Sunk costs and the dynamics of creative industries

Gerben Bakker

© Gerben Bakker

December 2012
Sunk costs and the dynamics of creative industries

Gerben Bakker

ABSTRACT

This chapter examines the long-run evolution of modern entertainment industries such as the film and music industries. It investigates ways to conceptualise and quantify the subsequent waves of creative destruction, and investigates specifically how sunk costs affect the evolution of the industry through its interaction with variety, market integration, product differentiation and price discrimination, and how old entertainment formats almost never became extinct. It finds that within this framework, four economic tendencies shaped the entertainment industries evolution: first, endogenous sunk costs often led to a competitive escalation of production expenditures, which we call ‘quality races’, which increased industrial concentration. Second, the fact that marginal revenues largely equalled marginal profits led to extreme vertical integration through ownership or revenue-sharing contracts, as well as to an oversupply of variety and a dual market structure with high-concept blockbuster products and low-budget niche products. Third, entertainment’s public good characteristics led to substantial income inequality among creative inputs and business models optimising exclusion possibilities in the value chain. Finally, the project-based character of entertainment production implied large intra- and inter-industry agglomeration benefits and often led to geographical concentration. Dynamic product differentiation allowed various old formats to survive the waves of creative destruction, albeit in much smaller incarnations.

Key words: sunk costs, market structure, dynamic efficiency, price discrimination, history, creative industries, motion pictures, videogames, music, live entertainment, dynamic product differentiation, quality races, industrialisation, horizontal and vertical product differentiation, agglomeration economies, total factor productivity, revealed comparative advantage, variety, vertical integration, business models.

A revised version of this paper will be forthcoming in Candace Jones, Mark Lorenzen and Jonathan Sapsed eds., The Oxford Handbook of the Creative and Cultural Industries (Oxford University Press, 2013).

1 Previous versions of this paper benefitted from comments received at the sports and business history seminar at the Institute of Historical Research in London, the workshop on media history of the German Society for Business History in Augsburg, the workshop on the history of the music industry at St Andrews University, the research seminar in music studies at Kingston University, and the workshop on the management and evolution of the creative industries at the Freeman Centre, University of Sussex. I thank Jonathan Sapsed, Candace Jones, Mark Lorenzen, Juan Mateos-García and Christopher Colvin for additional comments. The research for this chapter was partially funded by the Economic and Social Research Council (U.K.) and the Advanced Institute of Management Research under the ESRC/AIM Ghoshal Fellowship Scheme, grant number RES-331-25-3012.

2 Dr Gerben Bakker, Departments of Economic History and Accounting, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, Tel.: + 44 - (0) 20 - 7955 7047, Fax: + 44 - (0) 20 - 7955 7730, email: g.bakker@lse.ac.uk. Website: http://www2.lse.ac.uk/researchAndExpertise/Experts/g.bakker@lse.ac.uk
I. Introduction

Over the past two centuries several forms of mass entertainment have become industrialised. Motion pictures, for example, were adopted to automate, standardise and make tradable live performances, the gramophone to industrialise live music, and videogame consoles to industrialise all types of board and card games. Yet new and old forms of entertainment, big-budget productions and shoe-string projects, homogenous and diverse markets often survived in an uneasy co-existence. Music recording did not kill live performances, Hollywood blockbusters not the art-house film, and best-seller markets not the enormous diversity of idiosyncratic books.

The main aim of this chapter is to explain long-run change in the entertainment industries, to develop an overarching conceptual framework to understand their dynamics since about 1850, especially those of sub-industries such as live entertainment, motion pictures, music and videogames. This leads to three research questions:

What are the defining economic characteristics of entertainment industries?

How have they shaped historical processes such as concentration, vertical integration and agglomeration in film and music over the past two centuries?

How did they shape the ‘product-space’, the coexistence of old and new forms, big and small producers, homogenous and diverse markets?

This investigation is worthwhile because it yields insight in the process of long-run change in the creative industries and might thus help practitioners and policy makers to understand and anticipate the shape of things to come.


Works that address all three aspects are sparse. They include the monumental total histories of the global film industries by the French film historians Jean Mitry (1967-1980) and Georges Sadoul (1948-1954; 1972), the work by Kristin Thompson (1985) on the international expansion of the early Hollywood industry, the work by Michael Chanan (1980, 1996) on how cinema emerged from a range of popular entertainments such as music hall, and Baumol and Bowen (1966), who examine the evolution of performing arts’ costs in the Unites States.

What follows will qualitatively analyse structural change in the entertainment industry since c. 1850. First we will look at ways to conceptualise and quantify the subsequent waves of creative destruction, and then we will investigate how sunk costs affected industry evolution through its interaction with variety, market integration, product differentiation and price discrimination. A subsequent section will discuss four economic characteristics that drove the evolution of the entertainment industry: how sunk costs led to quality races, marginal revenues
equalling marginal profits to vertical integration, quasi-public good characteristics to business models centred around points-of-exclusion, and the industry’s project-based character to agglomeration. A final section concludes.

II. The process of creative destruction

II.A. Industrialisation and the shifting production possibility frontier

During the late nineteenth century, demand for entertainment increased sharply, driven by a growth in leisure time, discretionary income, and population. In addition, the demand was focused spatially by urbanisation and the growth of rapid local transport networks (Bakker 2008).

Stimulated by this booming demand, entrepreneurs adopted new technologies such as the phonograph, motion pictures, radio and television to automate and standardise entertainment and make it tradable, thus industrialising it. Automation resulted in low and flat marginal costs of offering additional viewings, compared to high fixed and sunk costs of making the prototype (the film negative, the manuscript), enabling massive scale economies. Thus the printing press industrialised handwritten manuscripts and verbal communication, cinema theatrical entertainment, the gramophone musical performances, and videogames the playing of games. The tradability of performances themselves (performers being shown in many venues simultaneously) led to the integration of regional and national entertainment markets. The resulting product standardisation homogenised the consumer experience.

This industrialisation was brought about by several waves of creative destruction that swept the industry. Schumpeter (1942) originally introduced the concept of creative destruction as the way in which a new way of doing business competes away the old ways and products. Creative destruction is intricately linked to dynamic efficiency: an industry can be statically efficient if price equals marginal costs (allocative efficiency), and if all firms use the most efficient technology available (productive efficiency). Yet a statically efficient industry need not be dynamically efficient: it may not develop new products, new processes, new markets, new sources of supply or new organisational forms, and herein lies according to Schumpeter the defining characteristic of capitalism: that it is never in equilibrium, but always changing from one equilibrium to the next, like a person walking. Once one equilibrium is being approached, a new better one has become possible, and entrepreneurs jump on the new opportunities offered.

Thus instead of eating away at the margins, creative destruction threatens the very survival of entire firms and industries. In many industries one can see waves of creative destruction in which a new product sweeps away old products, but in entertainment old ‘products’ often survived in new, differentiated guises.

A way to further conceptualise the process is the production possibility frontier (PPF) originally introduced by one of Schumpeter’s contemporaries, Gottfried von Haberler (1930).

Dynamic efficiency is a process in which the PPF is being pushed outwards continuously. A PPF can show the trade-off of a society in consuming entertainment and all other products (diagram 1). Inside the frontier C-F, say from A to B, we can get more of both goods by using
more efficient existing technology. At the frontier, starting from C to D, initially we get lots entertainment by giving up a little of all else, but eventually we get less and less additional entertainment for an additional amount of other goods we give up.

Diagram 1. *Hypothetical production possibility frontier for entertainment and all other goods.*

Point D is productively efficient as it is on the PPF, but not allocatively efficient, as it is not on the highest utility curve possible, unlike point E, which is both. The second PPF C-G shows the effect of dynamic efficiency, brought about by innovations such as motion pictures and the gramophone.

Over the last one and a half centuries a series of innovations shifted the PPF outwards, followed by a series of jumps in which the industry moved closer to the PPF, but as it tended to reach it, there was not there anymore and the PPF had shifted outwards again.

In live entertainment, for example, several exogenous innovations ranging from urbanisation and steel-frame buildings to the railroads, shifted the PPF outwards, and entrepreneurs jumped on it by building high capacity fixed theatres and founding central booking offices directing theatre groups by telegraphs over railroads along the most efficient routes (diagram 2). Other entrepreneurs built entire movable theatres on boats visiting cities along a river.
Diagram 2. Qualitative analysis of successive shifts in the production possibility frontier for various forms of entertainment.

