Can Emerging Markets Tilt Global Product Design? Impacts of Chinese Colorism on Hollywood Castings

Hermosilla, Manuel and Gutierrez-Navratil, Fernanda and Prieto-Rodriguez, Juan

Johns Hopkins University, Universidad Publica de Navarra, Universidad de Oviedo

30 September 2017

Can Emerging Markets Tilt Global Product Design? Impacts of Chinese Colorism on Hollywood Castings*

MANUEL HERMOSILLA†
Johns Hopkins University

FERNANDA GUTIÉRREZ-NAVRATIL
Universidad Pública de Navarra

JUAN PRIETO-RODRÍGUEZ
Universidad de Oviedo

October 2017

Abstract

In various cultural and behavioral respects, emerging market consumers differ significantly from their counterparts of developed markets. They may thus derive consumption utility from different aspects of product meaning and functionality. Based on this premise, we investigate whether the economic rise of emerging markets may have begun to impact the typical "one-size-fits-all" design of many international product categories. Focusing on Hollywood films, and exploiting a recent relaxation of China's foreign film importation policy, we provide evidence suggesting that these impacts may exist and be non-negligible. In particular, we show that the Chinese society's aesthetic preference for lighter skin can be linked to the more frequent casting of pale-skinned stars in films targeting the Chinese market. Implications for the design of international products are drawn.

*We thank Colleen Stuart, Ricard Gil, Laura Kornish, Yi Qian, Junzhao Ma, Liye Ma, Charles Weinberg, Aynne Kokas, Manuel Bagues, and Stanley Rosen for their comments and suggestions.
†Corresponding author. Email: mh@jhu.edu.

1 Introduction

Fast-paced, sustained GDP growth and increasing integration into world trade have made emerging markets important new players in the global economy. These economies have rapidly enlarged their middle classes and significantly reduced poverty (Sala-i Martin, 2006; Sudhir et al., 2015; Chandy and Narasimhan, 2015), endowing significant shares of their populations with purchasing power and access to retail commerce. As a result, this process has articulated an influx of "newly endowed" consumers into the global demand for products and services.

In this article we investigate the importance of this phenomenon for the design of international products. These "newly endowed" consumers are known to significantly differ from their counterparts of developed markets in cultural and behavioral respects (Henrich et al., 2010; Sheth, 2011). For example, whereas the former emphasize embeddedness and hierarchy, the latter emphasize autonomy and egalitarianism (Burgess and Steenkamp, 2006). Differences like this suggest that consumers from emerging markets may place utility value on different aspects of product meaning and functionality. Their activation into the global demand for products and services may therefore bring along a shift in aggregate consumer preferences within the total addressable market, prompting so an impact on (or accommodation of) product design.

Although these design impacts may unfold through additional product variety, they are likely to be more transparently observed on products which cannot be meaningfully customized to each international market. That is, on products which, due to the nature of their development process, are "condemned" to a largely standardized design. These products are the focus of this study. For ease of reference, we call them "global products." We associate them to categories such as feature films, TV series, music, video games, literature and other commercial art forms, as well as alternative kinds of written and audiovisual media (e.g., blogs, podcasts). We observe that, despite the fact that some of these may undergo minor forms of international customization, large fixed re-design costs prevent deep customization. For example, although feature films may be dubbed and slightly edited when released in a foreign market, the more important determinants of their appeal to audiences (aesthetics, storyline, central themes, starring actors) are deeply enmeshed in, and inextricably tied to their core structures, and thus largely uncustomizable. We ask: can the newly acquired importance of emerging markets impact their "one-size-fits-all" design?

To shed light on this question, we study the design of Hollywood films. Labeled as "unique and complex artifacts" (Craig et al., 2005), feature films embody "explicit visions of the world" (Moul, 2005), and are rich in cultural meaning (Craig et al., 2005). They thus entail a likely arena for a shift of aggregate cultural preferences to manifest on.[1] Our research design is also informed by the graduality of emerging markets' rise to prominence —a process that has unfolded over at least three decades. Because other relevant factors have also been changing during this period, isolating the referenced preference shift over such long time span would be difficult. We therefore exploit the variation created by a single event, which heightened the global relevance of consumer preferences from leading emerging market. This is, the 2012 relaxation of China's foreign film importation policy.

Since the Cultural Revolution, China has maintained stringent restrictions on foreign film importation. These started to be challenged by the US Government when China joined the World Trade Organization in 2001. During this period, however, there was a large amount of uncertainty regarding whether, when, and how China would respond. Fostered by the broader political circumstances, these restrictions were finally loosened in February of 2012. From the perspective of Hollywood studios, this policy change effectively constituted a one-time, discontinuous, and large increase in the importance of Chinese preferences within the total addressable market. Our empirical analysis aims at deciphering whether, in adapting to the new market environment, Hollywood studios *accommodated* core elements of film design to elements of the Chinese culture.

By 2012, Hollywood was already actively pursuing the booming Chinese theatrical market.[2] Anecdotal evidence suggests that this push may have impacted film design through the inclusion of additional footage (including Chinese actors or Chinese product placement, or scenes shot in China), and the implementation storylines exalting Chinese people, values, or institutions.[3] While noteworthy, these events were reportedly rare. More importantly, they may not necessarily evidence accommodation to the culture of Chinese people. The inclusion of Chinese-targeted additional footage (a mild form of customization) may have been motivated by the very goal of avoiding to accommodate the otherwise standard design. In turn, the implementation of storylines that exalt Chinese themes may have been solely oriented at earning the goodwill of the regulating authority (and with it, market entry). These themes would thus reflect accommodation to the preferences of gatekeepers rather than to those of audiences. With these caveats in mind, we examine a marker of cultural accommodation that is both deeply ingrained in a film's design, while at the same time not a suspected object of scrutiny by Chinese gatekeeping authorities. That marker is the skin color of starring actors.

[1]Walls and McKenzie (2012) provide evidence that suggests that the Hollywood industry has started increasingly cater to foreign preferences. However, the foreign markets that they consider do not include emerging markets.

[2]After tripling total box office revenues in a 5-year period, China became in 2015 the second largest market (after the US and ahead of the UK), accounting for 18% of the approximate $40B global box-office revenues (Motion Picture Association of America, 2015). The market, which already tops the international list in term of number of screens (Lin, 2016), is projected to displace the US in at the top of the box-office ranking by the turn of the decade (Kokas, 2017).

[3]Rosen (2015) cites examples suggesting that Hollywood has begun to progressively avoid negative representations thereof.

Evidence from a variety of domains converges on the notion that people in China (and other Asian societies) exhibit a strong form of *coloristic preferences*, by which pale skin is aesthetically preferred to dark skin. These preferences were forged over 2,000 years ago in the Chinese society. Today, they do not appear to be a purely implicit cultural trait, nor a matter of public controversy. Our review further suggests that coloristic preferences in China manifest more as a "light-skin premium" than a "dark-skin penalty," and that skin tone corresponds to central axis for perceived beauty, particularly for women. As such, skin color plays a defining role in the burgeoning Chinese cosmetics market. To this point, an official of China's cosmetics trade association stated that "skin whitening has a long history in Asia (..) this obsession with whiteness has not faded over time" (Xi, 2011).[4]

In a sample of about 3,300 films, our empirical strategy exploits an important feature of observed patterns of US film entry into the Chinese market. This is, films that enter tend to have a rather specific profile, which did not materially change after the new policy came into place. Based on this observation, and using a rich set of observable film characteristics, we are able to rank films in terms to their exposure to the policy shock. For example, because low-budget, 2D, horror or comedy films rarely enter China, it can be presumed that their castings were less impacted by the policy change than that of high-budget, 3D, action films, which enter the market much more often. We draw inference from a differences-in-differences specification, which implements a comparison between films produced before and after the policy change, across the range of exposure to the shock. The leading threat to the causal interpretation of our estimates is the possible endogeneity of observable and unobservable film characteristics. We are able to address this concern directly, finding no evidence to support it.

