The Effects of Screen Quotas on the Movie Exhibition Market: Evidence From Brazil

Pedro Braga Soares

Advisor: Leonardo Rezende

September 17, 2021

Department of Economics, PUC-Rio

Outline

Introduction

Context and Motivation

- What are screen quotas?
- Where? Argentina, Spain, Mexico, South Korea, plus Brazil
- Few quantitative analyses, in Brazil or otherwise
- Comprehensive **session-level administrative data** from 2017-2019
- **Reduced-form identification strategy**: (exogenous) variation in SQ per viewing room, plus fact that 2019 had no quota in effect
- Question: **what is the effect of quotas on movie theater revenues, ticket sales, and other variables of interest**?

Main Results

- **Reduced-form** results:
  - → Negative effects of quotas on **overall** and **foreign** movie theater revenue and ticket sales
  - → Insignificant effects on **Brazilian movie** revenues and ticket sales
- Then build a **dynamic discrete-choice model** to emulate programming choices
- **Structural parameters estimates** tell a similar story, but:
  - → Model does a poor job predicting screening decisions (heterogeneity driven by private shocks)
  - → Suggest SQ play small role in exhibitor choice

Related literature

- Effects of **screen quotas** (in Brazil, Courtney, 2015 and Zubelli, 2017, Messerlin and Parc, 2014 looks at South Korea), and for cultural content quotas in general;
- IO literature of **dynamic discrete-choice models** (Rust, 1987), in particular using forward-simulation algorithms (Hotz et al., 1994, and Bajari et al. 2007); this work addresses some issues regarding first-stage Conditional Choice Probability (CCP) estimators with large state spaces.

Institutional Setting

Audiovisual Policy in Brazil

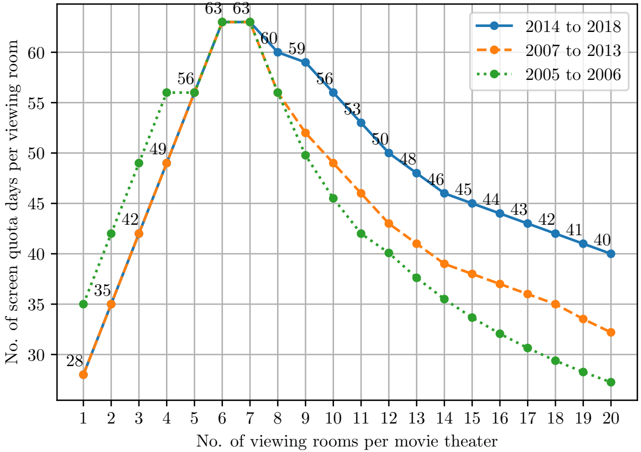
- *Medida Provisória 2228-1/2001*: **landmark** in current policy regime;
- 33 laws since 1991, plus 154 *Instruções Normativas* by Ancine;
- **Regulation**: cable TV (but not broadcasting), video, cinema;
- **Subsidies**: movies, video-games, infrastructure and equipment.
- CONDECINE tax and dedicated endowment (FSA);
- Federal funding in 2019 alone: R$ 243 million (tax breaks) plus R$ 500 million from FSA $\approx$ R$ 750 million;
- Alphabet soup **agencies**: ANCINE, SAv, CSC, CTAv, Cinemateca, Spcine and Riofilme.

Screen quotas in Brazil

- Screen quotas 1932, many incarnations: educational; week/weekend; *Lei da Dobra*;
- Since 2001, quotas have **two major requirements**:
  1. Minimum number of **days** as a (non-linear) function of the number of viewing rooms per multiplex (see Figure 1);
  2. Minimum number **titles** per year, also non-linear function multiplex viewing rooms;
- We **ignore title requirements**, since they are non-binding relative to day requirements (see Figure 2)

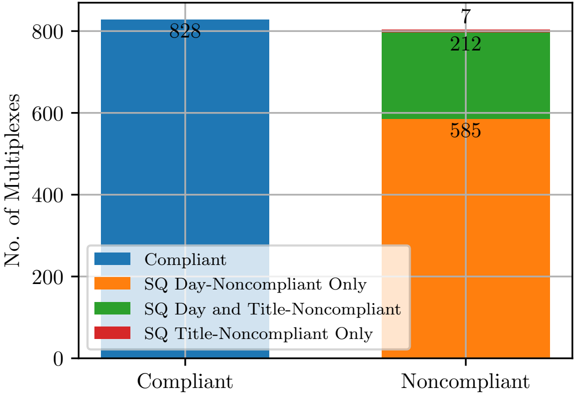
Screen quotas in Brazil

Figure 1: Screen quotas per viewing room by movie theater size



Screen quotas in Brazil

Figure 2: Screen quotas per viewing room by movie theater size (pooled sample 2017 and 2018)