### Theatre circuits

<table>
<thead>
<tr>
<th>Exogenous technological changes</th>
<th>First-order outward shift of PPF related to entertainment</th>
<th>Second-order outward shift of PPF related to entertainment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steel-frame buildings (early 19th c.)</td>
<td>Steel-frame theatres (early 19th c.)</td>
<td>Theatre circuits Novel additional circuits (opera, big and small-time vaudeville, burlesque)</td>
</tr>
<tr>
<td>Railways (c. 1840)</td>
<td>Show boats</td>
<td></td>
</tr>
<tr>
<td>Telegraph (1837)</td>
<td>Many more theatres</td>
<td></td>
</tr>
<tr>
<td>Urbanisation (since 18th c.)</td>
<td>Stock/repertory theatre (same actors, different plays)</td>
<td></td>
</tr>
<tr>
<td>Increasing income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switch to measured time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steamships</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deregulation of theatre</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Motion pictures

<table>
<thead>
<tr>
<th>Exogenous technological changes</th>
<th>First-order outward shift of PPF related to entertainment</th>
<th>Second-order outward shift of PPF related to entertainment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Projection (1654)</td>
<td>Motion picture cameras (1891)</td>
<td>Cinemas (c. 1905)</td>
</tr>
<tr>
<td>Persistence of vision gadgets (1826)</td>
<td>Motion picture film stock (1891)</td>
<td></td>
</tr>
<tr>
<td>Photography (1839)</td>
<td>Motion picture projection (1895)</td>
<td></td>
</tr>
<tr>
<td>Celluloid (1868)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positives and negatives (1887)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Celluloid sheets (1888)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roll film (1888)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-sensitivity emulsion (1888)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Music

<table>
<thead>
<tr>
<th>Exogenous technological changes</th>
<th>First-order outward shift of PPF related to entertainment</th>
<th>Second-order outward shift of PPF related to entertainment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spiral spring (15th c.)</td>
<td>Phonograph (1873)</td>
<td>Phonograph and gramophone used to record and sell music (1890s --)</td>
</tr>
<tr>
<td>Cylindrical home music systems (15th c.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small cylinder music boxes with cam, not bells (1796)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barrel organs (c. 1800)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barrel organ books (1892)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recording of sound (1850s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engraved interchangeable discs for music boxes (1870s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recording and playback of sound (1873)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Videogames

<table>
<thead>
<tr>
<th>Exogenous technological changes</th>
<th>First-order outward shift of PPF related to entertainment</th>
<th>Second-order outward shift of PPF related to entertainment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Board games (since times immemorial)</td>
<td>Arcade game machines (1970s)</td>
<td>Home game consoles (various generations since 1970s)</td>
</tr>
<tr>
<td>Mechanical arcade machines (e.g. pinball)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microprocessors (1957)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer simulation software (1960s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home computers (1977)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Likewise, for motion pictures seven preconditional technological changes shifted the PPF outwards, ranging from projection in 1654 to sensitive emulsions in 1888, and made possible the invention of cameras, film stock and projectors. These again shifted the frontier outwards, leading to the innovation of fixed cinemas by the mid-1900s, which again shifted the PPF outwards, leading to feature films by the mid-1910s, and so on.

Likewise for music, innovations such as barrel organs, disc-based music boxes and recorded sound shifted the PPF outwards, eventually leading to the phonograph. For videogames, exogenous technological changes such as board games, pinball machines, computer software and microprocessors shifted the PPF outwards and enabled the introduction of arcade game machines by the late 1960s, shifting the PPF further outwards and later enabling the innovation of home videogames.

II. B. Measuring dynamic efficiency

A quantitative proxy of creative destruction is the growth in total factor productivity (TFP), or how much faster outputs grow than inputs, reflecting a more efficient use of production factors. Although few estimates exist, available data for spectator entertainment in the United States, Britain and France for the 1900-1938 period show a TFP-increase ranging from 1.1 percent to 5.4 percent per annum, suggesting that the effect of dynamic efficiency was substantial (table 1). In the US, ten times as many spectator-hours were produced per unit of labour and capital than in 1900, and in Britain (already highly productive in 1900) about one and a half times as many. The U.S. data can be disaggregated into a 7.2 percent per annum TFP-growth for film and 0.5 percent for live (Bakker 2012), showing that, even while under threat from film, live entertainment in 1938 produced 20 percent more output per input than in 1900.3 Slightly less reliable data for France, show TFP-growth of 4 percent per annum, with 1938 TFP four times as large as in 1900. These number suggest that dynamic efficiency in

---

3 In theory, all TFP-growth in live entertainment could be explained by scale economies through increased urbanisation (Bakker 2012: 1055-1057).
entertainment was larger than in almost any other industry, and studies of the phonograph, radio and television would probably further reinforce this finding.

Table 1. Contributions to output growth in spectator entertainment, 1900-1938, per cent per annum.

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>UK</th>
<th>France</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output growth</td>
<td>8.6</td>
<td>3.2</td>
<td>5.9</td>
</tr>
<tr>
<td>Contributions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital</td>
<td>1.2</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>Labour quantity</td>
<td>2.0</td>
<td>1.2</td>
<td>1.1</td>
</tr>
<tr>
<td>All above inputs</td>
<td>3.2</td>
<td>2.1</td>
<td>1.9</td>
</tr>
<tr>
<td>TFP</td>
<td>5.4</td>
<td>1.1</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Notes: the capital share is set at 0.25; labour includes both the growth in labour quantity and labour quality.
Source: Bakker (2008); because of unique sources, the US growth rates could be estimated more precisely and are based on Bakker (2012); they may not be fully comparable with the French and British rates.

A second way to proxy an industry’s dynamic efficiency is its revealed comparative advantage (RCA): its share in a country’s exports over the world-wide industry’s share in global exports. The higher this ratio is above unity, the higher the RCA of the particular industry. Work by Nicholas Crafts (1989) shows the RCA for tradable entertainment products in a category he called ‘Book and film’ that comprised most media products, including books, magazines and records (table 2).

Table 2. Revealed Comparative Advantage rankings of tradable entertainment sector, 1899-1950.

<table>
<thead>
<tr>
<th>Rank of tradable entertainment products</th>
<th>1899</th>
<th>1913</th>
<th>1929</th>
<th>1937</th>
<th>1950</th>
</tr>
</thead>
<tbody>
<tr>
<td>Britain</td>
<td>14</td>
<td>13</td>
<td>6</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>US</td>
<td>13</td>
<td>10</td>
<td>6</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Germany</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>France</td>
<td>3</td>
<td>4</td>
<td>10</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Belgium</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Sweden</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Italy</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Switzerland</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Canada</td>
<td>4</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Japan</td>
<td>10</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>India</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Notes: tradable entertainment products are labelled ‘Book and Film’ by Crafts, and comprise the following: ‘books, periodicals and all printed matter, agendas, notebooks and boxed stationery, pens, pencils, toys, games and sports goods, gramophones, musical instruments, cameras, optical instruments, films and photographic paper, paintings and works of art.
The first number, 14, for example, shows that ‘Book and Film’ had an RCA that ranked 14 out of the 16 British manufacturing sectors in 1899.

The data shows that the ranks varied substantially among countries. In Germany ‘Book and Film’ was one of the top-sectors, while in Britain and the US in 1899 it was one of the lowest performing sectors. Between 1899 and 1950, however, Book and Film’s RCA in Britain
and the US increased considerably, as it also did in Belgium and Japan. The US had the highest gain in RCA-rank, followed by Belgium and then Britain. In Germany, France, Sweden and Canada the RCA of tradable entertainment declined substantially. Yet, in absolute terms, Germany RCA in tradable entertainment remained the highest of all countries throughout the period. France, India and Japan held a surprisingly high RCA in tradable entertainment, in many years higher than the US, though by 1950 their rank has become the same as that of US Book and Film. The high degree of aggregation and the inclusion of non-entertainment products makes more precise conclusions difficult.

Another proxy of dynamic efficiency is the number of new products. This is not always easily measurable, although existing studies suggest a sharp increase in variety, based on the number of music copyrights registered, films and TV-programmes released, and books published. We will now turn to the interaction between variety, sunk costs and industrialisation.

III. Sunk costs and the industrialisation of entertainment

Sunk costs are incurred once and cannot be recovered when exiting a business. Fixed costs are incurred periodically and can usually be avoided. Endogenous sunk costs such as advertising or R&D can be chosen by the firm, and do not have a minimal exogenous level. Costs such as film production, scouting for and recording of musicians (A&R) or developing videogames—are endogenous sunk costs because they are incurred once, cannot be recovered other than by selling tickets and their level can be chosen (Bakker 2005).

