Online Appendix

Estimating the Costs of Standardization: Evidence from the Movie Industry

El Hadi Caoui*

March 11, 2022

To distinguish the tables, figures, and equations presented in the appendix from those in the manuscript, their numbering is preceded by the letter A.

A Theater adoption panel: data sources

Table A1 presents the observation dates for the panel of digital projector adoption by data sources. The table also shows the periodic subsample selected. The periods selected are such that there is a previous observation period 6 months earlier (in some exeption it is 5 or 7 months). For instance, "May 2012" is selected because the industry is observed on November 2011. The observation periods selected are represented in blue in Figure A1.

B Evidence of network effects in digital screen adoption

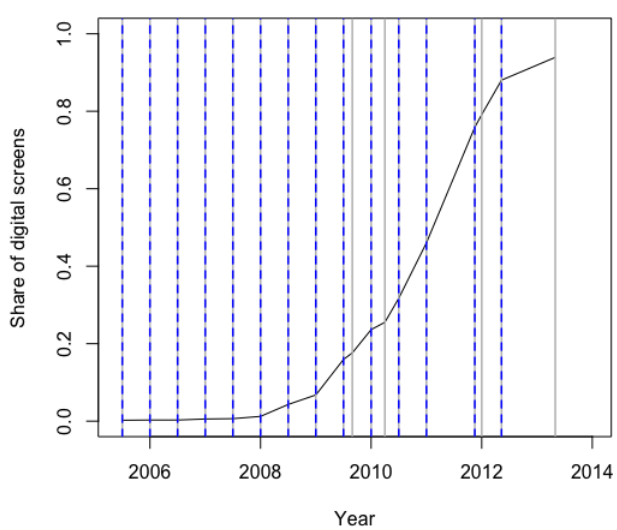
This appendix presents estimates of the causal effect of digital movie availability on digital screen adoption. Estimating this effect using only data on the French market is challenging because only one network is observed over a relatively short time horizon and all the variation in software availability is in the time-series. This appendix complements the analysis in the main text by employing supplementary data on adoption for a large panel of countries. The effect is identified from cross-sectional and time-series variation in digital movie availability. Identification relies on instrumental variables which affect digital movie availability

*Department of Management (UTM) and Rotman School of Management, University of Toronto. 105 St. George Street, Toronto, ON M5S 3E6, Canada. Email: elhadi.caoui@rotman.utoronto.ca.

Table A1: Observation times by data source

Date	Source	Periodic sample
July 2005	Media Salles	✓
January 2006	Media Salles	✓
July 2006	Media Salles	✓
January 2007	Media Salles	✓
July 2007	Media Salles	✓
January 2008	Media Salles	✓
July 2008	Media Salles	✓
January 2009	Media Salles	✓
July 2009	Media Salles	✓
September 2009	Cinego	
Jarnuary 2010	Media Salles	✓
April 2010	Cinego	
July 2010	Media Salles	✓
January 2011	Media Salles	✓
November 2011	Cinego	
January 2012	Media Salles	
May 2012	Cinego	✓
June 2013	Cinego	

Figure A1: Share of digitally equipped screens and observation times



in country c at time t but are, otherwise, independent of other factors affecting demand for digital screens in (c, t).

B.1 Aggregate data by country

This section presents the data used in the reduced-form analysis. Aggregate data on the number of digital screens (quarterly, $2005-2017$) and the number of digital movies released (semesterly, $2005-2010$) is obtained for 59 countries, from the Cinema Intelligence Service database of Omdia (previously, IHS Markit), a market research firm.[1]

This data is complemented by a rich set of variables on domestic movie industries from the UNESCO Institute of Statistics (UIS).[2] The UIS data contains more than 75 variables on the movie exhibition and distribution industries by country-year over the time period $1995-2017$. I extract variables that are relevant controls for digital screen adoption: namely, average box-office revenue per screen (converted to US$), and the distribution of theater sizes (fraction of miniplexes, multiplexes etc.). Additionally, I extract information on the market share (of box-office revenue) of U.S. distributors and their joint ventures with domestic distributors in each country-year. The major U.S. distributors are: United International Pictures (a venture of Paramount and Universal), Buena Vista Distribution Company (Walt Disney Studios Motion Pictures), 20th Century Fox, Warner Bros. Entertainment Inc., Columbia Pictures (Sony Pictures). Joint ventures of major U.S. and domestic distributors include, for example, Sandrew Metronome/Warner Bros in Sweden, Roadshow/Warner Bros in Australia and New Zealand.

The availability of aggregate data by country is shown in Table A2. Due to missing data for digital movies, the reduced form analysis is conducted on a subset of 29 countries with complete semesterly information on digital movies and screens between 2005 and 2010. Figure A2 shows the market share of U.S. distributors (and their joint ventures with domestic firms) for the subset of countries in 2005.

Software variety and hardware adoption are determined jointly and endogenously, generating a simultaneity problem: the availability of digital movies could be correlated with unobserved country-time level factors that shift the demand for digital screens which, in turn, affects the supply of digital movies. I handle this issue using instrumental variables. The identification strategy exploits the fact that whether a U.S.-produced movie is released in digital in country c at time t is at least partly a function of whether the movie was released

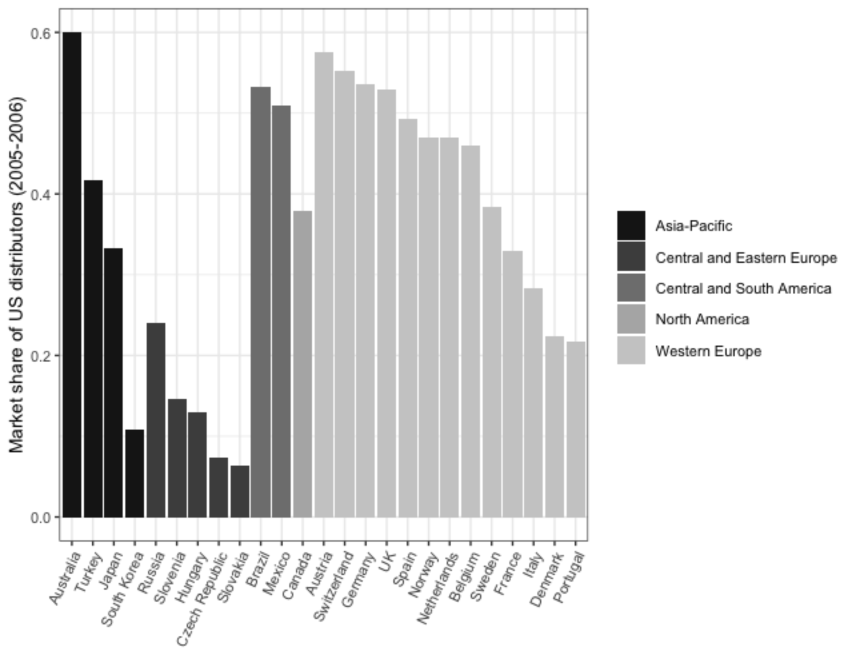
[1] `https://technology.informa.com/Services/424103/cinema-intelligence-service/Data`, section "Digital Cinema".

[2] `http://data.uis.unesco.org`, section "Culture, Feature Films". One of the source for the UIS data, in particular concerning European countries, are the Cinema Yearbooks used in the main text.