According to our preferred specification, the participation of pale-skinned star actors among films in the top decile of shock exposure increased by about 8% as a result of the policy change. We take this result as evidence of cultural accommodation, and refer to it as the *light-skin shift* (LSS). The result is robust to a series of checks, and its causal interpretation is supported by falsification tests and an analysis of the effect's composition. In particular, consistent with the idea that Chinese coloristic preferences impose more stringent beauty standards on women, and better resemble a "light-skin premium" than a "dark-skin penalty," we find that the LSS mainly operated through the more frequent inclusion of pale-skin actresses, and that it was not associated with a sharp decrease in the participation of actors of dark-most skin colors. Further analysis suggests

[4]It is important to highlight that coloristic preferences are not exclusive to the Chinese culture nor to Asia in general. In fact, evidence that coloristic preferences shape socio-economic and cultural outcomes is available from around the globe –including societies with predominantly dark-skinned populations. Several such cases are described in Hall (2012). The World Health Organization (2011) further notes that the consumption of potentially hazardous skin whitening cosmetics is prevalent around the world.) Our focus on China stems from the conflux of factors that make of it an interesting subject of study, namely, this nation's economic prowess, the marked presence of coloristic preferences therein, and its recent foreign film importation policy change.

that the LSS was neither driven by a shift towards the (mostly light-skinned) Superstars, nor the preferences of Chinese gatekeeping authorities.

Although emerging markets have received growing amounts of interest from marketing scholars, the vast majority of this work has treated them as "testing grounds" for theories coined in the context of developed markets, or focused on understanding the nature and local implications of their institutional, cultural, and behavioral differences with developed markets (Burgess and Steenkamp, 2013; Narasimhan et al., 2015; Sudhir et al., 2015; Chandy and Narasimhan, 2015). We contribute to this agenda by highlighting that these differences may also have first-order implications at a global scale. The culturally sensitive nature of our dependent variable further reinforces this point, by suggesting that the emerging markets' newly acquired economic prowess may be propelling their preferences to "swim against the current" in the global stage. Our work also contributes to the literature on product design, where large-scale empirical studies are scant (Bloch, 1995; Luchs and Swan, 2011).

Our results have an obvious implication for the design of international products and services. Whereas analyzing the cultural and behavioral landscape of developed economies may have been sufficient to inform design in earlier days, the irruption of emerging markets renders it likely insufficient now. Moreover, in several ways, the irruption of emerging markets has added preference heterogeneity to the global addressable market, possibly increasing the difficulting of achieving satisfactory "one-size-fits-all" designs. This challenge stands out in the context of feature films, where consumption value is assessed through the lens of culture, and where wide cultural divides may force studios to reconcile somewhat opposing views. Such difficulties resonate, for example, with the case of "The Great Wall" (2016), whose lackluster commercial performance was rationalized on the basis that "the film pulled itself in two directions to please the 'other' audience, which in the end pleased neither side of the world" (Mendelson, 2017a).[5] Although we here focus on coloristic preferences, such cultural divides may also manifest with respect to other culturally sensitive issues, such as women's rights, the scope of personal liberties, and the respect for authority. Storylines and themes that shy away from such polarizing topics, or which avoid stinging representations thereof, may help studios to avoid these unwanted trade-offs.

Although our results connect with the recent controversy about racism in Hollywood (Kang et al., 2014) at a high level, such connection is tenuous at best. We thus ask the reader to interpret this connection with extreme care. Colorism does not equate to racism, and our data do not report actors' races. Moreover, there may be significant variation in skin tones within races, and colorism may manifest within individuals of the same race. More importantly, racism in the US is usually perceived as a systematic prejudice or discrimination against black individuals. Our results do not evidence this effect. Instead they show a favoritism effect for individuals of light-most skin tones. This effect seems to

[5]Mendelson (2017a) also references "Warcraft" (2016) as an example to the same point.

have been sustained by the lower participation of actors of all darker skin tones. Although these darker tones include African-American actors, they also include actors that appear to have Latino, Indian, and Asian ethnicities, as well as many relatively darker-skinned Caucasians (e.g., Adam Sandler, Al Pacino).

2 Coloristic Preferences

Coloristic preferences operate over the skin color of people, such that lighter skin is preferred to darker (Hunter, 2013).[6] The manifestation of these preferences —*Colorism*— does not equate to racism. Colorism is a matter of skin tone, which means that it can manifest within or across races (Hunter, 2013). Furthermore, unlike racism, colorism is not strongly linked to manifestations of violence or repression.

Colorism is also present in the Western World, including the US (Hunter, 2007). For these societies, however, factors such as the more pervasive racial heterogeneity among the population and in the media, and the vigorous affirmative action against racial disparity issues, suggest that overt coloristic manifestations may be met with social disapproval, and that coloristic preferences may not shape beauty standards as strongly as in China.

According to Dikötter (1992), the Chinese society developed a white-black polarity over 2,000 years ago, before the Qin dynasty. As in Europe before the industrial revolution, light skin became a signaling device for class status, differentiating a leisure class elite from lower-class groups. Mostly comprised by the agricultural peasantry, these tended to have darker skin tone because of their more intense exposure to sunlight due to outdoor labor.

Symbolized by the use of "white jade" as a metaphor for a light complexion (Dikötter, 1992),[7] fair skin became idealized and associated with intellectual endeavor, beauty, elegance, and virtue (Sautman, 1994). This symbolism permeated into beauty standards, as is reflected by one point in the "Ten Commandments of Classical Beauty of Ancient China" —a woman's beauty follows from "a skinny waist and snow-white skin"—, and the proverb "One whiteness covers up one hundred ugliness."

While the mechanics of skin color-differentiation were disrupted in Europe by the Industrial Revolution, they remained largely intact in China, which continued to rely on an agricultural economy until the economic reforms of 1978 gave way to industrialization

[6] Evidence has linked coloristic preferences with outcomes such as years schooling (Loury, 2009) and overall educational attainment (Ryabov, 2016), assortment in the marriage market (Hamilton et al., 2009), access to financing (Jenq et al., 2015), and lengths of prison sentences (Viglione et al., 2011; King and Johnson, 2016). Kreisman and Rangel (2015) provide consistent evidence from the labor market by examining earning gaps and tenure.

[7] In Chinese culture, jade symbolizes nobility, perfection and immortality, among other virtues, and it is considered the most valuable of all precious stones.

(Yao, 2000). In Europe, the Industrial Revolution caused the migration of outdoor to indoor labor (Azoulay et al., 2009), and the massification of steam-based transportation technology prompted elites to vacation in warm and sunny locations. Furthermore, relative to the Western world, China has maintained a high degree of ethnic and cultural homogeneity, even throughout its recent process of economic development and global integration (The Economist, 2016).[8]

Observers of the Chinese society rely on this history to explain how skin color, identity, and social status connect in today's China. Referencing Hong Kong as a proxy for Chinese society, Leong (2006) states that "skin color operates as a visual agent in defining the boundaries of cultural identity, and in identifying a person's place in a local social hierarchy," and suggests that white skin is the most important element of personal beauty, as well as a marker for good health. Sautman (1994) reaffirms the role of pale skin as a defining trait of beauty standards in today's China: "fair skin continues to be a standard of female beauty. Many urban Chinese women take pains to avoid the sun and some use whitening creams."

These marked coloristic preferences can be linked to a defining feature of the cosmetics market in China and other Asian countries. For example, the China director of the cosmetics multinational L'Oreal Paris has stated that "(Asian) women at every age want to bleach their skin (..) Since fairness is the specific requirement of Asian women, these kinds of products are only available in this area. But we do believe skin whitening will retain its very important position here" (Xi, 2011). Correspondingly, Asia has been the fastest-growing region in the global skin-lightening market since the 1970s (Tan, 2012), exhibiting rates of growth in excess of 70% in recent years (Xi, 2011). Skin-whitening products now account for about 30% of the over US$5 billion-a-year skin care Chinese market, making it its largest category (Xi, 2011). By comparison, this share is less than 3% in the US (Statista, 2017). The strength of the phenomenon is also reflected by intense competition and innovation in the category (Tan, 2012), as well as by graphic advertisement campaigns which overtly associate departures of pale skin with ugliness and awkwardness (Xi, 2011).