This section will investigate how sunk costs and variety were affected by the industrialisation process, the role of product differentiation, and the relation between sunk costs and price discrimination.

III.A. Sunk costs and variety

The existence of fixed (and partially sunk) costs are the reason we do not have a world of infinite variety. Each consumer is unique in her tastes, and the resulting difference in the valuation of goods and services makes profitable exchange possible. All consumers’ utility could be maximised by crafting entertainment products for each individual, but the existence of fixed and sunk costs necessitates the making of entertainment products that are somewhat different from each consumer’s optimum (diagram 3).

---

4 See, for example, figure 6, below.
A second, similar trade-off is that between variety and efficiency. In the extreme, zero variety (high sunk costs per prototype and low average costs per viewing) maximises efficiency, while endless variety (low sunk and high average costs) minimises it. One could envisage, for example, only 25 films being produced world-wide annually, satisfying all demand—enough to see one new film a fortnight, plus a menu of older films. If total consumer expenditure remained constant, a massive increase in TFP would take place, as the same output is now produced with fewer inputs. In extremis just one film would be made, that consumers would see over and over again, ostensibly increasing TFP to orgasmic heights.

Obviously, consumers enjoy variety and would regard the availability of just one film as a massive decrease in output, even if they could watch it many times and in many different ways. Each consumer prefers entertainment products fully crafted to their own tastes. The set of different media products consumed—books read, films watched, music listened to—is unique for every consumer. It is well nigh impossible among the six billion inhabitants of the earth to find two souls that have consumed exactly and strictly the same collection of entertainment. Economic theory (Krugman 1979) and empirical research suggest consumers strongly prefer variety. Broda and Weinstein (2006), for example, find that the benefits of the increasing import variety between 1972 and 2001 added 2.6 percent to U.S. GDP.

The degree of variety that actually materialises in free exchange is per definition suboptimal as it does not tailor each product to a specific consumer. It depends on the fixed and sunk costs needed to develop the product, but also on an intertemporal information asymmetry: distributors do not fully know ex-ante which product will sell and therefore need to distribute more than one product to make sure capacity (cinema seats, DVD shelf space, TV time slots) is used optimally. Even if consumers would not love variety, the studio would need to produce a portfolio of firms to discover the best film consumers like. The mirror image are consumers who do not fully know ex-ante how satisfying they will find a specific film, and this also makes
tailoring entertainment perfectly to each individual consumer impossible, even without fixed and sunk costs. In other words, producers who don’t know what will sell make products for consumers who don’t know what they want.

Only if consumers were completely indifferent to variety would they be happy with just one film being made, as the pleasure of seeing the biggest budget highest pleasure giving movie would outweigh the reduction of variety to zero. Although this situation seems far-fetched for commercial entertainment products, with experiences such as religious prayers, texts and rituals, or with centrally prescribed school textbooks in centrally planned economies, users consumed a single prototype many times—and that actually constituted the value of the experience. Both cases suggest that variety, for better or for worse, is intricately linked with modern capitalism, as free exchange is based on everybody having different preferences.

**III.B. The effect of industrialisation on variety**

Since 1850 an industrialisation process has shifted household production of many different products (songs, plays, poems, home-made toys) with limited fixed costs, and quality closely tailored to individual preferences towards a more centralised production of far fewer products with high fixed costs and quality less tailored-made (diagram 4).

At the economy level, two forces were at work. First, new technology such as centrally booked, railway-routed theatre circuits, cinema and the phonograph shifted the production possibility frontier outwards. Sunk costs per prototype became far higher, but average costs per viewing became far lower. At any level of variety, more quantity could be produced. Second, in a self-reinforcing process, this stimulated market integration which then in itself further stimulated the growth of sunk costs. As a result, at the aggregate level, there was probably less variety as a whole in a country, as an enormous archipelago of small islands—the individual households—producing a large part of their own entertainment, gave way to national markets with concentrated production.


<table>
<thead>
<tr>
<th></th>
<th>Household production</th>
<th>Market production</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed and sunk costs</strong></td>
<td>Limited costs per prototype</td>
<td>High costs per prototype</td>
</tr>
<tr>
<td><strong>Quality</strong></td>
<td>Limited quality</td>
<td>Higher quality</td>
</tr>
<tr>
<td><strong>Market size</strong></td>
<td>Small and closed</td>
<td>Large and open</td>
</tr>
<tr>
<td><strong>Market integration</strong></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td><strong>Differentiation</strong></td>
<td>Perfectly differentiated to household’s taste</td>
<td>Imperfectly differentiated to household tastes</td>
</tr>
<tr>
<td><strong>Total amount of variety</strong></td>
<td>Infinitely many varieties produced in total</td>
<td>A limited amount of variety produced in total</td>
</tr>
<tr>
<td>‘Effective’ amount of variety</td>
<td>Limited set available for each household given its</td>
<td>Large set available for each household (given that fixed</td>
</tr>
</tbody>
</table>

Outside entertainment, an extreme case is standardisation. Consumers all want to drive on the same side of the road whether it be right or left, for example.
<table>
<thead>
<tr>
<th></th>
<th>fixed costs</th>
<th>costs are ‘shared’</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model</strong></td>
<td>Traditional</td>
<td>Capitalist / exchange</td>
</tr>
<tr>
<td><strong>Efficiency</strong></td>
<td>Lower</td>
<td>Higher</td>
</tr>
<tr>
<td><strong>Over all</strong></td>
<td>Perfect differentiation, limited available variety per household, traditional production—no market or price/quality signal</td>
<td>Homogenisation, large available variety per household, market exchange price/quality signal</td>
</tr>
</tbody>
</table>

At the household level, the extreme product differentiation, in which each household made its own products, disappeared. But in return each household had now access to a far larger variety of entertainment products, from which it still could choose its own truly and strictly unique combination, and of which it now could afford to consume a far larger quantity. The fixed and sunk costs per household probably declined, and rising wages increased the opportunity costs of household production. In nineteenth century Britain, for example, the industrialised areas had the largest provision of commercial entertainment (Sanderson 1984).

Technological change since the 1950s has made the exogenous sunk costs of producing a record, film or book far lower. This should lead to more variety, but was mitigated by a competitive escalation of endogenous sunk costs, leading to a dual market structure, with both mass-marketed blockbuster products and a second market with an almost infinite variety of differentiated products (see section IV, below).

The evolution of exogenous sunk costs may have followed an inverted U-shape, with the difference that the pre-industrial low costs were for isolated submarkets, an infinite archipelago of largely autarkic household islands in a national economy, while the present-day post-industrial low cost products cater for an archipelago of specific ‘isolated’ market segments (diagram 5).
III.C. Dynamic product differentiation

Despite the waves of creative destruction that have swept the entertainment industry since the early nineteenth century, old formats often did not disappear but instead reinvented themselves. Today theatre plays, musicals and operas are still performed, despite the availability of radio, television and internet, because entrepreneurs adopted product differentiation to survive creative destruction.

Product differentiation can take place across many dimensions. Horizontally differentiated products are aimed at customers who will not easily switch to another product even if its quality increases, such as a pop music fan who never listens to operas no matter how much their quality is increased. Vertically differentiated products are clearly viewed as superior over others along a dominant quality dimension.

Informationally differentiated products vary by the extent to which consumers know about them, and can be influenced by advertising. As search costs increase with variety (Stigler 1961), for entertainment products consumer search costs are high. Entrepreneurs attempt to differentiate their products informationally by escalating promotional expenditure. Marketing costs as share of Hollywood’s production expenditure, for example, are exceptionally high.
<table>
<thead>
<tr>
<th>Within product category</th>
<th>Within same product</th>
<th>A particular film: cinema — TV etc</th>
<th>Partially coincides with price discrimination</th>
<th>Mode of viewing; point in time; (not the product itself)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between products</td>
<td>Dual market structure; Hollywood films vs. art-house cinema</td>
<td>Less to do with price discrimination</td>
<td>Content of product</td>
<td></td>
</tr>
<tr>
<td>Between product categories</td>
<td>Opera-first-class theatre – vaudeville — motion pictures</td>
<td>Coincides with different prices but also different costs and different products</td>
<td>Style/format of product (meta)</td>
<td></td>
</tr>
</tbody>
</table>

Within the same product such as a particular mainstream feature film, the differentiation criterion is often the means and time of viewing, varying from first-run cinema to television, each at a consecutively lower price (diagram 6). Besides mainstream, vertically differentiated Hollywood films, there is a second, fragmented market for independent art-house films, horizontally differentiated from each other and collectively from the Hollywood films. Between product categories differentiation is usually more based on style and format, both horizontally and vertically. Some categories such as opera also contain a large club good character when they become a way to meet like-minded persons.

