gender appeal			
		More appealing to men	Similarly appealing to both genders	More appealing to women	Total
Female protagonists	N	51	600	604	1,255
	Percent	4.06%	47.81%	48.13%	100%
Male protagonists	N	979	1,324	210	2,513
	Percent	38.96%	52.69%	8.36%	100%
Both female and male protagonists	N	49	149	79	277
	Percent	17.69%	53.79%	28.52%	100%

and has been used in thousands of studies in the more than 40 years since it was developed (Dean and Tate, 2017). Table B.5 lists the full set of feminine and masculine characteristics in Table 1 of Bem (1974).[49]

As Bem (1974) notes, these characteristics are selected as masculine or feminine on the basis of sex-typed social desirability (that is, a characteristic qualified as masculine if it is judged in American society to be more desirable in a man than in a woman, and as feminine if it is judged to be more desirable in a woman than in a man). In general, masculinity has been associated with instrumental traits, a cognitive focus on "getting the job done;" and femininity has been associated with expressive traits, an affective concern for the welfare of others. As Dean and Tate (2017) discuss, later work also characterizes masculine traits as agentic and feminine traits as communal, applying these concepts to a variety of contexts, such as the effectiveness and acceptability of styles of male and female leaders and how people's perception of self in relation to agency and communion attributes influences different social outcomes, such as attraction to different academic and professional fields.

It is useful to note a few caveats when interpreting the measure. First, as we explain below, by calculating the relative distance between a focal logline and these feminine and masculine characteristics, we are technically capturing not a specific character's personality traits but the overall description of the storyline. Nonetheless, as we discuss in the next section, it seems reasonable to interpret this measure as the depiction of the protagonist. This is likely because a logline centers mainly on the protagonist, and because the natural language processing tool that we use does a reasonably good job in relating the general social and professional contexts and activities that are often described by verbs and nouns to the benchmark characteristics that are mostly adjectives. Second, the BSRI inventory has been criticized in recent years as being outdated in depicting societal gender norms; for example, Donnelly and Twenge (2017) reviewed a large collection of studies that apply BSRI and show that women's femininity scores have decreased significantly over the years. This is not very concerning for our purpose both because we want to capture a more traditional depiction of stereotypical women and because what

[49]Bem (1974) arrived at the list of 40 feminine and masculine characteristics as follows. She started off with a set of personality characteristics that she and a group of students deemed to be positive in value and either masculine or feminine in tone (400 in total, with 200 for each gender). 100 judges were asked to rate each of the 400 characteristics according to questions such as: "In American society, how desirable is it for a man to be truthful?" or "In American society, how desirable is it for a woman to be sincere?" The judges were asked to answer each question based on a 7-point scale, ranging from 1 ("Not at all desirable") to 7 ("Extremely desirable"). A personality characteristic qualified as masculine if it was judged to be significantly more desirable for a man than for a woman ($p < 0.05$). Similarly, a personality characteristic qualified as feminine if it was judged to be significantly more desirable for a woman than for a man. Of the characteristics that satisfied these criteria, 20 were selected for the masculinity scale and 20 were selected as the femininity scale.

we care about is the direction of change (both before and after the shock and by Weinstein association).

Table B.5: BSRI masculine and feminine characteristics from Table 1 of Bem (1974)

Masculine items	Feminine items
Acts as a leader	Affectionate
Aggressive	Cheerful
Ambitious	Childlike
Analytical	Compassionate
Assertive	Does not use harsh language
Athletic	Eager to soothe hurt feelings
Competitive	Feminine
Defends own beliefs	Flatterable
Dominant	Gentle
Forceful	Gullible
Has leadership abilities	Loves children
Independent	Loyal
Individualistic	Sensitive to the needs of others
Makes decisions easily	Shy
Masculine	Soft spoken
Self-reliant	Sympathetic
Self-sufficient	Tender
Strong personality	Understanding
Willing to take a stand	Warm
Willing to take risks	Yielding

We construct this gender-role measure as follows:

- In Step 1, we use a BERT pre-trained model, as explained in the previous section, to convert each of the characteristics in the BSRI masculinity and femininity scales to a 738-dimension vector. Then, we create v_f, a single vector that represents feminine characteristics, by taking the average of all the vectors associated with the BSRI feminine characteristics. Similarly, we create v_m, which represents masculine characteristics, by taking the average of all the vectors associated with the BSRI masculine characteristics. We normalize both v_f and v_m to have unit length.