Table A2: Aggregate data availability by country

Region	Country	Digital screens (quarterly, 2005-2017)	Digital movies (semesterly, 2005-2010)	Regressors (IV, controls)
Africa Middle-East	Egypt	✓		
	Israel	✓		✓
	Morocco	✓		✓
	South Africa	✓		
	UAE	✓		
Asia-Pacific	Australia	✓	✓	✓
	China	✓	✓	
	Hong Kong	✓		
	India	✓		
	Indonesia	✓		
	Japan	✓	✓	✓
	Malaysia	✓		✓
	New Zealand	✓		
	Philippines	✓		
	Singapore	✓	✓	
	South Korea	✓	✓	✓
	Taiwan	✓		
	Thailand	✓		
	Turkey	✓	✓	✓
Central and Eastern Europe	Bulgaria	✓		✓
	Croatia	✓		✓
	Czech Republic	✓	✓	✓
	Estonia	✓		✓
	Hungary	✓	✓	✓
	Latvia	✓		✓
	Lithuania	✓		
	Poland	✓		✓
	Romania	✓		✓
	Russia	✓	✓	✓
	Slovakia	✓	✓	✓
	Slovenia	✓	✓	✓
	Ukraine	✓		✓
Central and South America	Argentina	✓		✓
	Brazil	✓	✓	✓
	Chile	✓		✓
	Colombia	✓		
	Mexico	✓	✓	✓
	Venezuela	✓		✓
North America	Canada	✓	✓	✓
	United States	✓	✓	✓
Western Europe	Austria	✓	✓	✓
	Belgium	✓	✓	✓
	Cyprus	✓		✓
	Denmark	✓	✓	✓
	Finland	✓		✓
	France	✓	✓	✓
	Germany	✓	✓	✓
	Greece	✓		
	Iceland	✓		✓
	Ireland	✓		
	Italy	✓	✓	✓
	Luxembourg	✓	✓	
	Netherlands	✓	✓	✓
	Norway	✓	✓	✓
	Portugal	✓	✓	✓
	Spain	✓	✓	✓
	Sweden	✓	✓	✓
	Switzerland	✓	✓	✓
	United Kingdom	✓	✓	✓

Figure A2: Market share of U.S. distributors (and joint ventures) by country



in digital in the U.S.—the more so if distributors in country c are subsidiaries of the major U.S. studios or joint ventures of domestic distribution companies and U.S. distributors. The release of digital movies in the U.S., on the other hand, was arguably not affected by the roll-out of digital screens in relatively small foreign movie markets (e.g., Denmark, Luxembourg, Mexico etc.).[3]

I construct a set of country-time-specific instruments for the number of digital movies in country-time (c, t) by interacting: (i) the number of digital movies in the U.S. in t and (ii) the market share of U.S. distributors (or joint ventures) in $(c, t = 0)$. Importantly, market shares are for $t = 0$ (i.e., at the beginning of the diffusion in 2005) because market shares of U.S. distributors in $t = 1, 2, ..$ may have been affected by the roll-out of digital screens in country c (for instance, a positive demand shock for digital screens in (c, t) would raise the demand for digital movies; if U.S. distributors are the only firms supplying movies in digital in (c, t), their market share would increase).

[3]This approach is close in spirit to the third identification strategy of Gowrisankaran and Stavins (2004) which measures network effects via the response of small banks to the adoption by small branches of large banks. The identification strategy used in the current paper is arguably most informative about the effect of *U.S.-produced* digital movie availability on digital screen adoption.

B.2 Reduced form analysis

To estimate the effect of digital movie availability on digital screen adoption, I regress the number of digital screens in country c (excluding the U.S.) at time t, on: the number of movies released in digital in (c, t), average box-office revenue per screen in (c, t) (in, US$), and the fraction of multiplexes in (c, t).[4] Region-time effects are included to remove unobserved time-varying factors at the regional level such as adoption costs (see Table A2 for the list of regions). In addition, I control for time-invariant country-category effects, where I use the total number of screens in each country to group countries into 6 categories. Standard errors are clustered at the country-level.

I do not allow for country fixed effects due to the small amount of data. Such country effects wash out a large part of the correlation between digital releases and screen adoption coming from cross-sectional variation (this issue is related to the attenuation bias discussed in Griliches and Hausman (1986)). The use of country-categories instead aims at navigating the trade-off between removing country-level fixed unobservables from the error term that may be correlated with the instrument and the attenuated effect of digital movies. I experiment with different definition of country-categories: e.g., total box office revenue in 2005, GDP per capita, etc. These alternative definition of country categories yield similar estimates.

Table A3 (Columns 1–3) show the OLS results. In specification (3), the elasticity of digital screens with respect to digital movies is 1.012. Adoption of digital screens is also positively correlated with average revenues per screen (a function of average admissions and ticket prices) and the fraction of multiplexes in the industry.

Table A3 (Columns 4–5) show the first and second-stage regression of the IV model. Once the endogeneity of digital movie availability is accounted for, the elasticity of digital screens with respect to digital movies decreases to 0.806. This is consistent with the expected direction of the bias for hardware-software systems with positive network effects.[5] Table A3 also reports the first-stage (Kleibergen-Paap Wald rk) F-statistic which is above the critical thresholds derived in Stock and Yogo (2005), indicating that the instrument strongly predicts the endogenous regressor.[6]

[4]Revenues per screen and the fraction of multiplexes are assumed exogenous. For the former, this assumption is consistent with digital technology being a cost-reducing technology. The theater size distribution may have been affected by the conversion to digital; in particular, for developing movie markets such as China. The results are robust when the sample is restricted to established markets with a stable theater size distribution.

[5]Although a chi-squared test fails to reject the null hypothesis that the specified endogenous regressor can be treated as exogenous.

[6]This F-statistic is used because the standard Cragg-Donald Wald F-statistic is not valid with cluster-robust standard errors.

Table A3: Effect of digital movie availability on digital screen adoption

	OLS			IV	
Dep. variable: Digital screens (log)	(1)	(2)	(3)	(4) First-stage	(5) Second-stage
Digital movies (log)	1.161 (0.0614)	1.176 (0.146)	1.012 (0.202)		0.806 (0.342)
Multiplexes (8+ screens, in %)			0.957 (0.492)	1.655 (0.415)	1.256 (0.560)
Revenue per screen (log, US$)			0.286 (0.114)	0.358 (0.171)	0.380 (0.229)
Digital movies in the US × Share of US distributors (log)				0.538 (0.138)	
Country-category FEs	No	No	Yes	Yes	Yes
Region × Time FEs	No	Yes	Yes	Yes	Yes
Kleibergen-Paap rk Wald F-stat					15.21
Endogeneity test (p-value)					0.64 (0.42)
Observations	300	288	288	288	288

Note: Unit of observation: country/semester for all countries in the sample excluding the U.S. Standard errors (in parenthesis) are clustered at the country-level. "Digital movies in the US" refers to the number of digital releases in the U.S. at time t. "Market share of US distributors" is in country c by the end of year 2005 (i.e., in period $t = 0$).

C Identification and Estimation

C.1 Identification approach

This section provides more details regarding the identification approach. First, a proof of Lemma 1 is presented.