Industries can be ranked informally by relative horizontal-vertical differentiation and low-high sunk costs (Diagram 7). Vertical differentiation implies that consumers will turn their backs on lower quality products when higher quality products appear, so that producers are disproportionately rewarded for a quality increase, which under some conditions leads to quality races.\(^6\) Horizontal differentiation implies that consumers value variety and that few products will obtain a large market share. In practice, in most creative industries products are both vertical and horizontally differentiated in varying degrees. The more vertically differentiated, the higher endogenous sunk costs can be, as they can be amortised over a larger volume. Products such as Hollywood films and videogames are far more vertically differentiated than music, television and art-house films.

This phenomenon of the survival of old media in a different form surfaced time and again in the history of the entertainment industry. We will call it here dynamic product differentiation.\(^7\) Although old formats did not disappear, they also did not stay the same, but differentiated themselves in the competition with the new medium. Often they could not offer as low prices as the new medium and therefore would focus on a more wealthy audience segment (the part of their previous audience with the highest willingness to pay), such as many forms of live entertainment did after cinema. Sometimes they could not offer the premium quality of the new segment and therefore would focus on an audience with a lower willingness to pay or in a different segment, such as free TV after cable TV and radio after television. Radio refocused, at sharply reduced advertising rates, on segmented markets and on situations where television was not feasible, such as while driving, working or shopping.

That old media did not simply disappear was partially because of their unique assets, such as theatre locations and production know-how. Some of these assets were moved into new media. Many theatres converted into cinemas, creative live professionals moved to film production, musicians began to make recordings and radio game shows transferred to television. Distribution capacity had always been scarce, so old delivery systems could still be used to serve specific segments. Each new medium increased capacity and allowed more market segmentation, the widest channel focusing on the lowest common denominator.

Old media output often shrunk in relative terms. Between 1900 and 1938, for example, U.S. live entertainment attendance declined with 1.3 percent a year on average, to sixty percent of its original quantity, but in relative terms it had declined far steeper, to only a few percent of all spectator entertainment (Bakker 2012).

An iconic example of dynamic product differentiation is live entertainment. In the course of the nineteenth century many new forms developed which still existed in the Boston market in 1909 (figure 1). The shape of the demand curve, with high price elasticity at high prices and low prices, and rather moderate elasticity in between may to some extent explain the time lags in dynamic product differentiation. In a discovery process, entrepreneurs had to find out what was the demand for new innovative forms of entertainment. The low elasticity in the middle part of the demand curve may have slowed down attempts to provide far lower-priced forms of

---

\(^6\) See the section on sunk costs, above.

\(^7\) On product differentiation in motion pictures since 1945, see Sedgwick (2002).
entertainment, as many entrepreneurs may have guessed that demand would not increase enough to make it worthwhile. Thus it may have taken many historical accidents and innovations such as cinema before entrepreneurs ‘discovered’ the responsiveness of demand at low prices. This process is not dissimilar to the ‘path dependency’ hypothesis that more efficient paths are not taken because the first few steps are unattractive and hide the large pay-offs later on. The path taken resulted from the interplay between the entrepreneurial discovery process, cost-decreasing technology, and changes in size and geographic density of demand.\(^8\)

After the coming of the feature film and of sound, the live entertainment that survived did so by differentiating itself. It was either highly artistic and heavily subsidised, or highly commercial and high-value added, such as avant-garde theatre or Broadway musicals and plays. What lay in between was competed away by cinema. Likewise, after television took over the focus on the lowest common denominator, motion pictures focused on a younger, more affluent audience, offering more event movies at higher prices.

Figure 1. *Ticket price versus cumulative ticket-selling capacity for entertainment venues in Boston in 1909* ($ and number of tickets).

Source: Boston Committee (1909) as reported in Jowett (1974); see also Bakker (2008).

---

\(^8\) On entrepreneurial discovery see Kirzner (1985).
The above can be illustrated on a classic Hotelling line (diagram 8). If we assume a continuum from low- to high-value entertainment, in the 1890s theatre would have been somewhere in the middle to maximise revenue. After film arrived, theatre moved upmarket, and film would move slightly upmarket as well. Eventually, new technologies occupied more and more spots on the differentiation continuum, with cinema and theatre moving further upwards.

Thus dynamic product differentiation limited the effect of creative destruction. Old format were often not completely destroyed by new technology, but instead reinvented itself.

**III.D. Sunk costs and price discrimination**

If a producer-distributor charges a single (monopoly) price for an entertainment product, say a film, then revenue may not be enough to cover fixed and sunk production costs, while the total surplus, which includes the consumer surplus, would be large enough. This is shown in panel A of diagram 9, where price is below average costs. In other words, the producer cannot take into account the consumer surplus while setting the budget, and consumers exist that are willing to pay a price above marginal costs, but remain unserved. Through price discrimination, however, a producer-distributor can charge different prices to different consumers and so transform consumer surplus into revenue, enabling higher fixed and sunk costs (panel B in diagram 9). With a linear demand curve, monopoly pricing and zero marginal costs, at the extreme perfect price discrimination can double revenue and thus sunk outlays.
Diagram 9. A stylised example of sunk costs, revenues and average costs with and without price discrimination.

Note: the line shows the demand curves, the dotted line marginal costs, the bold line average costs and the surface revenue. This necessarily is a simplified example. For a more detailed discussion see, for example, Carlton and Perloff (2003: 219-224) and Romano (1991).

In the nineteenth century, for example, consumers paid a high price to see a new play in a metropolitan theatre. Later lower-priced versions were offered by travelling groups and resident stock companies, and the smallest communities might be visited by miniature puppet theatres (Bakker 2008).

Nineteenth-century theatre managers were well aware of the need for price discrimination. The financing of the rising fixed and sunk costs spent in the railroad, utilities and the theatre construction booms, was only possible through large-scale price discrimination. In 1849, Jules Dupuit (1849: 16, 25), a pioneer of modern utility analysis, mentioned the theatre and book publishing as prime examples of price discrimination:

A single price for tickets will not fill a theatre and might often yield only modest receipts. Hence the manager would suffer a pecuniary loss and the public a loss of utility. Divisions on the floor of the theatre and differential pricing of the tickets nearly always raise receipts as well as the number of viewers. This would be easy enough to understand if these divisions merely separated the seats where one can see and hear well from those where one cannot. But observation of how a theatre floor is mostly split up in practice shows that this is one of the least considerations in pricing the tickets; the theatre managers know how to adapt their prices to all the whims of the spectators, those that go to see and those that go to be seen, and those that may go for some other reason. They are made to pay according to sacrifice they are prepared to make to satisfy their whims, and not according to the show they enjoy.

Likewise, before 1950 motion pictures were released in metropolitan theatres at premium prices, then in the main city centre theatres across the country, then in regular theatres, and so
on, at ever lower prices, until after a year it ended up for a few cents in ramshackle sixth-run
neighbourhood cinemas (Sedgwick 2001). After 1950, cinemas reorganised into fewer
windows, while television, video, pay-TV, cable and free-to-air TV took over the other windows.
Without such price discrimination, the film could not be made at the given quality. Most
consumers who see Hollywood films on television can only do so because others were willing to
pay premium prices to see them sooner in a higher quality medium such as cinema.

Several studies have tried to quantify price discrimination’s effect. Leslie (2004), for
example, studying 199 performances of one Broadway play, found that price discrimination
increased profits by five to seven percent, and Huntington (1993) that non-price-discriminating
theatres could increase their revenue by 24 percent if they did. These might be lower bounds, as
endogeneity—roughly speaking a self-reinforcing upward spiral—is ignored: if price
discrimination’s increasing revenues result in higher sunk costs that increase quality, this may
increase consumers’ willingness to pay and thus allow an increase in over-all prices, as well as
perhaps an even sharper price discrimination, allowing higher sunk costs, increasing willingness
to pay, allowing a higher over-all price, and so forth.