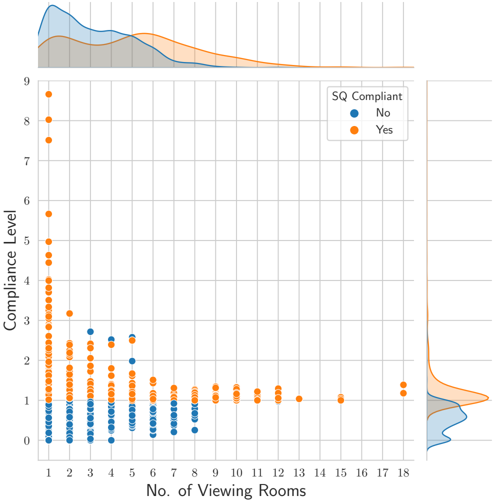


Screen quotas in Brazil

- Additional sources of **heterogeneity** of quotas:
  - → Intra-chain **"swaps"** — so we look at chain-level;
  - → Quotas are a **function of opening days**;
  - → **"Predatory occupation"**.
- **Exogeneity of quotas** per VR:
  - → Regulatory assessment report "On screen quota distortions" (see Figure 3), quotas penalize medium-sized theaters;
  - → Linear quotas set to begin in 2020 postponed.

Screen quota compliance

Figure 3: Multiplex Size vs. Screen Quota Compliance (pooled sample 2017 and 2018)



Data

Data

- Three main data sources:
  - → **Ticket sales session-level** data from exhibitors, from 2017 to 2019, obtained through Brazilian FOIA request;
  - → **Inspection reports** regarding SQ fulfillment, publicly available from 2009 to 2018;
  - → **Registry** data comprising companies, movie theaters, chains and movies.
- All data comes from **Ancine**. Data is submitted by regulated agents and submission is mandatory.

Table 1: Session Dataset Descriptive Statistics — 2017

Statistic	N	Mean	Min	Median	Max	St. Dev.
Year 2017						
Ticket Sales	4,151,236	42	0	25	850	48.195
Box-Office (in R$)	4,151,236	709.65	0	377.99	50,833.58	916.07
Seat Capacity	4,151,236	206	20	191	2,000	89.530
Seat Occupation	4,151,236	0.21	0	0.13	1	0.220
Nationality	4,151,236	0.159	0	0	1	0.366

Note: *Nationality* is a dummy variable coded 1 for Brazilian films and 0 otherwise.

Data

Table 2: Session Dataset Descriptive Statistics — 2018

Statistic	N	Mean	Min	Median	Max	St. Dev.
Year 2018						
Ticket Sales	4,306,632	37	0	20	1,242	46.807
Box-Office (in R$)	4,306,632	599.002	−9.11	295.220	967,036.000	959.85
Seat Capacity	4,306,632	205	30	191	2,000	89.827
Seat Occupation	4,306,632	0.19	0	0.11	1	0.215
Nationality	4,306,632	0.160	0	0	1	0.367

Note: *Nationality* is a dummy variable coded 1 for Brazilian films and 0 otherwise.

Data

Table 3: Session Dataset Descriptive Statistics — 2019

Statistic	N	Mean	Min	Median	Max	St. Dev.
Year 2019						
Ticket Sales	4,362,749	39	0	22	850	48.347
Box-Office (in R$)	4,362,749	645.24	0	326.57	166,163.10	878.28
Seat Capacity	4,362,749	204	29	188	2,000	92.750
Seat Occupation	4,362,749	0.20	0	0.12	1	0.225
Nationality	4,362,749	0.152	0	0	1	0.359

Note: *Nationality* is a dummy variable coded 1 for Brazilian films and 0 otherwise.

Reduced-Form Regressions

Conceptual Framework

- **Naïve** approach: SQ per VR as independent variables Go
  - → Problem: quotas may be non-biding for some (some fulfill 10x due quotas, while others ignore it);
- Alternative approach: **segment regressions** by compliance levels
  - → Problem 1: compliance bins are **arbitrary**
  - → Problem 2: "bin-shifting" so we **cannot compute FE** (only 13-40% obs in same bin, depending on spec)
  - → Problem 3: **2019** (everybody with 100% compliance?)