Proof. From Equation (17), we have that for all $a \in \{1, \dots, S_{\boldsymbol{\tau}}\}$

$$\begin{aligned} \Delta\widetilde{W}_t(a, a-1|\boldsymbol{\tau}, s; \boldsymbol{\Pi}) &= \beta\left(\widetilde{V}_{\boldsymbol{\tau},t+1}(s+a; \boldsymbol{\Pi}) - \widetilde{V}_{\boldsymbol{\tau},t+1}(s+a-1; \boldsymbol{\Pi})\right) \\ &= \widetilde{W}_t(1|\boldsymbol{\tau}, s+a-1; \boldsymbol{\Pi}) - \widetilde{W}_t(0|\boldsymbol{\tau}, s+a-1; \boldsymbol{\Pi}) \\ &= \Delta\widetilde{W}_t(1, 0|\boldsymbol{\tau}, s+a-1; \boldsymbol{\Pi}) \end{aligned} \tag{A1}$$

The predicted differences in choice specific value functions can be rewritten as a linear combination of predicted differences between actions 1 and 0, for all $(a_{k+1}, a_k) \in Supp(\boldsymbol{\tau}, s, t)$

$$\begin{aligned} \Delta\widetilde{W}_t(a_{k+1}, a_k|\boldsymbol{\tau}, s; \boldsymbol{\Pi}) &= \Delta\widetilde{W}_t(a_{k+1}, a_{k+1}-1|\boldsymbol{\tau}, s; \boldsymbol{\Pi}) + \ldots + \Delta\widetilde{W}_t(a_k+1, a_k|\boldsymbol{\tau}, s; \boldsymbol{\Pi}) \\ &= \Delta\widetilde{W}_t(1, 0|\boldsymbol{\tau}, s+a_{k+1}-1; \boldsymbol{\Pi}) + \ldots + \Delta\widetilde{W}_t(1, 0|\boldsymbol{\tau}, s+a_k; \boldsymbol{\Pi}) \end{aligned} \tag{A2}$$

Since actions $a = 1$ and $a = 0$ are played with positive probability, $\Delta W_t(1, 0|\boldsymbol{\tau}, s)$ is identified from the data. The linear dependence between optimality conditions shown in Equation (A2) implies that only the set of conditions for $a = 1$ and $a = 0$ provide identifying restrictions for $\boldsymbol{\Pi}$. The set of restrictions is

$$\Delta W_t(1, 0|\boldsymbol{\tau}, s) = \Delta\widetilde{W}_t(1, 0|\boldsymbol{\tau}, s; \boldsymbol{\Pi}) \tag{A3}$$

for every firm type $\boldsymbol{\tau} \in \mathcal{T}$, firm state $s \in \{0, 1, \dots, S_{\boldsymbol{\tau}} - 1\}$, and period $t = 1, \dots, \overline{T} - 1$. The optimality conditions (A3) form a (underdetermined) system of $(L - \overline{\tau}) \cdot (\overline{T} - 1)$, equations in the unknown parameter $\boldsymbol{\Pi}$.[7] □

Second, the system of linear equations (Equations (21)) in the unknown parameter $\boldsymbol{\Pi}$ is obtained as follows. Expressing the differences in choice-specific value functions as a function of the ex-ante value function (Equation (17)), we obtain

$$\Delta W_t(1, 0|\boldsymbol{\tau}, s) = \beta\left(\widetilde{V}_{\boldsymbol{\tau},t+1}(s+1; \boldsymbol{\alpha}) - \widetilde{V}_{\boldsymbol{\tau},t+1}(s; \boldsymbol{\alpha})\right)$$

[7] Recall that L equals $\sum_{\boldsymbol{\tau} \in \mathcal{T}} (S_{\boldsymbol{\tau}} + 1)$. Therefore the number of optimality conditions is $\sum_{\boldsymbol{\tau} \in \mathcal{T}} (S_{\boldsymbol{\tau}}) \cdot (\overline{T} - 1) = (L - \overline{\tau}) \cdot (\overline{T} - 1)$.

Collecting these equations for every firm types $\boldsymbol{\tau} \in \mathcal{T}$, firm state $s \in \{0, 1, \ldots, S_{\boldsymbol{\tau}} - 1\}$, and period $t = 1, \ldots, \overline{T} - 1$, in matrix form, and substituting the ex-ante value function (Equation (16)), we obtain a system of linear equations in the unknown parameter $\mathbf{\Pi}$

$$\begin{aligned} \Delta\widetilde{\mathbf{W}} &= \beta \mathbf{H} \widetilde{\mathbf{V}} \\ &= \beta \mathbf{H}(\mathbf{I} - \beta \cdot \mathbf{F})^{-1} \mathbf{\Pi} - \beta \mathbf{H}(\mathbf{I} - \beta \cdot \mathbf{F})^{-1} \sum_{a} \mathbf{P}(a) a [\mathbf{p} + \mathbf{e}(a)] \end{aligned} \tag{A4}$$

where $\mathbf{H}$ is $((L - \overline{\tau}) \cdot (\overline{T} - 1)) \times (L \times \overline{T})$-dimensional with value -1 in row $(\boldsymbol{\tau}, s, t)$ and column $(\boldsymbol{\tau}, s, t + 1)$ and value $+1$ in row $(\boldsymbol{\tau}, s, t)$ and column $(\boldsymbol{\tau}, s + 1, t + 1)$, $\Delta\widetilde{\mathbf{W}}$ is a $((L - \overline{\tau}) \cdot (\overline{T} - 1)) \times 1$-dimensional vector with row $(\boldsymbol{\tau}, s, t)$ equal to $\Delta W_t(1|\boldsymbol{\tau}, s)$, and $\widetilde{\mathbf{V}}$ is the $(L \cdot \overline{T}) \times 1$ dimensional vector of ex-ante value functions. The linear system in $\mathbf{\Pi}$ can be written

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{\Pi} \tag{A5}$$

where $\mathbf{Y}$ and $\mathbf{X}$ are a $((L - \overline{\tau}) \cdot (\overline{T} - 1)) \times 1$-dimensional and $((L - \overline{\tau}) \cdot (\overline{T} - 1)) \times (L \times \overline{T})$-dimensional matrices which are function of the (known) CCP, prices, discount factor β, and distribution of the firm-specific shock F.

C.2 Parametric bootstrap procedure

This section provides details about the bootstrap procedure used to compute standard errors for the structural parameter estimates. The non-parametric bootstrap approach used in the dynamic oligopoly game literature typically relies on the availability of a cross-section of independent markets and assumes that the same equilibrium is played in all markets. Standard errors can be calculated by sampling market-histories (with replacement) and estimating the parameter of interest for each cross-section of markets sampled.

This approach would not be valid in the current setting because only a single industry is observed and firms' decisions are correlated through network effects (at the industry-level). Sampling firm-histories (or even local markets, e.g. urban/rural units-histories) would destroy the dependence between observations.

Instead I employ a parametric bootstrap approach. The main idea is to rely on the (parametric) model, or data generating process, which incorporates the dependence between observations, to generate the bootstrap samples on which estimation is conducted.