A common thread across these references is the idealization of white skin as opposed to the denigration of black skin. This aspect may be rooted on the mechanics of social differentiation, as they apply to the Chinese context. In particular, because the genetically-determined skin colors of Chinese people are generally light within the broad spectrum of skin colors in humans (Wei et al., 2007), black skin is a virtually irrelevant point of reference, and thus an ineffective anchor to base differentiation on. This suggests that coloristic differentiation in China is likely to be driven by the degree of proximity to

[8]The Economist (2016) reports that "China today is extraordinarily homogeneous. It sustains that by remaining almost entirely closed to new entrants except by birth." This statement is in part supported by recent statistics for naturalized citizens: while the US, Britain, France, and Russia naturalized between 0.1% and 0.4% of their respective populations, China has naturalized 0.0001% in total (i.e., across years).

the "snow-white" ideal, rather than by the degree of separation from black skin extreme, as most Chinese people would be roughly equally apart from it.[9] The common practice of avoiding sun exposure (Levin, 2012) supports this view, as it suggests that even minor distortions to a pale skin tone can jeopardize this ideal. We conceptualize this idea by saying that coloristic preferences in China manifest more like a "light skin premium" than a "dark skin penalty."

We conclude with two examples suggesting that Chinese colorism may influence consumption decisions in areas beyond beauty products. In the early days of the modern Chinese Photography market, Fuji outperformed Kodak. Smith (1996) suggested that the cause was partly rooted on the former's products' skin-lightening effect. Kodak later replicated this feature. The popularity of recent "selfie-enhancing" cameras and apps —known to produce the same effect— atests to the same point.[10] Chinese coloristic preferences could also be linked to the recent controversy sparked by the poster of "Star Wars: The Force Awakens" (2015). The controversy arose because, relative to the poster's US version, the Chinese version visibly minimized the character played by the black actor Jon Boyega, while at the same time highlighted those played by the white actress Carrie Fisher and Harrison Ford (Child, 2015).

3 Regulatory Framework

3.1 China's Foreign Film Importation Policy

After a long ban established during the Cultural Revolution, the importation of foreign films was reinstated in China in 1994, although with strict regulations. The purpose of these regulations was two-fold. The first was to support the domestic film industry by protecting it from international competition (O'Connor and Armstrong, 2015). The second, to control the flow of information and content that may undermine the values espoused by the Chinese Communist Party (CCP), or which portrayed the country or its people in a negative light (O'Connor and Armstrong, 2015).[11] To this day, all films that enter China may be censored, or required to alter some of its content.

Since 1994, there have been two main modes for foreign films to enter the Chinese market. These are, "revenue share" and "flat-fee entry." Revenue sharing has been

[9]In contrast, coloristic differentiation in African societies is portrayed as operating through lighter shades of black (Lewis et al., 2013).

[10]Forbes (Sin, 2016) reports that, due to its initial popularity, the price of the "selfie-enhancing" Exilim TR Casio Camera rose from an initial US$249 to US$800-1000 shortly after its release. Quartz (2016) reports that "selfie-enhancing" Meitu app had over 900 million users by March 2016.

[11]To this effect, Lynch (2016) states that: "in China, films have long been an important propaganda tool to promote socialist values and the hegemony of the CCP (..) Nothing says 'Western Values' quite like a Hollywood movie."

the predominant choice for large-budget Hollywood blockbusters, as it allows studios to participate in the box-office upside. In particular, studios retain about 13% of box-office revenues (Cieply, 2012). Distribution is managed by a Chinese state-owned distributor (O'Connor and Armstrong, 2015). Between 1994 and 2001, 10 films were allowed to enter with this mode. When China joined the World Trade Organization (WTO) in 2001, this quota increased to 20. On the other hand, under the flat-fee model, studios receive a fixed upfront payment, but do not participate in box-office receipts. For this reason, flat-fee entry has traditionally been preferred by independent films or smaller productions.[12]

Although different State organisms may influence importation and censorship outcomes, the primary gatekeeping authority is the State Administration of Press, Publication, Radio, Film and Television (SAPPRFT). Because state-owned enterprises directly benefit from the importation of popular films, importation and content control decisions balance economic and ideological factors (Squire, 2004),[13] suggesting that a film's popularity among audiences may be a precondition for entry into the market. If satisfied, censorship and editing may be used to veto/modify those films that conflict with the gatekeeper's ideological mandate. Alternatively, films with weaker popular appeal may be favored if they align with it.

Content that exalts the Chinese culture, institutions, people, or values is expected to earn the SAPPRFT's goodwill.[14] However, it is not always clear what kind of content will achieve this goal or, in contrast, estrange the authority. For instance, when the animated film "Despicable Me 2" (2013) was denied entry, many suggested the reason was an alleged similarity between the animated characters ("minions") and the former CCP's Secretary General Jiang Zemin, despite the authority's official position that the result was due to low profitability expectations (Child, 2017). Qin (2011) and Grimm (2015) describe the convoluted organizational structure that is responsible for importation and censorship decisions (which includes a number of other organisms besides SAPPRFT), suggesting that the resulting lack of transparency may be a deliberate feature, aimed at providing authorities "with the maximum level of flexibility and efficacy desired" in controlling informational flows (Qin, 2011).

[12] An alternative route of entry entails side-stepping the "foreign film" label by engaging in coproductions with Chinese firms. Although there are no known restrictions for this entry mode, there neither is a clear definition of what constitutes a legitimate coproduction in the eyes of the regulator, nor reassurance that all coproduced films will be allowed in. In fact, in our data, about 45% of coproductions with Chinese firms do not enter the market.

[13] An additional economic incentive for the importation of popular films stems from the reconfiguration of Chinese retail commerce around shopping malls, whose operators have favored theater chains as tenants in order to foster footfall (Financial Times, 2016).

[14] Perhaps the most salient example of studios' strive to earn the SAPPRFT's goodwill corresponds to the "The Martian" (2015), where a mission of the Chinese National Space Administration saves the day for the lead character and NASA (Hoad, 2015).

3.2 2012 Policy Change

Although the revenue share quota was increased from 10 to 20 when China joined the WTO in 2001, the stringent restrictions continued to stand in contradiction with a WTO's central mandate. Namely, they breached the equal right of all enterprises and individuals —local and foreign— to import and distribute goods.

China was then called to altogether eliminate these restrictions. In open contradiction with its plans, China ignored the request, prompting the US government to bring a case against it at the WTO (Voon, 2009). Lacking response, the case proceeded. In 2009 a WTO panel officially determined that China remained in violation of its obligations as a member of the organization (Voon, 2009). The panel later rejected a Chinese appeal, and a formal call to comply was issued in March of 2011 (O'Connor and Armstrong, 2015).

Despite this sustained pressure, China's strong views on safeguarding its tight control over informational flows and protecting the domestic industry introduced a large amount of uncertainty regarding whether, when, and under what terms restrictions would be relaxed.

The new policy was forged during a February 2012 visit of then president-in-waiting Xi Jinping to the US. Optimism for progress in film trade policy was weak at the time as, in an essay published earlier in 2012, then-president Hu Jintao had laid out hostile views towards the cultural influence of the West, and its attempts to interfere with China's path (Wong, 2012). Furthermore, although Hollywood's interests were not absent from the agenda, the focus was placed on bigger-picture issues, such as Human Rights violations, the Syrian crisis, the protection of intellectual rights, and "fair play" in technological trade (Lander and Wong, 2012).

A policy reform was nevertheless agreed upon during the visit. It was struck before dinner on the last day of the visit, after Chinese agencies had signed off to the conditions earlier the same day (Waxman, 2012). The new policy was implemented later in 2012, and conceived as a 5-year transitory framework until a more definitive set of rules were worked out.

This deal —known as the *Memorandum of Understanding* or *Xi-Biden* agreement— changed the conditions for the most sought-after entry mode by large Hollywood studios, revenue-sharing. Importation restrictions were relaxed in two main ways (Cieply, 2012). The first was an increase in the percentage of box office receipts retained by foreign studios, which rose from 13% to 25%. The second was an increase of the importion quota, which rose from 20 to 34. Although the policy reserved the 14 additional slots for 3D or IMAX films, there were no stated limits on 2D films among the remaining slots. Thus, to the extent that the supply of 3D/IMAX films would not meet Chinese demand, the policy change also improved entry prospects for 2D films. Both these changes point to an increased relevance of Chinese audiences' preferences from the perspective of Hollywood studios.