Conceptual Framework

- Preferred approach: weight compliance levels using kernel functions, and then interact it SQ per VR.
- Intuition: **policy effects should be stronger on narrowly compliant chains**
- We run the following regression:

$$\ln(y_{it}) = \beta_0 + \beta_1 q_{it} + \beta_2 f(c_{it}) + \beta_3 q_{it} * f(c_{it}) + \theta \mathbf{x}_{it} + \varepsilon_{it} \tag{1}$$

- y_{it}: dependent variable (total box-office or ticket sales);
- c_{it}: normalized compliance (days fulfilled / days due);
- $f(.)$: weighting function (see Figure 4);
- $\mathbf{x}_{it}$: vector of controls with movie-chain FE, year FE and opening days.

Weighting Kernels

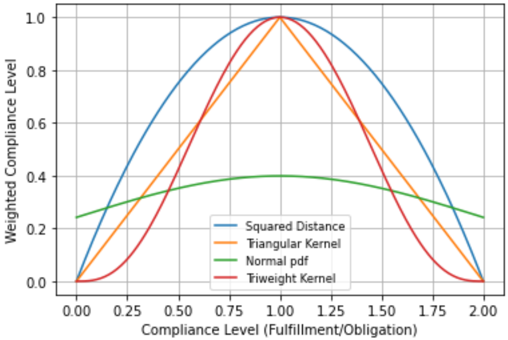


Figure 4: Weighting functions for compliance levels

Naïve Regression Results

	Dependent variable:					
	log(Box Office)			log(Ticket Sales)		
	All Movies (1)	Foreign (2)	Brazilian (3)	All Movies (4)	Foreign (5)	Brazilian (6)
Nominal screen quota per VR	0.0064	0.0078	0.0065	0.0074	0.0084*	0.0088
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)
Opening Days	0.0212	0.0207	0.0074	0.0255	0.0242	0.0139
	(0.032)	(0.032)	(0.042)	(0.031)	(0.031)	(0.040)
Chain Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
2019?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	628	624	602	628	624	602
R^2	0.960	0.961	0.930	0.958	0.959	0.931
Adjusted R^2	0.935	0.937	0.886	0.932	0.934	0.888

*p<0.1; **p<0.05; ***p<0.01

Weighted Regression Results

	Dependent variable:								
	log(Box Office)			log(Ticket Sales)			log(No. of Sessions)		
	All Movies (1)	Foreign (2)	Brazilian (3)	All Movies (4)	Foreign (5)	Brazilian (6)	All Movies (7)	Foreign (8)	Brazilian (9)
Compliance (squared dist kernel)	2.0501***	2.0906***	0.5047	1.9880***	2.0739***	0.4061	-0.8332	-0.6767	-1.3266**
	(0.2911)	(0.2991)	(0.5740)	(0.2783)	(0.2869)	(0.5471)	(0.5993)	(0.6112)	(0.5816)
Opening days	0.0348	0.0343	0.0187	0.0416	0.0402	0.0271			
	(0.0293)	(0.0294)	(0.0411)	(0.0280)	(0.0282)	(0.0392)			
Quota per Viewing Room	0.0553***	0.0609***	-0.0091	0.0555***	0.0615***	-0.0082	-0.0436**	-0.0382**	-0.0484***
	(0.0079)	(0.0081)	(0.0161)	(0.0075)	(0.0078)	(0.0154)	(0.0173)	(0.0176)	(0.0166)
Compliance × Quota per VR	-0.0541***	-0.0585***	0.0145	-0.0519***	-0.0572***	0.0167	0.0482**	0.0425**	0.0613***
	(0.0093)	(0.0096)	(0.0183)	(0.0089)	(0.0092)	(0.0174)	(0.0186)	(0.0190)	(0.0179)
Chain Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2019?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	628	624	602	628	624	602	628	624	602
R^2	0.9671	0.9683	0.9351	0.9665	0.9676	0.9367	0.9996	0.9993	0.9991
Adjusted R^2	0.9461	0.9480	0.8931	0.9450	0.9467	0.8958	0.9991	0.9985	0.9979

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Reduced-form results — interpretation

- **Negative effects** on overall and foreign revenues and public as expected, but **insignificant** for Brazilian movies (though point estimates are positive)
  - $\rightarrow$ At the same time, quotas increase session availability
- **Simple micro story**:
  - $\rightarrow$ Add restriction to optimization, income and moviegoers fall
  - $\rightarrow$ Quotas do shift sessions to Brazilian movies, but box-office and ticket sales lag behind
  - $\rightarrow$ Moviegoers fall less than incomes because theaters respond by lowering ticket prices.