Let $\mathbf{D} = \{\mathbf{x}_1, \{a_{it}\}_{i \in \mathbf{N}, t=1,\ldots,\overline{T}}\}$ denote the data observed, where

$$\mathbf{x}_1 = ((\boldsymbol{\tau}(1), s_{11}), \ldots, (\boldsymbol{\tau}(N), s_{N1}))$$

is the initial state of the industry including firm types, and $s_{i1} = 0$ for all i (in the first period, all firms operate using film screens).[8] The data $\mathbf{D}$ is assumed to be generated from the set of equilibrium CCP $\{P_t(a_{it}|\boldsymbol{\tau}(i), s_{it})\}$. I assume that the distribution of CCP is parametric, namely, it follows an ordered probit model with parameters $\boldsymbol{\theta}$ (governing the dependence of the CCP on $(t, \boldsymbol{\tau}(i), s_{it})$).[9] The parameters in $\boldsymbol{\theta}$ are *reduced-form* first-stage parameters entering the CCP.

The structural parameter of interest $\boldsymbol{\alpha}$ is a functional of the CCP. Let $\widehat{\boldsymbol{\theta}}$ be an estimate of the reduced-form parameters $\boldsymbol{\theta}$ obtained from the data $\mathbf{D}$ (shown in Table 4 of the main article). Let $\widehat{\boldsymbol{\alpha}}$ be an estimate of the structural parameters $\boldsymbol{\alpha}$ obtained from $\mathbf{D}$ and $\widehat{\boldsymbol{\theta}}$ (via the second step of the two-step approach detailed in Section 6.1). I denote the estimates of the CCP by $\{P_t(a_{it}|\boldsymbol{\tau}(i), s_{it}; \widehat{\boldsymbol{\theta}})\}$.

Standard errors for $\widehat{\boldsymbol{\alpha}}$ are obtained by following a parametric bootstrap procedure (see Efron and Tibshirani (1993), sections 8.4 and 8.5 on "more general data structures"). Throughout the procedure, the initial industry state $\mathbf{x}_1$ is fixed.

1. Simulate B boostrap datasets $(\mathbf{D}_1, \ldots, \mathbf{D}_B)$ from the estimated CCP $\{P_t(a_{it}|\boldsymbol{\tau}(i), s_{it}; \widehat{\boldsymbol{\theta}})\}$. The bootstrap datasets have the same size as the original dataset $\mathbf{D}$.

2. For each bootstrap dataset indexed by $b \in \{1, \ldots, B\}$, follow the two-step procedure of Section 6.1: first, obtain estimates of the reduced-form first-stage parameters $\widehat{\boldsymbol{\theta}}_b$. Second, given data $\mathbf{D}_b$ and estimate $\widehat{\boldsymbol{\theta}}_b$, compute an estimate $\widehat{\boldsymbol{\alpha}}_b$ of the structural parameter.

The requirement that the initial industry state $\mathbf{x}_1$ (in particular, firm types) is kept fixed—rather than, for instance, sampling firms with replacement and simulating bootstrap samples—guarantees that the equilibrium played in dataset $\mathbf{D}$ is also an equilibrium in the bootstrap dataset $\mathbf{D}_b$: the equilibrium CCP $\{P_t(a_{it}|\boldsymbol{\tau}(i), s_{it}; \widehat{\boldsymbol{\theta}})\}$ are still valid and can be used to simulate firms' adoption decision in the bootstrap samples. This a consequence of the fact that the process $\{\overline{h}_t(\mathbf{a}^{no})\}_{t=1,\ldots,\overline{T}}$, to which firms are best-responding, is only a function of the initial industry state $\mathbf{x}_1$ and equilibrium strategy $\mathbf{a}^{no}$.[10]

[8] The remaining aggregate variables $(p_t, h_{US,t})$ are assumed deterministic and are index by t.

[9] Note that observations $\{a_{it}\}_{i\in\mathbf{N}}$ for a given t are independent and identically distributed conditional on $(t, \boldsymbol{\tau}(i), s_{it})$. This a consequence of the distributional assumption on ϵ_{it}. The curse of dimensionality prevents estimating the (non-parametric) empirical distribution of $a_{it}|\boldsymbol{\tau}(i), s_{it}$.

[10] If alternatively, firm types were also sampled with replacement, the deterministic process $\{\overline{h}_t\}$ would change, and firms' best-responses to it would also change.

While the parametric bootstrap approach relies heavily on the correct specification of the (parametric) data generating process, it can be viewed as one path forward in dealing with dependent data in the context of dynamic games with a small number of markets but a large number of firms. Such dependence arises, in particular, when network effects are present.

C.3 First step: adoption policy rule

First step estimates for the adoption policy rule are obtained by estimating an ordered probit model. Denote by P_{ij} the probability that theater i transitions to state j. Possible states are $\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ in the case of miniplexes ($4-7$ screens), and $\{0, \frac{1}{8}, \frac{2}{8}, ..., \frac{7}{8}, 1\}$ in the case of multi/megaplexes (8 screens or more). In constructing the likelihood, one has to account for the fact that theaters cannnot divest digital screens, and therefore, cannot transition to lower states: the dependent variable $s_{it}/S_{\boldsymbol{\tau}(i)}$ satisfies $s_{it} \geq s_{i,t-1}$. The log likelihood is constructed as follows:

$$LL = \sum_{i:4 \leq S_{\boldsymbol{\tau}(i)} < 8} \sum_{j=s_{i(t-1)}/S_{\boldsymbol{\tau}(i)}}^{1} d_{ij} \log\left(P_{ij}\right) + \sum_{i:8 \leq S_{\boldsymbol{\tau}(i)}} \sum_{j=s_{i(t-1)}/S_{\boldsymbol{\tau}(i)}}^{1} d_{ij} \log\left(P_{ij}\right) \qquad \text{(A6)}$$

where d_{ij} is an indicator for firm i transitioning to state j.

C.4 First step: goodness of fit

To check the goodness of fit of the first-step adoption policy (CCP), model predictions for the share of digital screens are compared to actual shares in the data. Tables A4, A5, and A6 present the comparison for all firms, miniplexes only, and multi/megaplexes, respectively. In each table, the aggregate share of digital screens, the share of adopters (theaters with at least one digital screen), and the average within-theater share of digital screens (among adopters) are shown from 2006 to 2013. Overall, given the limitations imposed by the parametric specification of the policy function, the model captures the main trends in the aggregate, inter-firm, and intra-firm diffusion rates, for all firms and by firm size (miniplexes vs. multi/megaplexes).

The aggregate share of digital screens was constantly lower for miniplexes than for multi/megaplexes, as reflected in the predictions as well. Additionally, the intra-firm rates' evolution over time is smoother in the prediction than in the data.[11]

[11] In particular, for miniplexes, the intra-firm rate rises to 41% as early as 2007, whereas the model predicts a slow increase between 2006 and 2009 to reach 40%. The prediction is also smoother in the case of multi/megaplexes, with an increase in the actual intra-firm rates from 13.5% in 2008 to 41.5% in 2009, whereas the model predicts a smoother transition. Due to the limited number of firms with at least one digital screen in the initial years ($2006-2008$), predicting the intra-firm rate (computed only on theaters

Table A4: Predictions using the adoption policy function - All firms

	Aggregate		Inter-firm		Intra-firm	
Year	Data	Prediction	Data	Prediction	Data	Prediction
2006	0.003	0.000	0.021	0.000	0.127	0.130
2007	0.006	0.000	0.028	0.001	0.210	0.188
2008	0.012	0.006	0.056	0.020	0.182	0.234
2009	0.068	0.049	0.122	0.129	0.420	0.296
2010	0.236	0.187	0.431	0.387	0.459	0.399
2011	0.460	0.418	0.684	0.693	0.583	0.535
2012	0.791	0.664	0.841	0.907	0.880	0.689
2013	0.939	0.879	0.934	0.988	0.985	0.877

Note: The column labelled "Aggregate" corresponds to the share of digital screens across all firms in the industry. The column labelled "Inter-firm" corresponds to the share of theaters with at least one digital screen. The column labelled "Intra-firm" corresponds to the within-theater average share of digital screens among theaters with at least one digital screen. The predicted rates are obtained by averaging 500 simulation paths.