Despite not receiving much coverage by the Chinese media, some interpreted the Memorandum as skillful political maneuvering by the impending President Xi: relative to the other issues at stake during the visit, the agreement constituted a relatively minor concession (Page, 2012). Moreover, the larger number of US films entering China would help to fill the rapidly-expanding number of screens in China, and enable the transfer of technology to the domestic industry. For the White House, the agreement was perceived as "well-timed victory," at a time it verged on the "embarrassing collapse of legislation aimed at protecting intellectual property" (Waxman, 2012). Finally, reflecting the reigning uncertainty and importance of the Chinese market for Hollywood studios, MPAA's Chris Dodd and Ron Kirk called the new agreement a "big deal" (Waxman, 2012) and a "breakthrough" (Abrams, 2012), respectively, while an analyst commented "the unthinkable happened" (Landreth, 2012).

4 Data

The construction of our data set entailed two main parts: the retrieval of information for a sample of films from the IMDb website, and the codification of the skin colors of star actors included on each of these films. Details for each are provided in turn.

4.1 Films

We employed three criteria to select films. First, we focused on films released in the US between 2009 and 2015, a time frame which provided us (at the time of data collection) with the roughly widest symmetric window around the policy change. Given our focus on the Hollywood industry, the second criterion was to select only those films for which the US was listed among the origination countries. The large number of titles available from IMDb includes many small productions with little impact, so we further narrowed down the sample to the set of more impactful films. We implemented this refinement by focusing on films with a large enough number of popularity votes in the IMDb website.[15] The distribution across release years of the 3,378 films in the resulting sample is described by the bars of Figure 1. The number of released films is roughly stable across years.

We coded several design characteristics. Descriptive statistics are presented in Table 1.

[15] The IMDb popularity voting system is designed to minimize the extent of manipulation, for example, by limiting voting rights to registered users and allowing these to issue a single vote for each film. The selected threshold was 500 votes. Parsing through titles with fewer votes revealed that many such films are notoriously less impactful. Moreover, for these, important design characteristics in our analysis below are often missing.

From technical specifications and keywords we generated the indicator 3D/IMAX, which identifies the about 5% of films in our sample with these formats. IMDb data generally associates each film with more than one genre. Drama, Thriller and Comedy are the more frequent ones.[16] We also generated a set of sensitive content indicators (i.e., sex, nudity, violence, drug use, and strong language) by mining the keywords on the parental guide. About 68% of films include at least one such type of content.

Although these variables capture a good amount of variation regarding films' appeal to different audiences, their overall appeal may also be determined by other design features which are not be fully-controllable during production. One of these corresponds to the ratings awarded by the Motion Picture Association of America (MPAA), which classify each film based on the suitability of their content to different audiences. These may add information by capturing higher-level design characteristics, which unfold through the specific portrayal of each type of content.[17] Another non fully-controllable variable corresponds to awards. Previous research has shown that the number of awards that a film is nominated to and wins may signal a film's quality to consumers.[18] We thus constructed the variables AWRDNOM and AWRDWON, which respectively track these.[19]

We also encoded production budgets. By a rational expectations argument, large budget films may be designed to have broad appeal, in a way that is not captured by the above variables. This may occur, for example, through the inclusion of spectacular special effects or popular (high-earning) actors.[20] In turn, Chinese authorities attempting to make efficient use of the small number of revenue-share slots may favor films that are both likely to appeal to a larger share of the population, and yield higher box-office receipts. Because budget information for about 45% of the films in the sample is missing from IMDb, we devised a coding scheme which primarily relies on identifying films at the top of the distribution. In particular, assuming that budget figures are more likely to be missing for films in which these are relatively small, we created the indicators P75BUDGET and P90BUDGET, which respectively activate for films whose budgets lie within the top 25% and 10% of the distribution of budgets for films released each year.[21]

[16]The "other" category includes genres that are less frequent in the data. These are: Western, Musical, Biography, History and Sport.

[17]The system awards the following ratings: PG for films suitable for audiences aged 7 and older; PG-13 for audiences aged 13 and older; and R for restricted audiences. It has been argued that these ratings can, by themselves, impact overall appeal and commercial performance of a film, even after holding content constant (Palsson et al., 2013). The system has been criticized based on an alleged inconsistency, whereby films containing scenes of comparable nature are awarded different ratings (Pomerantz, 2010). Consequently, they may be subject to a degree of uncertainty from the perspective of producers.

[18]Although proxying for a film's quality through awards is problematic in some respects, previous research has made this connection and provided supporting evidence (e.g. Nelson et al., 2001; Ginsburgh, 2003; Deuchert et al., 2005).

[19]We consider the following major awards: Oscars (Academy Awards), Golden Globes, BAFTA Awards, Golden Lion (Venice Film Festival), Palme d'Or (Cannes International Film Festival), Grand Jury Prize (Sundance Film Festival), The Golden Bear (Berlin International Film Festival), The Golden Leopard (Locarno International Film Festival), and Filmfare Awards.

[20]In addition, promotional budgets tend to be proportional to production budgets (Vogel, 2014).

[21]In practice, this codification assumes that missing budgets lie somewhere in the bottom 75% of each

The indicator BIG6STUDIO was created to identify films involving at least one of the large "Big 6" studios.[22] About 37% of the films in the sample fall in this category. This variable may help us, among others, to control for potential differences stemming from studios' experience engaging with Chinese authorities. Further, we generated USONLY-FILM to identify films produced by US studios only. Because these films are produced in absence of international collaboration, they may have a relatively stronger emphasis on US culture or values.

The variable CHINESECOPROD identifies coproductions with companies from China, which could facilitate entry into that market.[23] Furthermore, in order to pick up the potential appeal derived the inclusion of local actors, we retrieved each star actor's country of birth. Because only few of the stars in our data set are reported as born in Mainland China, we also considered actors born in Hong Kong and Taiwan. The indicator CHINESESTAR identifies the small percentage (less than 1%) of films including at least one such actor.

Entry into the Chinese market is codified by the variable CHINAENTRY. An initial inspection of the data revealed that many of the films that IMDb reports as having been shown on Chinese screens do not correspond to standard commercial releases, but to limited releases (for example, in the context of screenings and festivals).[24] Because such releases may be subject to less stringent regulatory oversight, or mainly serve promotional goals, they are not accounted for by CHINAENTRY.

Patterns of entry into the Chinese market according to CHINAENTRY are depicted by the line plots of Figure 1. Whereas for bars (total releases of all US films in our sample) the horizontal access represents films' release year in the US, for line plots it represents the year of release in China. Thus, the relatively smaller numbers for 2009 likely stem from data censoring (as opposed to variations in the regulatory environment).[25]

By focusing on the type of films that are more likely to enter through the revenue-share system (i.e., large budget-films), the red line maps out the impacts of the new policy

distribution. To investigate the validity of this assumption we resorted to IMDb popularity votes. In particular, if we found that budget information tends to be missing for very popular films, our assumption would not be supported by the data. Statistics presented in the Appendix suggest the opposite: films with missing budget information are typically associated with a much smaller number of IMDb popularity votes than those for which budget data is available.

[22] These studios are Disney, Fox, Paramount, Sony, Universal, and Warner Bros.

[23] As defined, CHINESECOPROD accounts for coproductions with Chinese companies or the inclusion of scenes shot in Chinese territory. This variable is constructed primarily based on shooting location and country of origin data for involved firms (available from IMDb Pro), but supplemented with films' stated country of origin. That is, CHINESECOPROD=1 if a one of the films' declared origination country is China.

[24] We identify these in the data through the following markers: festival release, limited release, screening, premiers, and internet release.

[25] That is, some of the films that entered China in 2009 may correspond to films released in the US in earlier years. This censoring will not affect our later econometric results, as estimated models include year fixed effects and primarily rely on cross-sectional film design variation.

changes on entry quite closely. The observed number of this type of films that enters China is about 25 in 2010 and 2011, and a bit larger in 2012 (the new policy took effect during the summer that year). The pattern exhibits a discontinuous increase in 2013, when a new level of about 35 is reached and maintained through the end of the sample period. The differences between the observed pre- and post-policy change numbers with respect to the stated quotas of 20 and 34 may stem from co-production arrangements, or flat-fee entry that we are not able to identify from the data at hand. However, these differences are small.