One of price discrimination’s main benefits is to allow in the consumers outside the
market: those that otherwise would not be in it (Courty 2000). When a production could not be
made without price discrimination, this might have meant all its consumers would have been
outside the market.9

The creative inputs and the owners of the distribution delivery system often have different
interests. In pop-concerts, for example, performers have an incentive to price discriminate,
maximising revenue rather than attendance. Promoters, however, make most sales from
concessions, parking and the Ticket Master service charge split, and thus have an incentive to
maximise attendance and not ticket prices.10 Likewise, cinemas often make most profits from
concessions and distributors from tickets. Only perfect price discrimination, clearing the market
and filling every seat serves both interests. These misaligned incentives could explain why actual
ticket prices and degree of price discrimination are far below levels that maximise performer
revenue.11

The importance of price discrimination can be shown quantitatively with an example of
an average Hollywood movie in the 1990s (figure 2). Ninety percent of revenue came from
consumers with a relatively high willingness to pay, ranging from $5.50 and above to $0.70 and
above, and only ten percent from those with a very low willingness to pay ($0.15 and above). If
all consumers saw the movie at the last (probably perfectly competitive) price, total gross
revenue would be only $32m instead of the actual $200m and the film could only be made at a
far lower budget. Even if we assume a single monopoly price is charged of say $0.30, revenue
would only be $64m, still necessitating a far lower production budget. It is thus very clear that
consumers that saw the movie at $0.15 could only do so because others saw it at far higher

9 On ticket pricing and price discrimination see also Rosen and Rosenfeld (1997), and Courty (2003a,
2003b). See also Orbach and Einav (2007) on pricing in cinemas who observe that this used to be very
differentiated, but is now more uniform. The latter is undoubtedly due to pay, cable and free-to-air
television as well as videocassettes and DVDs taking over the windows that were previously formed by
lower-level cinemas. See also Ekelund (1970).
10 Billboard, 7 March 2009; see also, for example, Krueger (2005).
prices. The audience share of cinema and video for example, was the obverse of its revenue share.™

If we assume a production budget of $40 million and a producer’s share of revenue of one-third, it becomes clear that in none of the windows the producer’s price was higher than average costs (figure 3), but that when all windows are summed, the effect of price discrimination has lifted the producer’s average revenue per viewing substantially above average costs.

If we assume zero marginal cost, a linear demand curve and that for all windows a monopoly price was charged, given the producer’s copyright-monopoly, consumer surplus will be exactly half the revenue, suggesting that although consumers would have been willing to pay $300m for the movie in the aggregate, they needed to pay only $200m, so despite the price discrimination still enjoyed a consumer surplus of about $100m. Deriving consumer surplus empirically from Figure 2 yields broadly similar results (Table 3). Cinema, with only five percent of the audience, accounted for 29 percent of sales and 21 percent of the total consumer surplus, while free-to-air television, with 68 percent of audience, accounted for only eleven percent of sales revenue, and 29 percent of total consumer surplus. The latter group would not have been able to watch the film without the former group with a higher willingness to pay.™ Without price discrimination, the price per viewing would have been about three dollars, while under price discrimination, prices range from as low as fifteen cents to as high as $5.50. Price discrimination, in this case, increased sales by $91m and the consumer surplus by $86m (Table 3). Without price discrimination transforming about two-thirds of consumer surplus into sales revenue, the movie could not be made at its existing quality. This fact shapes production in many other creative industries and suggests that price discrimination might be welfare increasing in creative industries.

™ Including sell-through DVDs in Figure 2 between ‘cinema’ and ‘video’ would not basically change the revenue model, as the price per viewing of these was likely lower than that of cinema, given that they were generally watched by more than one person and multiple times.

™ We abstract here from the situation that a portion of viewers will be in both groups.
Figure 2. **Price versus cumulative number of viewings for a typical Hollywood film, 1990s ($ and number of viewings).**

Source: calculated from data in Dale (1997).

Figure 3. **Producer’s share of ticket price and average costs versus cumulative number of viewings for a typical Hollywood film with a $40m production budget, 1990s.**
IV. Four economic tendencies that drove the evolution of the entertainment industries

Four economic tendencies have determined the entertainment industry’s evolution since c. 1850: the importance of endogenous sunk costs, the fact that marginal revenues largely equalled marginal profits, entertainment’s quasi-public good character, and its project-based nature. We discuss their characteristics in turn, identify their dynamic implications and assess how these expressed themselves historically.

IV.A. Sunk costs and quality races

The dynamic implication of endogenous sunk costs is what we call a quality race, a market phase in which some firms escalate their sunk outlays in order to obtain a larger market share. While in other industries the lower bound concentration may fall to zero as the market size increases, because there is ‘room’ for more companies to enter the market, in some endogenous sunk costs industries concentration is bounded from below when market size grows to infinity, and does not converge to zero (Sutton 1991, 1998). Market growth in such industries raises profits for any given quality level, making room for new entrants, but also stimulating firms to raise their R&D investments—such as film production or videogame development outlays—to improve their quality level: while the marginal cost of an increase in R&D-spending is unchanged, the marginal benefit from it is now higher, leading to higher fixed

---

Table 3. World-wide prices, audience/sales potential and consumer surplus for a typical Hollywood motion picture, 1990s.

<table>
<thead>
<tr>
<th>Price ($1000)</th>
<th>Audience (%)</th>
<th>Sales ($1000)</th>
<th>Consumer Surplus ($1000)</th>
<th>Total Surplus ($1000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cinema</td>
<td>5.50</td>
<td>10,500</td>
<td>57,750</td>
<td>28,875</td>
</tr>
<tr>
<td>Video</td>
<td>3.00</td>
<td>34,000</td>
<td>102,000</td>
<td>42,500</td>
</tr>
<tr>
<td>Pay television</td>
<td>0.70</td>
<td>24,000</td>
<td>16,800</td>
<td>27,600</td>
</tr>
<tr>
<td>Free-to-air television</td>
<td>0.15</td>
<td>147,000</td>
<td>22,050</td>
<td>40,425</td>
</tr>
<tr>
<td>Total</td>
<td>0.92</td>
<td>215,500</td>
<td>198,600</td>
<td>139,400</td>
</tr>
<tr>
<td>Cinema and video</td>
<td>3.59</td>
<td>44,500</td>
<td>159,750</td>
<td>71,375</td>
</tr>
<tr>
<td>All else</td>
<td>0.23</td>
<td>171,000</td>
<td>38,850</td>
<td>68,025</td>
</tr>
<tr>
<td>No price discrimination</td>
<td>2.99</td>
<td>35,833</td>
<td>107,051</td>
<td>53,526</td>
</tr>
<tr>
<td>Effect price discrimination</td>
<td>0.15 - 5.50</td>
<td>179,667</td>
<td>91,549</td>
<td>85,874</td>
</tr>
<tr>
<td>Perfect price discrimination</td>
<td>(1.62)</td>
<td>215,500</td>
<td>349,025</td>
<td>0</td>
</tr>
</tbody>
</table>

Notes: the consumer surplus shown is based on rough estimates and simplified assumptions and is for illustrative purposes only. For cinema, the consumer surplus has been calculated assuming monopoly pricing for the segment and a linear demand curve, making it equal half the sales revenues; for all other categories it is the surface of the white triangle spaces under the connecting lines in Figure 2. The ‘no price discrimination’ case is based on a rough estimate speculatively assuming Figure 2 shows a demand schedule, that marginal costs are zero and that monopoly pricing applies. The table shows the average of 2 different cases: one with a linear demand curve following the first line segment in figure 2, the other following the second line segment. An average price of $3.30 under no discrimination equalises the increased sales revenue with the lost consumer surplus. Perfect price discrimination sums total actual sales and consumer surplus, which together constitutes sales under perfect discrimination; the average price is shown. The consumer surplus under perfect discrimination is the sum of the existing surplus and the estimated surplus below the lowest price point of $0.15. The latter has been roughly estimated by speculatively assuming the demand curve intercepts the zero price line after 147 million additional viewings beyond the free-to-air television price point. Total surplus, producer and consumer surplus is the sum of sales revenue and consumer surplus under the assumption of zero marginal costs.

Source: Calculated from data from Dale (1997).

14 For a case study that applies these four tendencies to the evolution of the British entertainment industry, see Bakker (2013).
sunk costs and limiting the number of firms as the market tends to infinity. Which of the two effects has the upper hand depends on the distribution of the willingness-to-pay and shape of R&D-costs associated to quality improvements (Bakker 2005).  