Table A5: Predictions using the adoption policy function - Miniplexes (4-7 screens)

	Aggregate		Inter-firm		Intra-firm	
Year	Data	Prediction	Data	Prediction	Data	Prediction
2006	0.002	0	0.008	0	0.267	0
2007	0.007	0.000	0.017	0.000	0.415	0.251
2008	0.007	0.002	0.017	0.004	0.415	0.339
2009	0.024	0.017	0.054	0.044	0.435	0.362
2010	0.112	0.092	0.243	0.206	0.443	0.421
2011	0.294	0.277	0.498	0.529	0.569	0.506
2012	0.653	0.546	0.745	0.845	0.857	0.636
2013	0.879	0.838	0.891	0.980	0.975	0.853

Note: The columns are defined in the same way as in Table A4, but the reference group is miniplexes instead of all firms. The predicted rates are obtained by averaging 500 simulation paths.

Table A6: Predictions using the adoption policy function - Multi/Megaplexes (8-23 screens)

	Aggregate		Inter-firm		Intra-firm	
Year	Data	Prediction	Data	Prediction	Data	Prediction
2006	0.003	0.000	0.037	0.000	0.087	0.129
2007	0.005	0.000	0.043	0.002	0.107	0.180
2008	0.015	0.009	0.106	0.037	0.135	0.221
2009	0.092	0.065	0.207	0.226	0.415	0.279
2010	0.307	0.236	0.670	0.594	0.466	0.389
2011	0.554	0.491	0.920	0.879	0.593	0.553
2012	0.870	0.726	0.963	0.978	0.903	0.741
2013	0.973	0.900	0.989	0.998	0.997	0.904

Note: The columns are defined in the same way as in Table A4, but the reference group is multi/megaplexes instead of all firms. The predicted rates are obtained by averaging 500 simulation paths.

D Counterfactual analysis

D.1 Solution to the planner's benchmark via perturbation

Details about the approach to solve for the planner's benchmark are presented here. I generate a large number of random industry adoption paths and select the path that maximizes the objective function in problem (25) (in the main text).

The simulated paths are generated by adding perturbations to the equilibrium cut-offs in each theater's CCP. For a given vector of perturbations $\{\xi_i\}_{i \in I}$, the simulated path is obtained using the following procedure.

1. Initialize the industry at $\mathbf{x}_1$ such that $s_{i1} = 0$ for all i.
2. Draw firm specific adoption shocks $\{\epsilon_{i1}\}_{i \in I}$ and corresponding adoption decisions dictated by each firm's perturbed CCP.
3. Calculate single-period industry profits $\sum_{i \in I} \Pi(a_{i1}, \mathbf{x}_1, p_1, h_{US,1}, \epsilon_{i1})$.
4. Update the current state $\boldsymbol{y}_1 = (\mathbf{x}_1, p_1, h_{US,1})$ according to the adoption decisions and transition of the exogeneous processes to next period state: $\boldsymbol{y}_2$. In particular,

$$h_2 = \left(\frac{s_2}{S}\right)^{\eta_s} h_{US,2}^{\eta_h}$$

with at least one digital screen) in this initial adoption phase is more challenging.

and

$$s_{i2} = s_{i1} + a_{i1}$$

5. Repeat steps 1-4 for $\overline{T}$ periods.

Given a vector of perturbations $\{\xi_i\}_{i\in I}$, the objective function is obtained by averaging L simulated paths. The paths have length $\overline{T} = 30$ periods (or 15 years).[12] For $t \geq \overline{T}$, the industry is its long-run steady state and profits are constant ($s_{it} = S_{\boldsymbol{\tau}(i)}$ and $h_t = 1$). An estimate of the objective function is obtained as

$$\frac{1}{L}\sum_{l=1}^{L}\left\{\sum_{i\in I}\sum_{t=1}^{\infty}\beta^{t-1}\Pi^l(a_{it}, \mathbf{x}_t, p_t, h_{US,t}, \epsilon_{it})\right\} \tag{A7}$$

where $\Pi^l(a_{it}, \mathbf{x}_t, p_t, h_{US,t}, \epsilon_{it})$ is the single-period profit of firm i in simulation l at period t, when the firm follows the adoption strategy obtained by adding the perturbation ξ_i to its equilibrium CCP. In practice, I draw $K = 15,000$ perturbation vectors and compute the objective function for each vector $\{\xi_i\}_{i\in I}$. I select the perturbation vector that yields the highest value of the estimated objective function (A7).

Figure A3a shows an example of simulated industry adoption paths. Each simulated path corresponds to a perturbation vector $\{\xi_i\}_{i\in I}$.[13] Figure A3b plots the effect of 500 additional perturbation draws on the maximum value of the estimated objective function (expression (A7)) as a function of the total number of perturbation draws. Beyond $10,000$ perturbation draws, the incremental improvement in the maximized objective function is less than $2,000$ € (i.e. less than 0.002% of aggregate industry profits).

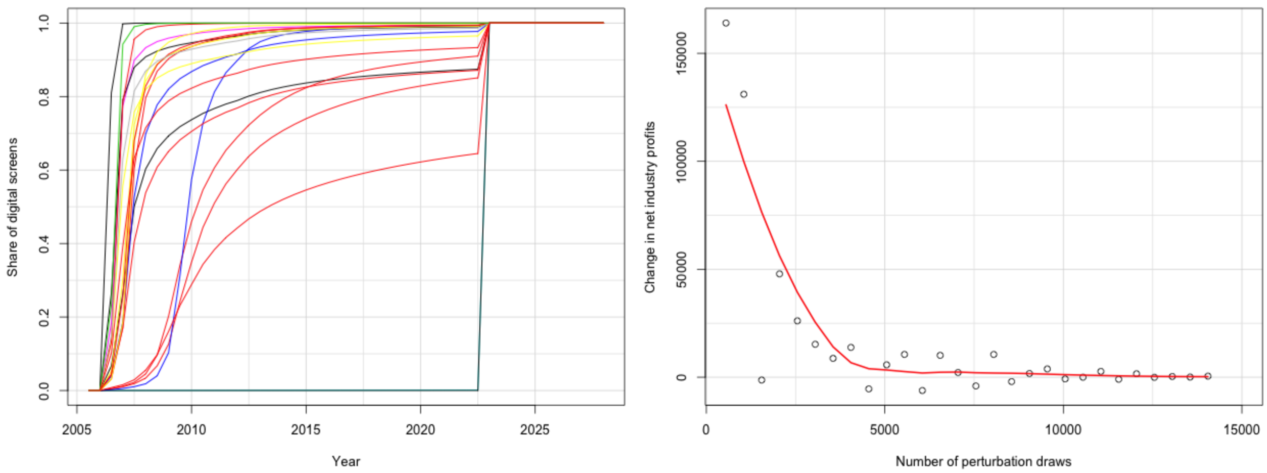
D.2 Full-solution approach to the Planner's benchmark

This section presents the full-solution approach to the planner's problem. The planner's benchmark is solved by backward induction starting from $\overline{T}$. To address the high-dimensionality problem, I maintain the reduced-form specification characterizing h_t as in Section 7.1. Additionally, the planner's adoption strategy space is coarsened by assuming group-symmetric strategies: I assign theaters to G groups based on profits $\pi_d(\boldsymbol{\tau}(i))$ and the planner chooses the same adoption decision (adoption rate) for theaters within the same group.