Also note that, despite that the new policy opens 14 new slots for 3D/IMAX films, the more immediate entry impact is registered among films without these formats. This supports our previous assertion that, despite its formulation, the new policy effectively expanded the market for both types of films during the covered period, 3D/IMAX and not. Lastly, we note that the difference between the total number of films (black line) and large-budget films that enter China (red line) is likely to reflect the extent flat-fee entry, that is associated to smaller budget productions. This difference suggests that only a very small fraction of small budget films (about 1% in our sample) will enter the market. Thus, for the latter, preferences of Chinese audiences are likely to be a much less important consideration for casting.

Finally, accounting for production lags will be important in our analysis. Film production requires a set of activities (casting, shooting, post-production, etc.) which may take several months, or even years. This means that the impact of 2012 policy change may have started to manifest among cohorts of films released in later years. To determine the specific cohorts that were impacted by the new policy, we retrieved production dates. In particular, we considered the date at which filming started, which was available for about 70% of the films in the sample. Using these dates to proxy for the time in which casting decisions were likely already defined, we computed the production lags. These correspond to the number of months between the reported date in which filming begun, and the earliest release date in theaters. The distribution of these lags is presented in the Appendix. The median of this distribution is 17 months, which suggests that the new policy's casting impacts must have been first perceived among the cohort of films released in 2014.

4.2 Skin Color Coding

We retrieved the list of actors starring on each film, and codified their skin color. We used the list of star actors presented on each film's main IMDb profile. Beyond their on-screen relevance —that is, their importance as it relates to the number and types of scenes they participate on— star actors usually have an important role in films' promotional activities, thus representing the most culturally-charged set of actors within the cast. A manual inspection of a number of films revealed that IMDd's listed stars often correspond to those appearing in promotional posters, and that these lists tend to coincide with those of different internet sources (Rotten Tomatoes, Wikipedia).[26]

Discarding actors for whom there was no profile picture available from IMDb, this procedure resulted in a sample of 10,127 starring roles (i.e., actor/film combinations) and 5,442 actors. About 85% of films in the sample have three stars, while little under 2% have only one. About 65% of actors play only one role in the sample, while 15% play two. A relatively small fraction of actors (less than 5%) play more than 5.

We codified the skin color of each these using MTurk.[27] A profile picture of each actor was shown to 5 MTurk coders, asking each to rate the skin color in a scale with the following entries: "very light," "light," "medium," "dark," and "very dark." We did not provide information about the actor's filmography, nor anchored answers by providing examples. We then codified each response in a 1-5 scale, where 1 corresponded to "very light" and 5 to "very dark," and averaged the scores awarded by coders within each actor. Figure 2 presents the resulting distribution. The median of this distribution is 2, which can be interpreted as 50% of actors in having a skin color that is somewhere in between "very light" and "light."

To contextualize these results and assert their validity, we selected a random sample of actors across the range of skin color codes. Figure 3 presents the result. The most important feature of this figure is that it suggests that the MTurk coding results in a reasonably accurate ordering of actors based on their skin color: from left to right, skin colors progressively darken. The figure also suggests that the coding procedure was robust to less-than-ideal profile pictures. Pictures distorted by background light, or with actors wearing hats or hoods, do not evidence blatant coding errors.

To further assess the validity of this procedure we investigated the degree of agreement among coders. We focused on within-actor coding discrepancies, computed as the difference between each coder's rating and the average across coders. Results point to low overall disagreement. The distribution is symmetric, over 75% of coder/actor ratings are within the [-0.6,0.6] interval, and 90%, within [-0.8,0.8]. Furthermore, according to the criteria of Cicchetti (1994), the resulting value of 0.9 for the average absolute-agreement intraclass correlation is "excellent."

To facilitate the description of casting patterns across films, we translated the skin color alternatives presented to coders into a categorical variable. Denoting by c_a the average color score awarded to an actor a across coders, we assign the actor to a category

[26]In the Appendix we present a sample film to illustrate the retrieval of starring casts, and the correspondence of this list with the design of promotional material.

[27]MTurk (Amazon Mechanical Turk) is a crowdsourcing online marketplace, where task requesters post simple Human Intelligence Tasks performed remotely by workers. To minimize cultural variation, we selected US-based coders only.

$k = 1, .., 4$ if $k < c_a \leq k + 1$, with $k = 1$ for $c_a = 1$. We then constructed n_{ki} as the total number of actors that belong to category k and star in film i, and N_i as $\sum_k n_{ki}$. Thus, given the results of Figure 2, n_{1i} not only represents the number of actors in film i whose skin is judged as "light" or lighter, but also the number of actors who have lighter-than-median skin. These variables are summarized in Table 1. Because most films have three color-coded stars, the 1.85 average of n_{1i} suggests that two thirds of actors fall in this category. The smaller average participation of dark-skinned actors ($k = 3, 4$) relative to their frequency in the distribution of Figure 2 stems from the smaller average number of roles played by them.

5 Empirical Strategy and Shock Exposure

5.1 Empirical Strategy

We conceptualize studios' production decisions through *product design pairs* (Y, X), where Y corresponds to observed casting decisions, and X to the remaining film's design observables (e.g., genre, content, etc.). Within the latter we also include observable "behind the scenes" production arrangements (e.g., co-productions, budget, etc.). Consistent with our empirical analysis, we restrict our definition of Y to only encompass stars' skin colors, assuming their independence with respect to other of characteristics of relevance (e.g., acting skills, experience).

The theoretical object on interest is the optimal casting policy, $Y^*|X$. We understand this as the criteria that guide optimal actor selection given a film's characteristics X. For example, films inspired on World Wars may call for light-skinned actors, as these events took place before major waves of African and Middle Eastern immigration into Europe. For these films, the strive for historical accuracy may lead to an "organic" or "natural" gravitation towards light-skinned stars.

To rationalize these policies, we denote by $\mathcal{P}(Y)$ the distribution of preferences for light-skinned stars in the total addressable (global) market. Once the policy change became effective, Chinese coloristic preferences acquired greater relevance within this market, inducing a shift in this distribution. We denote this shift by $\Delta\mathcal{P}(Y)$. Our inference is based on a comparison of optimal casting policies $\mathbb{E}[Y^*|X]$ between films whose characteristics X suggest a relatively large shift $\Delta\mathcal{P}(Y)$, and others for which the shift was presumably smaller. Our focus on X-conditional casting policies means that we aim to identify a potential change in the set or criteria that govern the selection of stars.

We next present a framework that allows us to implement this strategy in a simple and transparent way. In this framework, $\Delta\mathcal{P}(Y)$ is operationalized through a "treatment intensity" function EXPOSURE(X) $\in [0, 1]$. This function is computed based on patterns of China entry observed prior to the policy change. We use three variants of EXPOSURE, one dichotomic and two continuous. The consistency of results obtained from each of these illustrates the robustness of our main finding.

Under certain assumptions, a differences-in-differences specification allows us to identify the causal impact of the policy change on optimal casting policies. At a broad level, the effect is identified by the progression of "pre/post" differences in $\mathbb{E}[Y^*|X]$ along the support of EXPOSURE. Specification details and assumptions are addressed in Section 6. We investigate the validity of the required assumptions, devoting special attention to the potential endogeneity problem rooted on studios' ability to choose films' observable and non-observable design characteristics.

The next subsection develops the basic elements needed to construct EXPOSURE. We employ Probit specifications, which relate observables characteristics X to CHINAENTRY. As expected from our review of Section 3, results are consistent with the idea that a variety of factors —related to both the preferences of Chinese audiences and gatekeepers— influence the likelihood of China entry. We emphasize that disentangling the influence of audiences' preferences relative to those of gatekeepers' is neither an essential question, nor the focus of our analysis. A different question —whether authorities impose a *coloristic bias* on entry outcomes— is, however, relevant to interpret our results. We address this question in subsection 6.6.

5.2 Determinants of Entry into the Chinese Market

The arbitrariness and non-transparency China's importation and content control outcomes suggest that entry into this market can never be assured. Thus, these factors support a probabilistic framework to model films' shock exposure.