As the market for films grew, for example, two opposite effects could happen: more companies could enter or existing companies could increase production budgets. Which effect dominated depended on how easy it was to increase perceived quality through higher production outlays and to what extent higher-quality films stole sales from lower-quality films.

It turns out that, in the motion picture industry, the second effect dominated. With each new product category, such as colour film, newsreels or cartoons, a quality race ensued, but also ended soon when further expenditures did not yield additional sales anymore. Eventually, in the U.S. during the mid-1910s, an escalation phase began which did not have a natural endpoint: producers continuously increased expenditure on feature films, yielding ever more ticket sales and replacing shorts as the dominant format (figures 4 and 5). European firms could hardly participate, as the European home market was disintegrating, and war made venture capital almost unavailable, as all capital allocations were targeted at the war effort (Bakker 2000).

Through a process of entrepreneurial discovery, partially by accident partially by design, the industry had stumbled upon a format where the sunk costs could be amplified almost continuously because of the disproportionate return to higher quality. Cinemas, for example, preferred a higher-quality film, even if lower-quality ones had bargain prices, because they needed to earn back cinema fixed costs and because a lower-quality film’s opportunity costs were huge.

A 700-seat cinema, for example, with a production capacity of 39,200 spectator-hours a week, weekly fixed costs of $500, and an average admission price of $0.05 per spectator-hour, needed a film selling at least 10,000 spectator-hours, and would not be prepared to pay for that (marginal) film, because it only recouped fixed costs. Films thus needed a minimum selling capacity to cover cinema fixed costs. Producers/distributors could only price down low-budget films that passed the threshold level where expected revenues equalled costs. With a lower expected selling capacity, these films could not be sold at any price (Bakker 2004).

Opportunity costs reinforced this even further. If the hypothetical cinema obtained a high-capacity film for a weekly rental of $1,200, which sold all 39,200 spectator-hours, the cinema made a profit of $260 ($0.05 price X 39,200 spectator-hours = $1960 revenue - $1200 film rental - $500 fixed costs= $260). If a film with half the budget and, we assume, half the selling capacity, rented for half the price ($600), the cinema-owner would lose $120 ($0.05 price x 19,600 spectator-hours = $980 revenue - $600 film rental - $500 fixed costs = -$120). Thus, the cinema owner would want to pay no more than $220 for the lower budget film, given that the high budget film is available ($0.05 X 19,600 = $980 revenue - $220 film rental - $500 = $260 profit). If the high-capacity film were not available, the cinema would only want to pay $480

15 The term R&D is used here for sunk outlays on developing products in the creative industries rather than the more general term ‘innovation’, because we are interested here in particular identifiable products with a particular amount of money spent on their development, rather than in innovation as a whole.
16 This is not dissimilar to the quality thresholds, capabilities and minimum quality/cost ratios discussed in Sutton (2005).
17 The relation between production costs and selling capacity is assumed to be linear. Moderate decreasing returns would yield the same effect, as would increasing returns. Probably on average, increasing returns
for the low-capacity film at most ($980 revenue - $500 fixed costs) in order to break even. A film’s ticket-selling capacity could not always be perfectly predicted, but as a film’s release progressed and as it was exported to more countries, its appeal became better known (Bakker 2004).

In this example a film that costs twice as much earns more than five times as much in rentals ($1,200/$220) in case of competition, and more than two and a half times as much ($1200/$480) without competition. In both cases the high-quality film gets the full rental, while the rentals for a low-quality film depend on competitive conditions. Producers thus had a sharp incentive to increase costs, making the continuous budget inflation of the 1910s and 1920s understandable (figure 4).

The fixed-costs threshold became lower with each new distribution delivery system. It was far lower for television and even lower for internet distribution, meaning that more lower quality products could be distributed. Opportunity costs, however, did not necessarily show such a decline.

Figure 4. **Total annual production outlays for various U.S. film producers, 1913-1927: semi-logarithmic scale.**

**Notes:** all series, except ‘Census’ are scales on the left-hand axis. The DeMille series concerns the outlays of just one producer of Paramount and is less comparable to the other series, which are for entire studios or the industry. Census = total industry production outlays.

**Sources:** see Bakker (2005: 325-326).

were followed by constant returns, and finally by decreasing returns to the last dollars spent on already high-budget films.
Quality races such as the rise of the feature film in the mid-1910s have happened at various points in the entertainment industry’s development. In the 1970s, for example, a new quality race started when the ailing Hollywood studios focused on high-concept blockbuster movies that were heavily marketed and advertised on television, starting with *Jaws* in 1975. Helped by other factors as well, the studios rapidly reached their former pre-eminence.

In the music industry since the 1950s, a sharp escalation of firms’ expenditure on Artists and Repertoire (A&R) took off, with a few big multinationals focusing on heavily marketing a few popular acts while at the same time also achieving horizontal product differentiation through a wide range of different labels and acts (Bakker 2011b). Real revenue per music copyright increased almost four-fold between 1955 and 1970, a phenomenal average annual real growth of about nine percent (figure 6).

In videogames probably a quality race took place from the mid-1990s onwards, when companies started to sink more and more and more outlays in the development of videogames, with a few large companies with big distribution organisations becoming prominent (Bakker 2010).
IV.B. Marginal revenues are marginal profits

The second tendency follows from the first. When entertainment products such as film tickets, music recordings or video games are sold, all development costs have already been incurred and marginal revenue therefore largely equals marginal (gross) profit. This holds, for example, for the film producer who sells the film to distributors—each sale is marginal profit—, for the distributor who, once it bought the film, sells it on to cinema –owners, and cinema-owners, who, once having bought the film, sell it to consumers, with every additional chair filled being marginal profits. Similar points can be made for many other entertainment products such as music, books, or videogames.

Low marginal costs should lead to very low prices, and thus an inability to incur large fixed and sunk costs, were it not that the producer-distributor is granted a monopoly through copyright law, and can use price discrimination to transfer some of the consumer surplus into revenue, as we saw above.

In an unintegrated value-chain in which products are sold outright, the producer’s incentive to increase quality is limited by the fact that s/he will not get any of the marginal revenues that their product generated for the distributor and for the retailer. The classic solution has been vertical integration or the use of revenue-sharing contracts. This construction aligns the
interests of the participants in the various links of the value chain, and so makes sure that marginal retail revenues reach the producer, and this in turn gives the producer the incentive to make marginal improvements in product quality that will lead to additional sales, and thus more profits, also for the producer.

In the early motion picture industry, for example, films were sold. This meant that the cinema-owner kept their marginal gross profits themselves, and the distributors did the same. This made the profit incentive for the film producer small: a good film would lead to some more sales at higher prices to more distributors, but the distributors and cinemas would get all the marginal profits (Bakker 2003). This circumstance drove vertical integration in the industry, with producers often integrating with distributors, the latter sometimes with cinemas, and films increasingly being rented for a percentage of revenue rather than sold outright. Distributors also started to advance money for film production costs. In France, Pathé already started this integration in the 1900s, in the U.S. the Motion Picture Patents Company and General Film Company tried to integrate production and distribution and monopolise the film business, and from about 1912 the U.S. independents vertically integrated (Bakker 2008). Integration and profit-sharing across the value chain shaped the development of many media industries.

Besides stimulating vertical integration, standard industrial organisation theory suggests that constant or falling marginal costs lead to an excessive number of firms. Together with low exogenous costs this had led to an enormous variety of media products and a dual market structure. In motion pictures, for example, the competitive escalation of endogenous sunk costs in quality races led to a vertically differentiated, highly concentrated market for big-budget Hollywood films, while low exogenous sunk costs and low marginal costs led to excessive entry and a second, separate, fragmented market with an almost infinite variety of films, each having a very small market share. The latter market appears characterised by monopolistic competition: price is above the competitive level, but firms do not make economic profits (Chamberlin 1933; Robinson 1933).

### IV.C. Quasi-public good characteristics

Public goods such as national defence are nondiminishable (or nonrivalrous), and non-excludable: one person’s consumption of it does not reduce the quantity available to others, and no persons can be excluded from benefiting from it. Since 1850, entertainment has become ever more nondiminishable, but remained excludable. In a theatre, for example, until capacity was filled, one additional person would not diminish the entertainment available to others, while any additional consumer of bread did diminish the quantity available to others.