[12] In the equilibrium market outcome, firms have an incentive to free-ride on other firms' adoption, so the speed of digital conversion is slower than what the profit-maximizing industry planner would choose. Therefore, it is expected that the planner will have switched the industry to digital sooner than in the non-cooperative market equilibrium.

[13] For each firm i, I transform the equilibrium cut-offs κ in firms CCP (estimated by ordered probit) into $\kappa + \xi_i$ where ξ_i is drawn from $\mathcal{N}(X, Y)$ where $X \sim \mathcal{N}(-1.5, 2)$ and $Y \sim \mathcal{U}[0.5, 2]$.

Figure A3: Planner's benchmark



(a) Simulated industry paths (20 perturbation draws)

(b) Effect of 500 additional perturbation draws on maximized industry profits

These assumptions reduce the dimensionality of the state space tracked by the planner (excluding theaters' idiosyncratic shocks). However, within-group adoption strategies may in general still be a function of the realization of the vector of idiosyncratic shocks ϵ_{it}, a high-dimensional vector. I deal with this issue by noting that adoption decisions depend only on the *average adoption cost* across theaters within a group, i.e. (for N_g theaters in group g):

$$p_t + \frac{1}{N_g} \sum_{i \in g} \epsilon_{it}$$

Appealing to the law of large numbers, the second term is close to zero. By this argument, I can approximate the planner's problem as a non-stochastic problem.[14]

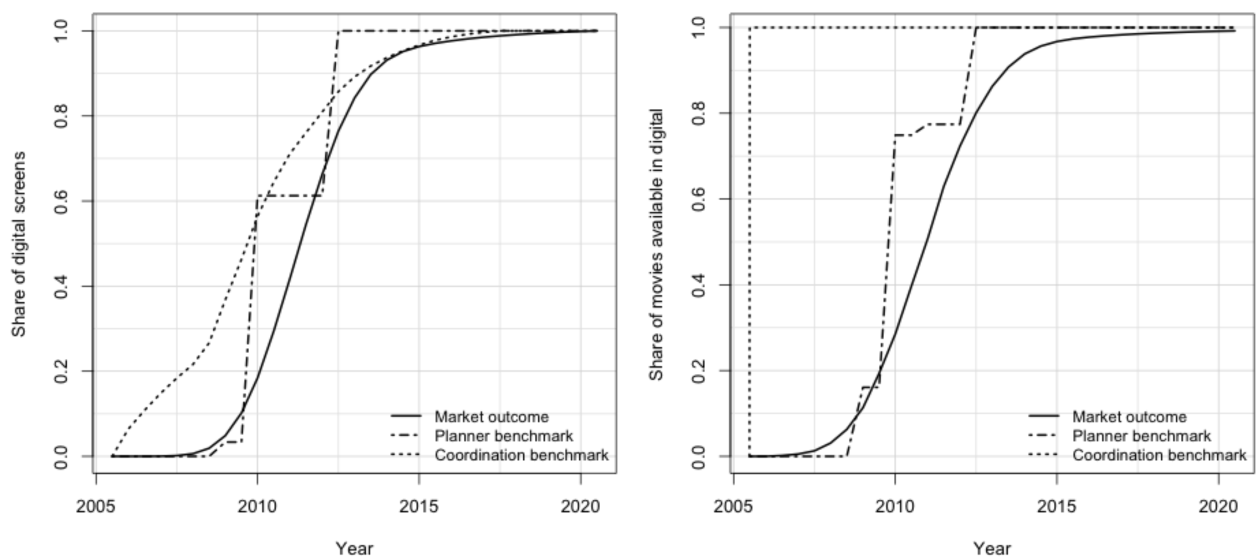
The solution is found for different values of G. Figures A4a and A4b show the results for $G = 10$ (5 for miniplexes and 5 for multiplexes). The optimal diffusion path under the planner's benchmark is close to the optimal path found by simulation shown in the main text in Figure 5 (top).

D.3 Distributors' surplus

Sections 7.1 and 7.2 focus on theaters' surplus. Does distributors' surplus increase under the two benchmarks? A priori, it is not clear: multi-homing is more prevalent and may be

[14] More precisely, it is a weighted average $\frac{\sum a_{it} \epsilon_{it}}{\sum a_{it}}$, because the number of screens converted differ across theaters of different sizes.

Figure A4: Diffusion paths under the three scenarios (full-solution approach)



(a) Digital screens

(b) Digital movies

costlier than 35mm film. This section shows that, given estimates of average distribution costs under film and digital, distributors also benefit under the two benchmarks.

Digital distribution affects only printing, shipping and storage costs of movie copies (PSS). Other costs incurred by distributors (e.g., advertising space purchase, advertising content creation, promotional events) remain constant.

Multi-homing may be costlier than film distribution because of economies of scale in PSS. When multi-homing, distributors split their production over two cost functions (film and digital), and in doing so, lose some of the scale economies. Denote by $C_f(Q)$ (resp. $C_d(Q)$) the total PSS costs of Q film copies (resp. Q digital copies); and $AC_f(Q)$ (resp. $AC_d(Q)$) the corresponding decreasing average cost curves. Digital distribution is more cost efficient, i.e., $C_d(Q) < C_f(Q)$. For a distributor releasing Q copies of a movie, multi-homing (with a fraction κ of digital copies and $(1-\kappa)$ of film copies) is costlier than film distribution if

$$C_f(Q) \leq C_d(\kappa Q) + C_f((1-\kappa)Q)$$

or equivalently

$$AC_f(Q) \leq \kappa AC_d(\kappa Q) + (1-\kappa) AC_f((1-\kappa)Q)$$

Whether the previous inequalities hold depends on the value of κ and the extents of scale economies in film and digital.

To compare distributors' costs under the market outcome and benchmarks, I collect data on average PSS costs per copy for film and digital, from the CNC. The CNC publishes a yearly report on the French movie distribution industry.[15] The reports contain a breakdown of total (industry-level) distribution costs into PSS costs, advertising space purchase, advertising content creation, and promotional costs. In particular, this information is available by number of copies (categorical variable). Table A7 shows average costs per copy for film and digital by total number of copies. Digital distribution is about 80% less costly than film distribution and the average cost per copy is decreasing in the number of copies. Shortcomings of the data are that it covers only French distributors, and for the "100-200 copies" category, costs per copy on film seem to increase.

Table A7: Average printing, shipping, storage cost per copy, by total number of copies (in thousands, 2010 €)

	Film		Digital	
Copies	LB	UB	LB	UB
<10 copies	1.14	2.28	0.50	1.00
10-50 copies	0.88	2.19	0.13	0.64
50-100 copies	0.79	1.58	0.13	0.26
100-200 copies	1.01	2.02	0.12	0.23
200-400 copies	0.59	1.17	0.09	0.18
>400 copies	0.57	1.13	0.08	0.15

Note: Estimate for French distributors. Imputed from aggregate PSS costs divided by # of movies per category. For Film, I use costs from 2007/2008*; for digital use costs from* 2015 *to* 2017.