Reduced-form results — interpretation

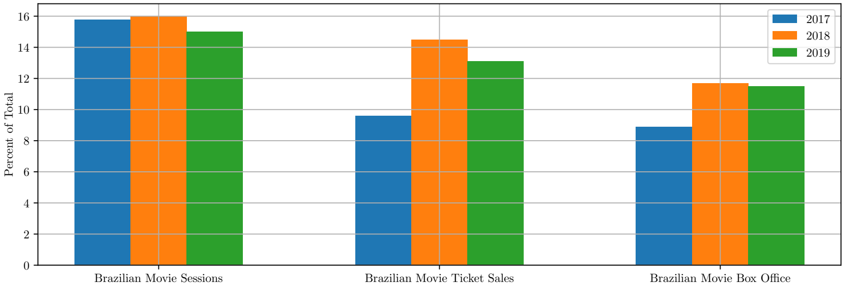


Figure 5: Brazilian Movie Sessions, Ticket Sales, and Box Office (as % of total)

Dynamic Model

Dynamic discrete-choice model

- Why model? Check reduced-form, ideally counterfactuals
- Programming is clear **discrete-choice problem**:
  - $\rightarrow$ **Only one movie** per session;
  - $\rightarrow$ Movie **attributes** partly observable, partly not.
- Screen quotas introduce a **dynamic** feature
- Very simple dynamic discrete-choice model for 12-screen multiplexes in 2018:
  - $\rightarrow$ Observables: avg occupation of movie in given week and quota requirements;
  - $\rightarrow$ Strong assumptions: exogeneity, perfect foresight

Dynamic discrete-choice model

- Each movie alternative has following "profit" function with respect to movie m, fraction of quota fulfilled x_t, and with expected occupation $o_{mt} \equiv E(o_m|t)$:

$$\pi(m_t, x_t, \varepsilon_t(m); \theta) = o_{mt} - \theta \max(0, 1 - x_t) + \varepsilon_t(m) \quad (2)$$

$$= \tilde{\pi}(m_t, x_t; \theta) + \varepsilon_t(m) \quad (3)$$

- Where $\varepsilon_t(m)$ is the private shock in time t for movie m.
- When we add the recursive term, the conditional probability of choosing movie m in time t is the familiar Logit form:

$$p(m_t|x_t) = \frac{e^{\tilde{\pi}(m_t, x_t; \theta) + \bar{V}_{t+1}(x_{t+1})}}{\sum_{j_t \in M_t} e^{\tilde{\pi}(j_t, x_t; \theta) + \bar{V}_{t+1}(x_{t+1})}} \quad (4)$$

Estimation

- Why not **ML estimation**?
  - → We could just guess, calculate all value functions and CCPs from **backwards recursion**, get LL, search, rinse and repeat;
  - → Problem: computational burden ($\approx$ 100,000 obs for 12-screen multiplexes, but full state space order of magnitudes bigger)
- Alternative: Conditional Choice Simulation (CCS) methods by Hotz and Miller (1994) and Bajari, Benkard and Levin (2007).
  - → Intuition: instead of using value funcs to compute probs, we use probs to compute val functions. Take logs of CCPs: Go to estimation

$$\frac{p(i_t|x_t,\theta)}{p(j_t|x_t,\theta)} = \frac{e^{v_t(i_t,x_t,\theta)}}{e^{v_t(j_t,x_t,\theta)}} \tag{5}$$

$$\ln p(i_t|x_t,\theta) - \ln p(j_t|x_t,\theta) = v_t(i_t,x_t,\theta) - v_t(j_t,x_t,\theta) \tag{6}$$

Table 4: Dynamic Model Parameter Estimates

	Kernel Density CCPs	*Logit CCPs*
	(1)	(2)
Expected occupation	0.00204853	−0.00825534
	(0.000027)	(0.000020)
Screen Quotas	−0.10801806	−0.01818042
	(0.002122)	(0.000119)

Results — interpretation

- Estimates point to a high private shock component variance
- All heterogeneity comes from private errors
- Quotas play a small role compared to private shocks, but enter negatively agents' value functions
- Sign of Logit for expected occupation makes no sense
- Next steps: try to add more covariates, but relevant observables are hard to get.