We estimate Probit specifications using CHINAENTRY as dependent variable, using the comprehensive list of observable film characteristics X of Table 1 as independent variables.[28] Table 2 presents the obtained results. Columns 1 and 2 correspond to estimates from the sample of films released in 2009-2012; Columns 3 and 4, those from the full sample. Odd-numbered columns do not account for design characteristics that are not fully controllable during production (MPAA ratings and awards). We focus our discussion on the estimates of Column 1, which will later be used to derive our main results.

We first turn our attention to genre indicators. These suggest that films in genres such as action, adventure, thrillers, and romance command relatively higher China entry probabilities. These results coincide with anecdotal evidence, some suggesting that action

[28]Standard checks (variance inflation, condition number) do not point to multicollinearity issues.

movies with special effects have particular appeal among Chinese audiences (Rankin and Kaiman, 2014) and other that movie going has established itself as part of the ritual of romantic courtship in the emerging Chinese middle class (The Economist, 2013).

On the other hand, comedy films exhibit systematically lower entry probabilities. This result likely stems from a large cultural discount, rooted on the fact that humor does not translate well.[29] Interestingly, horror films also exhibit lower entry probabilities. As opposed to the previous effects, this may more directly speak to the ideological mandate of gatekeepers, as authorities are reluctant endorse the promotion of superstition and, in particular, serial-killer type of violence (Martinsen, 2010).[30] Thus, these results illustrate that the preferences of both audiences and gatekeepers may determine China entry outcomes.

The coefficients of P75BUDGET and P90BUDGET suggest that large-budget films are more likely to enter China. This is hardly surprising as, for obvious economic arguments, large budget films must be designed to have broad appeal. The lack of statistical significance for the coefficient of BIG6STUDIO may stem from its high correlation with the budget indicators. The strongly significant and positive coefficient of 3D/IMAX coincides with reports suggesting the special adeptness of these formats to the Chinese market (Sagakian, 2016).

As expected, the coefficient estimate for CHINESECOPROD is positive, large, and strongly significant across specifications. However, it is interesting that this variable does not perfectly predict entry. This finding further supports the idea that entry cannot be assured.

The positive coefficient of CHINESESTAR and negative coefficient of USONLYFILM could reflect both audiences' and gatekeepers' preferences. In particular, while films with Chinese stars may be more popular among audiences, gatekeepers may also see it fit to further support their films. On the other hand, US-only films may have a more marked focus on US values or themes. For this reason, they may not resonate with Chinese audiences, nor with the mandate of gatekeepers.

As in most innovation contexts, the characteristics of the finalized product is not be fully controllable during production. Nevertheless, these will usually still determine the product's appeal. Non fully-controllable design characteristics are included in the specifications of even-numbered columns. The first of these corresponds to film ratings awarded by the Motion Picture Association of America (MPAA). We previously described how these ratings can be important for a film's commercial success, and they cannot be

[29] Speaking to this point, an expert on the Chinese film market states that "Humor is notorious for not translating as well, say, action or thrillers. Some genres cross boundaries with relative ease, but humor is much more difficult" (Andress, 2016).

[30] "Films are forbidden to have the following contents: (...) publicizing obscenity or superstitions or playing up violence" (People's Republic of China, 1996, chapter I, article 5).

perfectly anticipated during production. Even after controlling for the presence of different types sensitive content, results show that films with the more family oriented PG-13 are more likely to enter China (relative to films in the omitted R/non-rated category). The number and type of awards a film receives and is nominated to, is also a potentially important non fully-controllable design characteristic.[31] Results of Columns 2 and 4 support this view by suggesting that the number of major awards received can significantly increase the probability of entering China.

Although estimates across columns suggest that the determinants of film entry into China are relatively stable throughout the covered period, the value and statistical of some coefficients change. To investigate whether these imply a meaningful aggregate effect, we compare predicted entry probabilities across a wide set of specifications and estimation samples.

Table 3 lists the considered samples, specifications, and correlations among predicted entry probabilities. The correlations between probabilities predicted with estimates from the 2009-2012 and 2013-2015 samples (bolded) are of particular interest. Their high values suggest that the policy change did not materially alter the determinants of China entry. That is, after the new policy became effective, the types of design characteristics that made a film more likely to enter China did not change much.[32] This result suggest that basing the construction of EXPOSURE on one or other set of estimates will not significantly impact our results.

5.3 Shock Exposure

We formulate the following EXPOSURE metrics:

$$\begin{aligned}
\text{DEXPOSURE}_i &= \mathbf{1}[\hat{\Pr}(\text{CHINAENTRY}|X_i) \geq p(90)] \\
\text{CEXPOSURE1}_i &= \hat{\Pr}(\text{CHINAENTRY}|X_i) \\
\text{CEXPOSURE2}_i &= \text{Normalized}\Big(1 - \log\big(1 - \hat{\Pr}(\text{CHINAENTRY}|X_i)\big)\Big)
\end{aligned}$$

DEXPOSURE is an indicator variable, which is activated for films whose estimated entry probability is equal or larger than the 90^{th} percentile of the 2009-2012 distribution. CEXPOSURE1 corresponds to the predicted probability of entry, whereas CEXPOSURE2, to a convex transformation of it. The latter is included to account for the idea that

[31] These may increase overall interest among audiences amongst others, by generating word of mouth, or signaling quality (Ginsburgh, 2003).

[32] Furthermore, note correlations between specifications estimated on the same subsample are also generally high. This suggests that non fully-controllable characteristics does not materially condition the likelihood of entry.

the policy change may have had a larger impact at higher ranges of the predicted entry probability.[33] Its normalization re-scales the expression inside the larger parentheses to the unit interval. We found no significant correlation between these metrics and films' production lags.

For our main analysis, we construct these metrics using the results of Table 2, Column 1. Although we have shown that the determinants China entry did not meaningfully change following the policy change, these estimates are preferred because they better reflect studios' information set as they crafted $Y^*|X$ policies in the wake of the policy change. We include extensive robustness checks showing that our main result does not hinge on this specification choice. Figure 4 presents the histogram of predicted probabilities and describes each measure graphically. This figure suggests that, for the bulk of films in the sample, the policy change may have had little impact.

6 Results

6.1 Light-Skin Shift

Here we present our main result, a shift towards the more frequent inclusion of light-skinned star actors, among films with higher exposure to the policy shock. As announced in the introduction, we call this result the "light-skin shift" (LSS). To do so, we operationalize Y^* through the dependent variable SHARELIGHT. We construct this variable with the aim of reflecting the likely manifestation of coloristic preferences, namely, a "light-skin premium" as opposed to "dark-skin penalty." Results of subsection 6.5 support this specification.

We define SHARELIGHT as the total number of star actors in the lightest-skin category (n_{1i}) over the total number of skin-color coded star actors in a film i (N_i). That is,

$$\text{SHARELIGHT}_i = \frac{n_{1i}}{N_i}$$

To illustrate how SHARELIGHT traces a "light-skin premium" consider an alternative metric, AVGCOLOR=$(1/N_i)\sum_k k \cdot n_{ki}$. Consider also two hypothetic films i' and i'' such that $N_{i'} = N_{i''} = 2$, $n_{1i'} = n_{4i'} = 1$, and $n_{2i''} = n_{3i''} = 1$. Under AVGCOLOR, both films would rank equally under coloristic preferences, whereas under SHARELIGHT, film i' ranks higher than film i'' because it includes one actor in the lightest-skin category.

[33] Our reasoning is based on the possible presence of risk aversion on the part of producers. Because entry into China cannot be guaranteed, casting decisions $Y^*|X$ are made in the context of market uncertainty. Because the costs implied by this type of risk are relatively more important, the lower $\hat{\Pr}(\text{CHINAENTRY}|X)$ is, the latter's impact on casting decisions should be disproportionally larger at larger values.

Also note that, since the threshold dividing skin-color categories $k = 1$ and $k = 2$ also corresponds to the median of stars' skin color distribution, SHARELIGHT also represents the share of "lighter-than-median" stars.

Table 4 presents descriptive evidence supporting the LSS. This table compares average values of SHARELIGHT across four groups of films. Films produced prior to the enactment of the new policy averaged SHARELIGHT values of 0.65, irrespective of shock exposure. That is, in the "pre" period, two out of the three stars included in the average film belonged to the lightest skin-color category. Whereas in the "post" period this average remained virtually unchanged for films with lower shock exposure, it increased by about 0.07 for films with higher exposure to the shock.