An almost continuous series of technological improvements made entertainment even less diminishable. Sharp rises in theatre capacity through steel-frame construction, and later cinema and recorded music massively increased the audience a performer could reach. Not only radio, television and internet would increase this even further, but also ‘older’ technology such as jet travel and large-scale stadiums. Between 1981 and 2003, for example, the top-1 percent

---

18 Mezias and Mezias (2000) borrow the term resource partitioning from biology to study phenomena not unlike both dynamic product differentiation and dual market structure discussed here.
of live rock music performers received between 30 and 50% of the total ticket revenue (Krueger 2005:14).

Yet entertainment forever remained a quasi-public good: nondiminishable but excludable, unlike a pure public good as national defence. Throughout the history of the entertainment industry, entrepreneurs therefore developed business models that kept price above marginal costs by making them own the point where consumers could get excluded and extract all rents there. Five major business models were important.

First, the entrepreneur could exclude visitors from a theatre and so charge a price. They could physically exclude people and use their right to set prices to price-discriminate among visitors. Second, copyright allowed entrepreneurs a monopoly on their product.

Third, an important group of entrepreneurs formed the stars themselves, who could withhold their services and thus exclude people, depending on the extent of their talent. Industrialisation made their performances tradable, and so increased stars’ earning capacity. The result was a very unequal distribution of income among the top performers, although exact statistics lack to test whether this was actually more unequal that the distribution among live performers before film (Bakker 2001; Rosen 1981). The sharply unequal income distribution has remained a property of many media industries ever since. This made it also increasingly expensive for film makers, especially new entrants, to recruit top talent. Entrepreneurs adopted various business models to mitigate this value capture by the stars. The Hollywood studios introduced asymmetric long-term (‘seven year’) contracts that allowed them to keep stars’ pay limited when they became successful, European firms often used revenue-sharing contracts, which limited cash-costs, as pay-outs were only made when cash was coming in. From the late 1940s, seven-year contracts became unenforceable and a mix of advance payments and revenue-sharing is now widely used.

Fourth, production of diminishable goods such as merchandising that could be sold at a premium were used to generate revenue. Fifth, collusion could be a means of exclusion. The interwar Hollywood studios, for example, formed a cartel and jointly monopolised resources, preventing any firm to bid up star pay. Since the collusion and 7-year contracts ended in the late 1940s, star pay has risen sharply.

An example of long-run changes in the public-good nature is the music industry, where album sales have diminished and become less excludable because of copying technologies, and the new business model depends on the remaining excludable parts, such as live concerts, diminishable goods such as T-shirts or deluxe CD-editions, or renting out artists to advertise other diminishable goods (Krueger 2005).

VI.D. The project-based nature of entertainment

---

19 In the nineteenth century the global news agencies used the same technique to exclude non-payers. See Bakker (2011a).
A fourth economic characteristic is the project-based nature of entertainment production. Each entertainment product—such as a book, a film, a piece of music—is unique. The making of each product is a separate project with unforeseen contingencies, for which different creative, technical and commercial talent is assembled. Although every industry contains activities that are project-based, the proportion of it in the entertainment industry is extreme (diagram 10).

Diagram 10. Hypothetical representation of the boundaries of the project-based segments of the creative industries and various other industries.

The dynamic implication of this is agglomeration. Each project benefits from other projects being organised in the neighbourhood, much in the same way as Alfred Marshall described agglomeration benefits: within industries, co-location creates a thick market for specialised inputs, external economies of scale in production as specialised firms lower costs by spreading them over more buyers, and knowledge spill-overs, as talent meets informally, circulates between firms and exchanges best practices and ideas.20

In addition, Jane Jacobs (1969) emphasized the importance of inter-industry externalities. Co-location of different entertainment industries would yield similar benefits. In London, for example, music, film, radio, television and media financing companies may all benefit from being close to each other and therefore being able to better organise projects. Thus the agglomeration benefit spread at various different levels (diagram 11), and that makes them so important in the entertainment industry.

20 See, for example, chapter 6 in Krugman and Obstfeld (2003).
Diagram 11. Stylised representation of the interlinked intra- and inter-industry webs of project-based entertainment industries

An example is early film production which agglomerated in France in Paris and Nice, in Britain around London, and in the U.S. initially in New Jersey and later in Florida and California (Bakker 2008: 258-261). Until the mid-1920s, both intra- and inter-industry externalities were important. Production located close to theatrical districts such as London, Paris and New York. In the 1920s, however, the shift of the U.S. industry to Hollywood was driven solely by intra-industry externalities. It was far from New York, where the studios’ corporate headquarters and distribution operations were. Later television and music production would also come to Hollywood, adding inter-industry co-location benefits. In the late twentieth century, the large and fast-growing Indian film industry became also highly concentrated geographically in Mumbai (Lorenzen and Mudambi 2012).

In recorded music, agglomeration was not always as strong as in motion pictures. In Europe, capital cities were important locations. For most styles, location was part of the brand image, and the music industry therefore agglomerated in many more locations, but with intra-industry agglomeration benefits for specific styles, such as Nashville, Chicago or New York.

The project-based characteristic contributed to this differentiation of locations. Also, because an industrial district lowered set-up costs, it could lead to ‘excess’ variety.

V. Conclusion

21 For a detailed analysis of agglomeration benefits in the creative industries, and for present-day evidence of their importance, see Bakker (2010); for a historical analysis of Hollywood see Christopherson and Storper (1987).

22 For a comparison of agglomeration in the Indian film and software industries in Mumbai and Bangalore, see Lorenzen and Mudambi (2012).
We started this essay with the question how we can explain long-run change in the entertainment industry. We discussed the process of industrialisation, creative destruction and shifting production possibility frontiers, and how to measure its impact through, for example, total factor productivity, revealed comparative advantage or variety. We discussed the relation between sunk costs, variety, industrialisation, dynamic product differentiation and price discrimination and the trade-off between variety, quality, quantity and efficiency. Finally, we identified four main economic tendencies (diagram 12): the importance of endogenous sunk costs that led to quality races, zero marginal costs leading to vertical integration, quasi-public good characteristics leading to business models focused on points-of-exclusion, and, finally, entertainment’s project-based character leading to agglomeration.

Key insights included the need for price discrimination, vertical integration and business models protecting points-of-exclusion in the value chain to enable larger sunk outlays; and agglomeration to minimise costs and maximise effectiveness of those outlays. Old media survived by differentiating themselves, focusing on distinct, often narrower, audiences.

Diagram 12. Major economic characteristics of the film industry and their dynamic and historical implications.

<table>
<thead>
<tr>
<th>Economic characteristic</th>
<th>Dynamic implication</th>
<th>Historical expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunk costs</td>
<td>Quality race</td>
<td>Motion pictures 1910s and 1970s/80s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Music industry 1950/60s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Videogames 2000s</td>
</tr>
<tr>
<td>Marginal revenue = marginal profits</td>
<td>Vertical integration</td>
<td>Motion pictures: Europe 1900s; US 1910s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Music: 1900s; 1960s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Videogames: 2000s</td>
</tr>
<tr>
<td>Quasi-public good character (non-diminishable but</td>
<td>Income inequality (emergence star system and</td>
<td>Theatre 19th c. — Motion pictures 1910s — 1950s —</td>
</tr>
<tr>
<td>excludable)</td>
<td>superstars) Business models</td>
<td>Videogames: to come (?).</td>
</tr>
<tr>
<td>Project-based character</td>
<td>Agglomeration</td>
<td>U.S. 1905-1920s: Jacobs and Marshall externalities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U.S. 1925-: Marshall externalities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Europe: 1900s: Jacobs and Marshall externalities</td>
</tr>
</tbody>
</table>

Intriguingly, old and new forms, big and small producers, homogenous vertically differentiated markets and diverse, horizontally differentiated markets all coexisted. Low and flat marginal costs and exogenous sunk costs, the simultaneity of horizontal and vertical product differentiation and agglomeration benefits, led to excess variety, organised in a dual market structure, with one part offering fewer high-concept block-buster products, and the other part offering a flood of low-budget horizontally differentiated products.

Clearly historical events interacted with economic forces, such as the quality races in film in the 1910s and the 1970s, in music in the 1950s and 1960s, motion pictures agglomeration in Hollywood in the late 1920s and music’s decentralised agglomeration in different locations since 1955.
Implications for business include the need to either focus on horizontal or vertical differentiation; to find the best ways to price discriminate and to vertically integrate to increase sunk outlays, to develop business models focused on dominating the point-of-exclusion, to be alert to the signs that a quality race is about to begin, to locate within an industrial district to lower costs, to look out for shifts in the production possibility frontier, and, finally, to be aware of dual market structures and choose one of the two markets but not both. Sound and simple as these implications may sound, the path of the entertainment industry over the last two centuries is littered with the wrecks of firms that did not heed them.\footnote{See, for example, Bakker (2007).}


Bakker, Gerben, Entertainment Industrialised: The Emergence of the International Film Industry, 1890-1940 (Cambridge University Press, 2008).