Conclusion

Conclusion

- Takeaways:
  - → Reduced-form regressions and dynamic model estimates suggest screen quotas play a role in exhibitor choice;
  - → Effects on revenues and ticket sales are negative, but no statistically significant impact on Brazilian movies.
- Data suggests a simple story: movie theaters indeed respond to SQ by displaying more Brazilian movies, but sessions stay largely empty
- Dynamic model does a poor job predicting programming choices, but suggests quota do play a (small) role in choices, in line with reduced-form results
- (Partial) Data, code, and (too) many other regressions available at `https://github.com/pbragasoares/`

Appendix: Additional Reduced-Form Regression Tables

Alternative kernels: Gaussian

Back

	Dependent variable:					
	log(Box Office)			log(Ticket Sales)		
	All Movies (1)	Foreign (2)	Brazilian (3)	All Movies (4)	Foreign (5)	Brazilian (6)
Compliance (Gaussian kernel)	4.8385***	6.4733***	0.0709	4.6590***	6.5440***	-0.0302
	(1.2052)	(1.3114)	(1.7934)	(1.1546)	(1.2593)	(1.7099)
Opening Days	0.0410	0.0390	0.0183	0.0476	0.0446	0.0265
	(0.0305)	(0.0303)	(0.0418)	(0.0292)	(0.0291)	(0.0399)
Quota per viewing room	0.0661***	0.0930***	-0.0258	0.0655***	0.0939***	-0.0238
	(0.0153)	(0.0167)	(0.0238)	(0.0147)	(0.0161)	(0.0227)
Compliance × Quota per VR	-0.1390***	-0.2087***	0.0853	-0.1321***	-0.2070***	0.0870
	(0.0421)	(0.0458)	(0.0635)	(0.0404)	(0.0440)	(0.0605)
Chain Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
2019?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	628	624	602	628	624	602
R^2	0.9643	0.9664	0.9328	0.9634	0.9655	0.9345
Adjusted R^2	0.9414	0.9448	0.8893	0.9400	0.9433	0.8921

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Alternative kernels: Triangular

Back

	Dependent variable:					
	log(Box Office)			log(Ticket Sales)		
	All Movies (1)	Foreign (2)	Brazilian (3)	All Movies (4)	Foreign (5)	Brazilian (6)
Compliance (Triangular kernel)	2.0112***	1.9793***	1.1049**	1.9504***	1.9819***	0.9950*
	(0.3147)	(0.3219)	(0.5447)	(0.3012)	(0.3089)	(0.5202)
Opening Days	0.0362	0.0360	0.0166	0.0428	0.0417	0.0250
	(0.0296)	(0.0298)	(0.0410)	(0.0283)	(0.0286)	(0.0391)
Quota per viewing room	0.0479***	0.0515***	0.0049	0.0491***	0.0531***	0.0066
	(0.0067)	(0.0069)	(0.0122)	(0.0064)	(0.0066)	(0.0117)
Compliance × Quota per VR	-0.0522***	-0.0543***	-0.0043	-0.0508***	-0.0541***	-0.0025
	(0.0092)	(0.0094)	(0.0160)	(0.0088)	(0.0091)	(0.0153)
Chain Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
2019?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	628	624	602	628	624	602
R^2	0.9663	0.9675	0.9355	0.9656	0.9667	0.9369
Adjusted R^2	0.9447	0.9466	0.8938	0.9435	0.9452	0.8961