Figure 5 unpacks these averages temporally. In the "pre" period, trends are similar across exposure groups. In the "post" period, this trend remains roughly stable for low-exposure films but exhibits a discontinuous increase for high-exposure ones. Figure 6 presents an alternative view, by focusing on the share of films which only include light-skinned ($k = 1$) stars. The same discontinuity is observed. To further characterize the effect, we estimate the following differences-in-differences (DiD) specification:

$$\text{SHARELIGHT}_i = \alpha + \beta \text{EXPOSURE}_i \times \text{POST}_{t(i)} + \Theta X_i + \lambda_{t(i)} + \epsilon_i, \qquad (1)$$

where EXPOSURE corresponds to either of the measures introduced earlier. When implemented as DEXPOSURE, the speficication represents a standard 2×2 design. When implemented as CEXPOSURE1 or CEXPOSURE2, EXPOSURE represents a measure of continuous "treatment intensity."[34]

Given the distribution of production lags, POST=$\mathbf{1}[t \geq 2014]$ identifies the cohorts of films that were likely produced after the policy change was announced. The term ϵ is an independent error, and λ a release-year fixed effect. The stand-alone POST is omitted because λ makes it redundant. X is the set of observables used in the China entry probit model. We do not include the stand-alone EXPOSURE variable because its variation is picked up by X. Standard checks do not unveil multicollinearity problems. Throughout our analysis, we report and draw inference from robust standard errors.

The parameter of interest is β. An estimate $\hat{\beta} \neq 0$ would suggest that, after the new policy came into place, films with more "China-oriented" characteristics X, altered star casting guidelines. An estimate $\hat{\beta} > 0$ would be consistent with the descriptive evidence by implying that such alteration favored light-skinned actors. Studios' ability to choose X introduces a natural concern about this inference. We address this concern in subsection 6.3, finding no evidence to support it.

[34]This strategy —pairing a DiD specification with a continuous "treatment intensity" measure computed on "pre" data— has previously been used to analyze natural experiments by Acemoglu et al. (2004), Finkelstein (2007), Lakdawalla et al. (2013), Dranove et al. (2017), among others.

This DiD specification is useful because it controls for two important forms of confounds. We previously illustrated one of these by citing films inspired on World Wars I and II, which develop stories that are better suited for "light-skinned" characters. If this was pervasive in the sample, we may observe baseline differences in the propensity to cast light-skinned stars among films with different characteristics. This effect is controlled for by including X as explanatory variables. In addition, there may exist unobserved temporal variation, possibly arising from shifting conditions in the actors' job market, or from changes in disposable income across consumer segments.[35] This form of unobserved variation is accounted for by the release-year fixed effects. Results below suggest that these temporal trends do not differ based on shock exposure.

Table 5 presents the estimated β coefficients. These point to a robust, positive impact of the policy change on average SHARELIGHT values. The direct correlate of Figure 5 — the estimate obtained when using DEXPOSURE (Column 1)— is similar to that implied by the graph. It suggests that, as a result of the new policy, the share of actors in the lightest skin category increased by about 8% among films in the top exposure decile. Results from continuous EXPOSURE metrics suggest that this effect may have been as large as 22%, for the films with highest measure exposure.

As we noted in Section 3.2, the announcement of the new policy was preceded by a series of legal disputes and sustained pressure by the US to get China to relax (or eliminate) restrictions on foreign film importation. This leads us to question whether, observing this sequence of events, Hollywood studios may have anticipated the 2012 relaxation and accordingly adapted film design in years prior. To address this question we estimate:

$$\text{SHARELIGHT}_i \;=\; \alpha + \sum_{t' \geq 2010} \beta_{t'} \text{EXPOSURE}_i \times \mathbf{1}[t(i) = t'] + \Theta X_i + \lambda_{t(i)} + \epsilon_i$$

Here, the parameter vector $\{\beta_{2010}, .., \beta_{2015}\}$ picks up the pattern by which the LSS unfolded across successive film cohorts. (β_{2009} is normalized to zero to avoid collinearity.) Anticipatory effects would be reflected by $\hat{\beta}_t > 0$, for some $t < 2014$.

Estimates are shown in Table 6. Because statistical significance arises only among the 2014 and 2015 cohorts, these suggest that there were no anticipatory effects. We also note that the LSS seems to have unfolded gradually in 2014-2015. In particular, we obtain

[35] For instance, since the Affordable Care Act became effective in 2010, there has been a steadily declining trend in the rates of uninsured young adults (McMorrow et al., 2015), increasing the propensity of this population segment to quit their jobs and pursue independent professional endeavors. In fact, some evidence (Bailey, 2017; Blumberg et al., 2014) suggests that, by reducing the dependency employer-based health insurance, this trend may have had a positive impact on rates of entrepreneurship. Therefore, to the extent that the decline of uninsurance rates varied across individuals of different races, it is possible to conceive an impact on the relative supply of light- and dark-skinned actors. Furthermore, insofar declining uninsurance rates increased disposable income for some segments more than for others, studios may have adjusted design characteristics (including casting) to better compete for these audiences.

$0 < \hat{\beta}_{2014} < \hat{\beta}_{2015}$, across EXPOSURE metrics. This pattern could be rationalized, for example, if at the time the new policy was announced, a greater share of castings had been defined for the 2014 than the 2015 cohort of films. In turn, this result insinuates that the magnitude of the LSS may have continued to grow during following years. Lastly, note that these results also imply that "pre-trends" did not differ based on shock exposure.[36]

6.2 Robustness and Falsification

6.2.1 Robustness

Here we address a series of caveats about the above results. To save space, unless noted, supporting tables and figures are presented in the Appendix.

We start by noting that one could be concerned that the skin-color coding procedure introduces a violation of ϵ's independence assumption. This would be a particularly serious problem if coding errors favored light-skin coding among actors that appear in "post" films with high exposure to the shock. This pattern would entail an upward bias on $\hat{\beta}$, and potentially lead us to conclude that there was a LSS when in fact there was none. This violation is a-priori unlikely because many actors star in both "pre" and "post" films, with different degrees of shock exposure. We nevertheless assembled the distributions of within-actor coding discrepancies, across exposure levels and periods. The cited violation would be reflected by a DEXPOSURE=POST=1 distribution with relatively higher skewness. However, all distributions have the same roughly symmetric shape.

Next, we reproduced our main results of Table 5, but instead categorizing actors' skin color according to the median score awarded by MTurk coders. Estimates are similar and statistical significance is preserved. We then considered alternative exposure metrics, computed from the different subsamples and specifications for the China entry Probit. Results remain largely robust. We proceeded to implemented a bootstrapping procedure to account for the stochasticity of the exposure metrics.[37] The resulting distributions yield all strictly-positive β estimates ($p = 0$), with medians that closely trace the estimates of Table 5.

Lastly, we replicated the analysis using a matched-samples approach instead. Based on the vector of observed characteristics X, we partitioned the sample of films into groups $g = 1, .., G$ using k-means clustering. With these, we computed $\Delta_g = \mathbb{E}[\text{SHARELIGHT}|g,\text{POST}=$

[36]Estimating the model on 2009-2013 data only, a test for the joint significance of these coefficients yields p-values of 0.96, 0.49, and 0.57, respectively for DEXPOSURE, CEXPOSURE1, and CEXPOSURE2.

[37]We generated 1,000 pseudo-samples, re-sampling (with replacement) in each case the same number of films released in 2009-2012. Using each of these pseudo-samples, we re-estimated the model of Table 2, Column 1, and constructed the exposure metrics for all films. We then took each of these to the full sample and re-estimated specification (1).

$1]-\mathbb{E}[\text{SHARELIGHT}|g,\text{POST}=0]$. The analysis presented in the Appendix suggested the use of $G = 190$, which left us with 176 matched observations. We then regressed this difference on a constant and the exposure variable, which is now defined at the cluster level, computed based on cluster-specific averages of X, and has the same interpretation as before. Results are presented by Table 7. The small value of the constant suggests that observations were overall closely matched, and that the "pre/post" differences in average SHARELIGHT values are solely mediated by films' degree of exposure to the shock. These results reaffirm those obtained from our baseline DiD approach.