Hofstede, Bart, *Dutch Film Abroad: identity and organisation of the Dutch film industry since 1945* (Delft, Eburon, 2000).


2009

WP114  War and Wealth: Economic Opportunity Before and After the Civil War, 1850-1870
        Taylor Jaworski

        Albrecht Ritschl and Tobias Straumann

WP116  The Impact of School Provision on Pupil Attendance: Evidence From the Early 20th Century
        Mary MacKinnon and Chris Minns

WP117  Why Easter Island Collapsed: An Answer for an Enduring Question
        Barzin Pakandam

WP118  Rules and Reality: Quantifying the Practice of Apprenticeship in Early Modern Europe
        Chris Minns and Patrick Wallis

WP119  Time and Productivity Growth in Services: How Motion Pictures Industrialized Entertainment
        Gerben Bakker

WP120  The Pattern of Trade in Seventeenth-Century Mughal India: Towards An Economic Explanation
        Jagjeet Lally

WP121  Bairoch Revisited. Tariff Structure and Growth in the Late 19th Century
        Antonio Tena-Junguito

WP122  Evolution of Living Standards and Human Capital in China in 18-20th Centuries: Evidences from Real Wage and Anthropometrics
        Joerg Baten, Debin Ma, Stephen Morgan and Qing Wang

WP123  Wages, Prices, and Living Standards in China, 1738-1925: in Comparison with Europe, Japan, and India
        Robert C. Allen, Jean-Pascal Bassino, Debin Ma, Christine Moll-Murata, Jan Luiten van Zanden
WP124 Law and Economic Change in Traditional China: A Comparative Perspective
_Debin Ma_

WP125 Leaving Home and Entering Service: The Age of Apprenticeship in Early Modern London
_Patrick Wallis, Cliff Webb and Chris Minns_

WP126 After the Great Debasement, 1544-51: Did Gresham’s Law Apply?
_Ling-Fan Li_

WP127 Did Globalization Aid Industrial Development in Colonial India? A Study of Knowledge Transfer in the Iron Industry
_Tirthankar Roy_

WP128 The Education and Training of Gentry Sons in Early-Modern England
_Patrick Wallis and Cliff Webb_

WP129 Does Trade Explain Europe’s Rise? Geography, Market Size and Economic Development
_Roman Studer_

WP130 Depression Econometrics: A FAVAR Model of Monetary Policy During the Great Depression
_Pooyan Amir Ahmadi and Albrecht Ritschl_

WP131 The Economic Legacies of the ‘Thin White Line’: Indirect Rule and the Comparative Development of Sub-Saharan Africa
_Peter Richens_

WP132 Money, States and Empire: Financial Integration Cycles and Institutional Change in Central Europe, 1400-1520
_David Chilosi and Oliver Volckart_

WP133 Regional Market Integration in Italy During the Unification (1832-1882)
_Anna Missiaia_

2010

WP134 Total Factor Productivity for the Royal Navy from Victory at Texal (1653) to Triumph at Trafalgar (1805)
_Patrick Karl O’Brien FBA and Xavier Duran_
WP135 From Sickness to Death: The Financial Viability of the English Friendly Societies and Coming of the Old Age Pensions Act, 1875-1908
Nicholas Broten

WP136 Pirates, Polities and Companies: Global Politics on the Konkan Littoral, c. 1690-1756
Derek L. Elliott

WP137 Were British Railway Companies Well-Managed in the Early Twentieth Century?
Nicholas Crafts, Timothy Leunig and Abay Mulatu

WP138 Merchant Networks, the Baltic and the Expansion of European Long-Distance Trade: Re-evaluating the Role of Voluntary Organisations
Esther Sahle

WP139 The Amazing Synchronicity of the Global Development (the 1300s-1450s). An Institutional Approach to the Globalization of the Late Middle Ages
Lucy Badalian and Victor Krivorotov

WP140 Good or Bad Money? Debasement, Society and the State in the Late Middle Ages
David Chilosi and Oliver Volckart

WP141 Becoming a London Goldsmith in the Seventeenth Century: Social Capital and Mobility of Apprentices and Masters of the Guild
Raphaelle Schwarzberg

WP142 Rethinking the Origins of British India: State Formation and Military-Fiscal Undertakings in an Eighteenth Century World Region
Tirthankar Roy

WP143 Exotic Drugs and English Medicine: England’s Drug Trade, c.1550-c.1800
Patrick Wallis

WP144 Books or Bullion? Printing, Mining and Financial Integration in Central Europe from the 1460s
David Chilosi and Oliver Volckart

WP145 ‘Deep’ Integration of 19th Century Grain Markets: Coordination and Standardisation in a Global Value Chain
Aashish Velkar
WP146 The Utility of a Common Coinage: Currency Unions and the Integration of Money Markets in Late Medieval Central Europe
*Lars Boerner and Oliver Volckart*

WP147 The Cost of Living in London, 1740-1837
*Ralph Turvey*

WP148 Labour Market Dynamics in Canada, 1891-1911: A First Look From New Census Samples
*Kris Inwood, Mary MacKinnon and Chris Minns*

WP149 Economic Effects of Vertical Disintegration: The American Motion Picture Industry, 1945 to 1955
*Gregory Mead Silver*

**2011**

WP150 The Contributions of Warfare with Revolutionary and Napoleonic France to the Consolidation and Progress of the British Industrial Revolution
*Patrick Karl O’Brien*

WP151 From a “Normal Recession” to the “Great Depression”: Finding the Turning Point in Chicago Bank Portfolios, 1923-1933
*Natacha Postel-Vinay*

*Debin Ma*

WP153 The Finances of the East India Company in India, c.1766-1859
*John F. Richards*

WP154 Labour, law and training in early modern London: apprenticeship and the city’s institutions
*Patrick Wallis*

WP155 Why did (pre-industrial) firms train? Premiums and apprenticeship contracts in 18th century England
*Chris Minns, Patrick Wallis*

WP156 Merchantlist institutions for a first but precocious industrial revolution; The Bank of England, the Treasury and the money supply, 1694-1797
*Patrick O’Brien*
WP157  Hand Looms, Power Looms, and Changing Production Organizations: The Case of the Kiryu Weaving District in the Early 20th Century Japan  
**Tomoko Hashino, Keijuro Otsuka**

WP158  From Divergence to Convergence: Re-evaluating the History Behind China’s Economic Boom  
**Loren Brandt, Debin Ma, Thomas G. Rawski**

WP159  Money and Monetary System in China in the 19th-20th Century: An Overview  
**Debin Ma**

WP160  Contesting the Indigenous Development of "Chinese Double-entry Bookkeeping" and its Significance in China’s Economic Institutions and Business Organization before c.1850  
**Keith Hoskin, Richard Macve**

WP161  Steel, Style and Status: The Economics of the Cantilever Chair, 1929-1936  
**Tobias Vogelgsang**

WP162  The Seven Mechanisms for Achieving Sovereign Debt Sustainability  
**Garrick Hileman**

WP163  Reparations, Deficits, and Debt Default: The Great Depression in Germany  
**Albrecht Ritschl**

WP164  Bounded Leviathan: or why North and Weingast are only right on the right half  
**Alejandra Irigoin, Regina Grafe**

WP165  Monetary sovereignty during the classical gold standard era: the Ottoman Empire and Europe, 1880-1913  
**Coşkun Tunçer**

WP166  Going Beyond Social Savings. How would the British economy have evolved in the absence of the railways? A case study of Brunner Mond 1882-1914.  
**Edward Longinotti**

WP167  Public Finance in Britain and China in the long Eighteenth Century  
**Peer Vries**
WP168  The rise of the patent department: A case study of Westinghouse Electric and Manufacturing Company
*Shigehiro Nishimura*

WP169  The Maghribi industrialists contract enforcement in the Moroccan industry, 1956-82
*Romain Ferrali*

WP170  Adopting the rights-based model: music multinationals and local music industries since 1945
*Gerben Bakker*

WP171  The Eighth Wonder of the World: How might access for vehicles have prevented the economic failure of the Thames Tunnel 1843-1865?
*Rio Lydon*