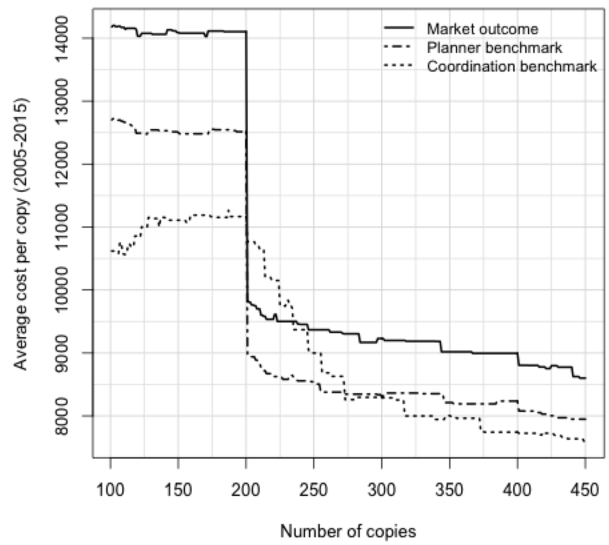
I use the data in Table A7 to compute the discounted sum of distribution costs for a distributor releasing one movie per period on Q copies

$$C(Q) = Q \times \sum_{t=0}^{\overline{T}} \beta^t \left\{ h_t \left(\frac{s_t}{S} AC_d(\frac{s_t}{S}Q) + (1 - \frac{s_t}{S}) AC_f((1 - \frac{s_t}{S})Q) \right) + (1 - h_t) AC_f(Q) \right\}$$

under the market outcome, planner's benchmark, and coordination benchmark. The first term inside the brackets corresponds to the average cost per copy when multi-homing

[15] The 2017 report, for instance, can be accessed at: `https://www.cnc.fr/cinema/etudes-et-rapports/etudes-prospectives/les-couts-de-distribution-des-films-dinitiative-francaise-en-2017_959089`.

Figure A5: Average cost per copy by number of copies



Note: Film LB and Digital UB used.

($\kappa = \frac{s_t}{S}$). The second term inside the brackets corresponds to the average cost per copy when distributing on film.

Figure A5 plots the average total cost $C(Q)/Q$ for various values of the number of copies Q. The results indicate that distribution costs are lower under the two benchmarks compared to the market outcome for a typical range of Q: the cost-reduction from digital more than compensates for the loss in scale due to multi-homing.

D.4 Robustness checks

This appendix investigates how the various assumptions required by the model impact the quantified effects.

Chain-level adoption. The assumption that theaters make their adoption decisions independently, even within chains, is violated if theater chains coordinate adoption decisions across theaters. Two incentives to do so are: (1) to benefit from lower per-unit adoption cost when placing large orders of projectors and (2) to tip the industry by significantly increasing the share of digital screens. To alleviate these concerns, the model controls for chain effects in the profits from operating (firm type $\boldsymbol{\tau}(i)$ include an indicator for the three largest theater chains). Discussions with chain managers (at Gaumont-Pathé) revealed that theaters were converted starting from the most profitable ones, which is consistent with theater-level profit maximization as in the model.

Discount factor. The main estimates use a discount factor $\beta = 0.95$ per 6 months, or 0.9025 annually. I examine how the results change with different discount factors ranging from 0.85 annually to 0.95 annually. Keeping the adoption costs fixed, a lower discount factor will be offset with higher estimates of per-period profits. As expected, I find that median profits per digital screen ($\pi_d(\tau(i))$) are $6,930$ euros with a 0.85 annual discount rate, $3,390$ euros with a 0.95 annual discount rate, compared to $4,986$ euros under the baseline 0.9025 discount rate.

The counterfactual results remain qualitatively similar albeit the actual magnitudes depend on the discount factor: for instance, the difference in (discounted) industry profits between the coordination benchmark and the market outcome ranges from 63% to 95% of industry profits over the sample period; whereas, the difference in (discounted) industry profits between the planner's benchmark and the market outcome ranges from 23% to 53% of industry profits over the sample period. With these alternative discount factors, network externalities explain between 31% and 55% of the surplus loss (as compared to 41% under the baseline discount factor of 0.9025).

Movie heterogeneity. The model assumes that profits per digital screen depend primarily on the share of movies released in digital, but not on the quality or expected box-office revenue. This assumption is mainly imposed due to the limited amount of data available on the distribution side. Nonetheless, if movies with high expected revenue are the first to be released digitally, this assumption might bias the estimates of the effect of h_t on profits. In particular, if movies are released in digital in decreasing order of expected box-office revenue, single-period profits per digital screen would be concave in h_t. The marginal return from an additional movie in digital would be decreasing in the number of digital movies.

One could, therefore, capture to some degree the heterogeneity in digital movie quality over time with a more flexible functional form for the dependence of profits on h_t. Imposing a quadratic form for h_t in the single-period profits does not alter the results quantitatively.

Calibration of distributors' reaction function. To verify the sensitivity of the results to the parameters governing the specification of h_t (Equation (26)), the planner's problem (25) is solved repeatedly under alternative choices of (η_s, η_h). I vary each parameter from 0.2 to 0.8 and calculate counterfactual industry profits under the planner's benchmark. Low values of η_s (respectively, η_h) correspond to more elastic supply of digital movies with respect to digital screens (respectively, digital releases in the U.S.).

Net industry profits are between 78.76 and 83.32 million euros (i.e., 4.2% lower to 1.3% higher than net industry profits under the baseline estimates of $(\widehat{\eta}_s, \widehat{\eta}_h) = (0.52, 0.21)$). The share of surplus loss explained by adoption externalities across downstream firms varies between 17% and 49%. The lower figure obtains when digital movies are little responsive to either $\frac{s_t}{S}$ or $h_{US,t}$. If the reduced-form estimates for the elasticity of s_t with respect to h_t (presented in Appendix B and found equal to 0.806) are of any indication, the lower figure is arguably a conservative lower bound for the surplus loss due to adoption externalities among theaters.

D.5 Policy remedies: subsidies to the first unit

This appendix investigates the impact of a subsidy targeting the first screen converted by each theater. Such subsidy could be used to build up initially a significant fraction of digital screens and digital movies, facilitating the subsequent conversion of late adopters. As in Section 7.4, I assume that the adoption cost incurred by theaters for the first screen converted is $(1-\gamma)p_t$ where $\gamma \in (0,1)$. After converting the first screen, theaters incur the full cost of conversion for their remaining screens.

The counterfactual diffusion paths are shown in Figures A6a and A6b, and corresponding industry profits and adoption are presented in Table A8. The subsidy policy succeeds at speeding adoption initially, with a concomitant increase in the share of movies available in digital. This results, however, in relatively small differences in the long-run: the time of 80% conversion is accelerated by only 6 months relative to the market outcome.