- In Step 2, we use the same BERT pre-trained model to convert each logline to a vector, v_l. Note that to avoid mechanical relationships between any gendered words in a logline and the benchmark characteristics, we replace a gendered word in a logline with a corresponding gender-inclusive word. For example, "she/he" is replaced with "they;" "woman/man" is replaced with "person;" and "wife/husband" is replaced with "spouse."[50]

- In Step 3, we want to construct a measure that ranks different loglines based on their relative distances to the benchmark feminine and the masculine vectors. One such measure can be calculated as follows (Cao et al., 2020):

$$G^l = \frac{Cos(v_l - v_m, v_f - v_m) \cdot |v_l - v_m|}{|v_f - v_m|} - \frac{1}{2} = \frac{\langle v_f - v_m, v_l - v_m \rangle}{|v_f - v_m|^2} - \frac{1}{2}. \tag{2}$$

[50]We use the following webpage of Springfield college for a list of gendered pronouns and nouns, available at https://springfield.edu/gender-pronouns#: :text=Pronouns%20can%20be%20in%20the,or%20she%2Fher%2Fhers.

Geometrically, this measure is equal to the ratio between the length of the projection of the difference vector $(v_l - v_m)$ onto the difference vector $(v_f - v_m)$ and the length of the difference vector $(v_f - v_m)$, minus 0.5. The higher this measure, the closer the logline is to the benchmark feminine vector relative to the masculine vector. If a logline vector is equally distant to the two benchmark vectors, for example, the projection of $(v_l - v_m)$ onto $(v_f - v_m)$ would land in the middle between v_f and v_m. After deducting 0.5, the above measure would be zero. For loglines that are closer to the benchmark feminine vector than to the masculine vector, this measure would be positive, whereas for loglines closer to the masculine vector than to the feminine vector, this measure would be negative.

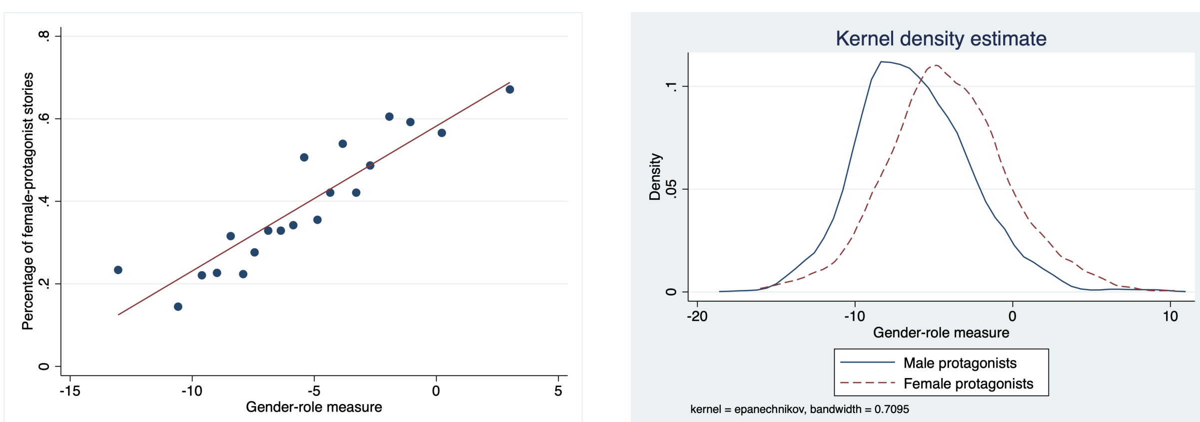
B.3.2 Description of the gender-role measure