*p<0.1; **p<0.05; ***p<0.01

Alternative kernels: Triweight

Back

	Dependent variable:					
	log(Box Office)			log(Ticket Sales)		
	All Movies	Foreign	Brazilian	All Movies	Foreign	Brazilian
	(1)	(2)	(3)	(4)	(5)	(6)
Compliance (Triweight kernel)	2.0112***	1.9793***	1.1049**	1.9504***	1.9819***	0.9950*
	(0.3147)	(0.3219)	(0.5447)	(0.3012)	(0.3089)	(0.5202)
Opening Days	0.0362	0.0360	0.0166	0.0428	0.0417	0.0250
	(0.0296)	(0.0298)	(0.0410)	(0.0283)	(0.0286)	(0.0391)
Quota per viewing room	0.0479***	0.0515***	0.0049	0.0491***	0.0531***	0.0066
	(0.0067)	(0.0069)	(0.0122)	(0.0064)	(0.0066)	(0.0117)
Compliance × Quota per VR	-0.0522***	-0.0543***	-0.0043	-0.0508***	-0.0541***	-0.0025
	(0.0092)	(0.0094)	(0.0160)	(0.0088)	(0.0091)	(0.0153)
Chain Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
2019?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	628	624	602	628	624	602
R^2	0.9663	0.9675	0.9355	0.9656	0.9667	0.9369
Adjusted R^2	0.9447	0.9466	0.8938	0.9435	0.9452	0.8961

*p<0.1; **p<0.05; ***p<0.01

Squared distance kernel: with and w/o 2019 Back

	Dependent variable:					
	log(Box Office)			log(Ticket Sales)		
	All Movies (1)	Foreign (2)	Brazilian (3)	All Movies (4)	Foreign (5)	Brazilian (6)
Compliance (Squared dist kernel)	2.0501***	2.0906***	0.5047	2.5331***	2.5937***	1.1897
	(0.2911)	(0.2991)	(0.5740)	(0.4452)	(0.4586)	(0.9132)
Opening Days	0.0348	0.0343	0.0187	0.0532	0.0539	0.0133
	(0.0293)	(0.0294)	(0.0411)	(0.0624)	(0.0638)	(0.0955)
Quota per viewing room	0.0553***	0.0609***	-0.0091	0.0705***	0.0770***	-0.0006
	(0.0079)	(0.0081)	(0.0161)	(0.0096)	(0.0101)	(0.0238)
Compliance $\times$ Quota per VR	-0.0541***	-0.0585***	0.0145	-0.0692***	-0.0755***	0.0005
	(0.0093)	(0.0096)	(0.0183)	(0.0146)	(0.0151)	(0.0291)
Chain Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
2019?	Yes	Yes	Yes	No	No	No
Observations	628	624	602	408	406	388
R^2	0.9671	0.9683	0.9351	0.9755	0.9759	0.9436
Adjusted R^2	0.9461	0.9480	0.8931	0.9464	0.9473	0.8753

*p<0.1; **p<0.05; ***p<0.01

Dynamic model estimation

CCPs 1st stage estimation

- But how can we obtain CCP estimates?
- The literature suggests avoiding overly parametric assumptions, but state-space is too large for simple bin estimators (many possible state-choice pairs are not available in the data)
- We therefore try two approaches:
  - → Gaussian kernel density estimators in the x_t/day space, to get densities and compute probabilities from relative densities;
  - → Flexible Logit using movie-theater and day FE, and the state x_t.

Estimation Algorithm — 1st stage

- Having at our disposal the CCPs for every possible movie and state x_t, we start from $t = 0$ and follow the steps:
    1. Starting at $x_0 = 0$, draw random shocks for each choice;
    2. Calculate the chosen movie i, i.e., the movie such that $v_t(i_t, x_t, \theta) + \varepsilon_t(i) > v_t(j_t, x_t, \theta) + \varepsilon_t(j)$, $\forall j_t \in M_t$;
    3. Get a new state x_1 given the choice and the transition function $x_1 = f(0, \delta(0, \varepsilon_0), a_0, q)$;
    4. Repeat 1-3 for the next state until the terminal state $t = T$ is reached.
- Having all the choices and associated shocks, we can easily calculate an estimate for the *ex ante* discounted value function an agent i, $\hat{\hat{V}}_{0i}(0;\theta)$. We then average out the function over 20 simulated paths to get consistent estimates for $\hat{\hat{V}}_{0i}(0;\theta)$ for each agent

Estimation Algorithm — 2nd stage Back

- To get parameter estimates, we repeat 1st procedures with disturbed value functions (adding noise or systematic bias in 1st stage CCPs)
- We then get parameters such that they minimize Equilibrium violations (i.e., minimize the square error of disturbed functions having value bigger than proper val functions):

$$(\hat{\theta}_1, \hat{\theta}_2) = \arg\min_{(\theta^1, \theta^2)} \sum_{i=1}^{6} \sum_{n=1}^{5} (\min\{0, \hat{\hat{V}}_{0i}(0; \theta) - \hat{D}_{0i}^{(n)}(0)\})^2 \tag{7}$$