Table A8: Theaters' surplus under counterfactual subsidies (in millions, 2010 €)

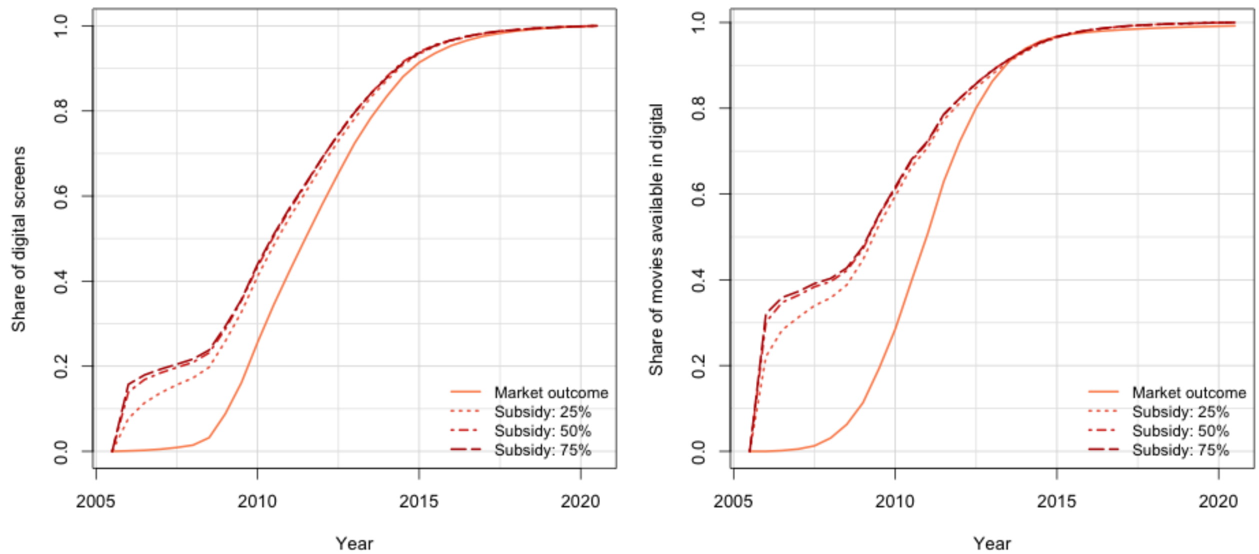
	Market	First-unit subsidies		
	outcome	$\gamma = 0.25$	$\gamma = 0.50$	$\gamma = 0.75$
Gross Industry Profits	167.54	191.14	196.88	197.84
Adoption costs	91.36	111.44	120.14	122.12
Net Industry Profits	76.18	79.7	76.74	75.71
Change in Net Profits	0	3.51	0.56	-0.47
Change in Net Profits (in %)	0	4.61	0.73	-0.62

Note: Profits from film are normalized to 0*: all profits are relative to the status-quo with no adoption.* γp_t *corresponds to the subsidy level incurred by the social planner and* $(1-\gamma)p_t$ *is incurred by the theater. "Adoption costs" include costs to theaters and the social planner.*

D.6 Algorithm to solve for counterfactual NOE

In this section, I introduce the algorithm used to solve for a counterfactual NOE under the various subsidy policies (see Section 7.4). Equilibrium computation of MPE can present

Figure A6: Diffusion paths under targeted subsidies (first unit)



(a) Digital screens (b) Digital movies

substantial hurdles when the game involves many firms and a large state space, as is the case in this particular application. The NOE restriction alleviates these issues by drastically reducing the dimension of the state space. Moreover, given that a firm ignores its impact on other players and on the aggregate states, each firms' best response can be computed separately (within a given iteration).

I solve for a NOE in the space of conditional choice probabilities using policy iterations. In a first step, I calculate ex-ante value functions associated with a given vector of choice probabilities (that may not be optimal). Second, choice-specific value functions are computed and used to update the choice probabilities. Iteration continues until convergence of the values and choice probabilities up to a prescribed tolerance level (in practice 10^{-5}). Details are presented in Algorithm (1). For each subsidy policy, I initialize the algorithm at different starting values for the choice probabilities to verify that the algorithm converges to a unique equilibrium.

References

Efron, B. and R. J. Tibshirani (1993). *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability 57. Boca Raton, Florida, USA: Chapman & Hall/CRC (cit. on p. 10).

Algorithm 1 Nonstationary Oblivious Equilibrium solver

1: Initialize the adoption policy rule $\mathbf{P}^{(0)}$ as

$$\{P_t^{(0)}(a_{it}|\mathbf{x}_{it})\} \text{ for all } t = 1, \dots, \overline{T} \text{ and } \mathbf{x}_{it} = (\boldsymbol{\tau}(i), s_{it})$$

2: $\Delta := \epsilon + 1$

3: **while** $\Delta > \epsilon$ **do** (Iteration $(k+1)$)

4: Compute the expected aggregate share of digital movies given $\mathbf{P}^{(k)}$

$$\overline{h}_t^{(k)} = \mathbb{E}_{\mathbf{x}_t}[h_t|\mathbf{x}_1]$$

by simulating L_1 industry paths from $t = 1$ to $\overline{T}$ and averaging over simulations.

5: For each firm i, update its ex-ante value function at all states given the process $\{\overline{h}_t^{(k)}\}$

$$\widetilde{\mathbf{V}_\mathbf{i}}^{(k+1)} = (\mathbf{I} - \beta \cdot \mathbf{F}^{(k)})^{-1} \left\{ \sum_a \mathbf{P}^{(k)}(a)\big(\boldsymbol{\Pi}(\overline{h}_t^{(k)}) - a(\mathbf{p}(\gamma) + \mathbf{e}(a))\big) \right\}$$

where matrices are defined as in Equation (16) and $\mathbf{p}(\gamma)$ is the adoption cost accounting for the subsidy.

6: For each firm i, update its choice-specific value function at all states and actions

$$\widetilde{W}_t^{(k+1)}(a|\mathbf{x}_{it}) = \beta \widetilde{V}_{\boldsymbol{\tau}(i),t+1}^{(k+1)}(s_{i,t+1})$$

where $\mathbf{x}_{it} = (\boldsymbol{\tau}(i), s_{it})$ and $s_{i,t+1} = s_{it} + a$.

7: Compute differences in choice-specific value functions $\Delta\widetilde{W}_t^{(k+1)}(a+1, a|\mathbf{x}_{it})$

8: Update the conditional choice probabilities

$$P_t^{(k+1)}(a_{it}^{no} \le a|\mathbf{x}_{it}) = 1 - \Phi\left(\Delta\widetilde{W}_t^{(k+1)}(a+1, a|\mathbf{x}_{it}) - p_t\right)$$

(note action $a+1$ is not played with positive probability if

$$\Delta\widetilde{W}_t^{(k+1)}(a+1, a|\mathbf{x}_{it}) = \Delta\widetilde{W}_t^{(k+1)}(a+2, a+1|\mathbf{x}_{it}))$$

9: $\Delta := ||\mathbf{P}^{(k+1)} - \mathbf{P}^{(k)}||$

10: **end while**

Gowrisankaran, G. and J. Stavins (2004). “Network Externalities and Technology Adoption: Lessons from Electronic Payments”. *The RAND Journal of Economics* 35.2, p. 260. arXiv: `arXiv:1011.1669v3` (cit. on p. 5).

Griliches, Z. and J. A. Hausman (1986). “Errors in variables in panel data”. *Journal of Econometrics* 31.1, pp. 93–118 (cit. on p. 6).

Stock, J. H. and M. Yogo (2005). “Testing for weak instruments in linear IV regression”. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg.* Vol. 2001, pp. 80–108 (cit. on p. 6).