Figure B.1a plots the percentage of female-protagonist stories for each bin of this measure. Consistent with what we might expect, the relationship is monotone positive. Figure B.1b shows that the density of this measure for female-protagonist stories is to the right of that for male-protagonist stories, though it exhibits a high variance for both sets of stories. The mean value of this measure is -4.11 for female-protagonist loglines and -6.21 for male-protagonist loglines (p-value is 0.000). That the mean of this measure is negative even for female-protagonist stories is consistent with the idea that the characteristics included in BSRI are likely to be too outdated to depict the current gender norms associated with women Donnelly and Twenge (2017).[51]

Figure B.1: Description of the gender-role measure

(a) Female-protagonist stories and the gender-role measure

(b) Kernel densities by female- and male-protagonist stories



Note: Based on the 4,045 observations in the unmatched sample.

To give some sense of how this measure corresponds to the content of the logline, we provide several examples of loglines with male and female protagonists at different parts of the distribution in Table B.6. Note that these displayed loglines are prior to the removal of gendered pronouns and nouns (as described in Step 2 above). For context, films such as Captain Marvel (*"Captain Marvel aka Carol Danvers is an air force pilot whose DNA is fused with that of an alien after an accident giving her superhuman strength and even the*

[51] Note, also, that because the measure is based on the entire logline, rather than on a specific character's personality description, it is also hard to interpret the value of this measure in relation to zero. Again, this is all right in our study because, as discussed, we care about the direction of change before and after the shock.

ability to fly."), which features a female protagonist in a traditionally male plot line scored at around the 15th percentile; however, the film Girl on the Train (*"A woman who is devastated by her recent divorce spends her daily commute fantasizing about the seemingly perfect couple who live in a house that her train passes every day until she sees something shocking happen there one morning and becomes entangled in the mystery that unfolds."*) scored at around the 85th percentile.

Table B.6: Examples of loglines across the distribution of the gender-role measure

Percentile	Male-protagonist loglines	Female-protagonist loglines
5th	*"A newly released prison gangster is forced by the leaders of his gang to orchestrate a major crime with a brutal rival gang on the streets of Southern California."*	*"Seeker. Only when it's too late does she discover she will be using her new found knowledge and training to become an assassin. The events take her around the globe from remote estate in Scotland to a bustling futuristic Hong Kong."*
25th	*"A young boy travels the world with his scientist father, adopted brother from India, Bandit the bulldog, and a government agent assigned to protect them as they go on their adventures investigating scientific mysteries."*	*"A woman returns from combat and befriends a family in New York City. When a gang of thieves plot to take the family's valuables she fights to defend the family."*
50th	*"A 30-year-old guy meets the woman of his dreams and life could not be better, but right before his wedding he gets a knock at his door and a 66-year-old man walks in and says do not marry this girl because she will ruin your life."*	*"A Wall Street financial adviser who has recently lost her job adopts a dog which has been adopted twice and returned twice for being too unruly. The woman cannot find another job and starts training Roo! who wins a special award for the best mixed-breed dog at the Westminster Dog show's first-ever agility competition, marking the first time mixed-breed dogs have appeared at the show."*
75th	*"Rickie, a sensitive and intelligent young man with an intense imagination, sets out full of hope to become a writer. Giving up his aspirations and opting for convention and marriage to Agnes, he gradually finds himself sinking into conformity and bitter disappointment until he once again realizes his dreams of literary ambition."*	*"A former pageant queen embarks on an all-night adventure with four unlikely friends she meets while volunteering at a women's shelter."*
95th	*"After 20 years of marriage, a lawyer goes to great lengths to prove his love to his wife, a music lover, and save their relationship by taking piano lessons from a free-spirited female teacher to learn Robert Schumann's Traumerei, a tricky piece that is also his wife's favorite song."*	*"A woman loves her daughter, but after years of expulsions and strained home schooling, her precarious health and sanity are weakening day by day. The battle of wills between mother and daughter ultimately reveal the frailty and falsehood of familial bonds."*

Note: The distribution cutoff is uniform rather than specific to the gender of the protagonist. As explained in the previous section, we replaced gendered pronouns and nouns with their gender-inclusive counterparts before constructing the gender-role measure. We display the original loglines here for clarity.