6.2.2 Falsification

To further investigate the validity of the causal interpretation of β, we carried out three falsification tests. In the first of these we dropped films released in 2013-2015 and imposed 2011 as a falsified date for the policy change. That is, we are falsely assuming that the 2011 and 2012 film cohorts were produced after the policy change was announced. The resulting coefficient estimates are presented in Panel A of Table 8. They show that our DiD regression correctly delivers a statistically insignificant $\hat{\beta}$.

Our second test focused on the 110 Animation films in the sample. In most cases, an actor's physical characteristics do not map into the characteristics of the animated character that he or she plays —the actor's contribution to the films is almost entirely circumscribed to his or her voice. Furthermore, as it can be seen from a quick Google Images search for the query "Animation Film Posters," images of starring actors tend to not appear in promotional material. Therefore, the heightened relevance of Chinese coloristic preferences should have been less important for the casting of these actors. Panel B of the Table presents the results. Resulting $\hat{\beta}$ values imply a consistently null effect.

Lastly, we implement a test that exploits the same rough logic, but which focuses on the casting outcomes among the 324 voice roles observed in the sample. As with Animation films, coloristic preferences should be less important in this case. Using a Probit model, we tested whether the probability these roles are played by a light-skinned actor ($k = 1$) increased as a result of the new policy. Estimates are presented in Panel C. Again, these show consistently null effects.

6.3 Endogenous Design Characteristics

6.3.1 Observable Characteristics

As pointed out earlier, our main inference may be confounded by studios' ability to choose films' design characteristics. In particular, after the new policy was announced, some studios may have decided to *redirect* their innovation towards sets of characteristics that are more amenable to Chinese audiences.

To see the problem, consider an example. Our results of subsection 5.2 showed that Action and large-budget films are more likely to enter China. Also recall our running example of films inspired on World Wars, which are usually associated to the Action genre, and better suited for light-skinned characters. If, in response to the new policy, studios began to more often allocate large budgets to World War films, these will be associated with generally higher EXPOSURE values in the "post" period. In this scenario, the documented LSS may not be entirely rooted on an alteration of casting guidelines, but also on the "artificial inflation" of EXPOSURE among this type of films, which is better suited for light-skinned characters.

Before addressing this issue empirically, we must note that, although studios "choose X," they do so under constraints. In particular, large studios (which produce most of the highly-exposed films in the sample) employ "portfolio strategies" aimed at curtailing their aggregate risk exposure (Vogel, 2014). These strategies limit the extent to which design characteristics X can be re-shuffled across films within a studio's portfolio, as diversification will usually require them to cover different genres (Perretti and Negro, 2007) and budget levels (Pokorny and Sedgwick, 2010). Furthermore, because films targeting Chinese audiences are riskier (China entry cannot be assured), they may reinforce the need to maintain well-diversified portfolios.

To evaluate this "redirection hypothesis" we present Figure 7, which displays the cumulative distributions of new films along the support of predicted China entry probabilities. Blue lines correspond to distributions obtained on the subsample of films produced before the new policy was announced ("pre" films); red lines, to those obtained from the subsample of films produced after it ("post" films).[38]

The "redirection hypothesis" would be reflected here by a distribution of "post" films that places relatively more mass on higher spectra of predicted entry probabilities. That is, we should observe that "post" distribution first-order stochastically dominates the "pre" distribution. However, we observe almost no difference. The Kolmogorov-Smirnov test for the equality of continuous distributions confirms this graphical finding. In the Appendix we include regression-based analysis which provides evidence to the same point. Thus, this result suggests that the new policy did not cause a "Chinese-oriented redirection" of overall innovation patterns, which in turn alleviates the concerns stemming from the potential endogeneity of X.

[38]For robustness we have added dashed lines corresponding to distributions for variations of the considered sample periods.

6.3.2 Unobservable Characteristics

Our results so far indicate that, although the new policy prompted studios to use different casting criteria, it did not prompt them to choose more "China-oriented" vectors of observable characteristics X. Nevertheless, studios may have adjusted other design characteristics that are not observed in the data. Although this hypothesis cannot be completely ruled out, it is possible to address a milder version of it. Namely, one focusing on a leading form of unobservable design characteristic: storylines. Here we test this hypothesis. We find no support for it.

It is important to first highlight that such impact on storylines would not be inconsistent with cultural accommodation. This impact would nevertheless blur the interpretation of the LSS —it would suggest that, at least in part, the LSS may have been the sub-product of a shift to storylines that are better suited for light-skinned characters. Since we do not observe this impact on storylines, we revert to interpreting the LSS as a directly controlled outcome. In particular, as an alteration of casting guidelines "holding storylines constant."

Our empirical test relies on a measure of storyline-similarity between pairs of films. To generate a similarity metric, we draw from Eliashberg et al. (2007) and Eliashberg et al. (2014), who apply natural-language processing to films' text descriptions. In particular, these authors use a "Bag-of-Words" (BW) approach to systematically address the nature of films' storylines and covered themes, information which then proves helpful to predict commercial performance.

As emphasized by the first of these articles (and implemented by the second), natural-language processing techniques would be ideally applied to scripts, which provide rich descriptions of films' storylines. However, scripts are not publicly available on a systematic basis.[39] Thus, like in Eliashberg et al. (2007), we apply BW to shorter texts that summarize storylines. In particular, we use IMDb's "summary plot" and "synopsis" entries, which are available on a reasonably systematic basis. Plots are shorter and less structured than synopsis.[40] Many films have more than one summary plot posted on the website whereas, when available, there is only one synopsis. After discarding non-usable entries, plots are available for about 80% of non-Animation films in the sample; synopsis, for about 44%.

BW codifies each text description as a vector of word densities. Film similarity can then be assessed by comparing the distributions contained thereby. To make these comparisons we employ the "Cosine similarity" metric, which is commonly used for this purpose due to its intuitive formulation (Singhal, 2001). Namely, this metric approaches zero

[39] Eliashberg et al. (2014) get around this by constructing a sample from (300) publicly available scripts.

[40] Plots are written by IMDb users. Although IMDb does not clarify the source of films' synopsis, it is likely that they are provided by the producers.

for highly dissimilar vectors, and one for highly similar ones. Intermediate values codify intermediate degrees of similarity. To avoid common confounds, we run the algorithm on a pre-processed dataset.[41] For every pair of text descriptions (i, j), we label the resulting variable SIMILARITY_{ij}.

The resulting distribution of SIMILARITY provides some validation for the approach. For example, even though used texts do not include films' titles, there is a strong convergence of themes among pairs with highest SIMILARITY values.[42] Furthermore, the average plot-SIMILARITY among same-film pairs is over five standard deviations larger than that for different-film pairs. A final piece of validatory evidence is available from the idea that higher SIMILARITY values should be expected from pairs of films associated to the same sets of genres, or which exhibit similar types of sensitive content.[43] That is, SIMILARITY should decrease with the "distance" between two films' genre and sensitive content profiles. To verify this, we measured this distance through the Euclidean norm of vectors of genre and sensitive content indicators (respectively $||X^G_{ij}||$ and $||X^C_{ij}||$), and estimated:

$$\text{SIMILARITY}_{ij} = \gamma + \phi||X^G_{ij}|| + \varphi||X^C_{ij}|| + \epsilon,$$

Our hypothesis is supported by negative, statistically significant estimates for ϕ and φ. These results are presented in the Appendix.

Using SIMILARITY, we now test whether higher shock exposure can be linked to larger changes in films' storylines. We rely on the clustering procedure used before, which partitioned the sample of films into groups $g = 1, .., 190$ based on their observable design characteristics X. We construct a sample composed of unique, non-ordered film pairs (i, j) that belong to the same group g. We further restrict the sample to the set of pairs which either include two films produced before the policy change, or which include one produced before and the other after it. That is, we consider the sample in which each observation is a pair of films in the set $\{(i, j) : i \neq j, g(i) = g(j), \text{POST}(t(i)) \times \text{POST}(t(j)) = 0\}$. Defining $\text{CROSSPER}_{ij} = \mathbf{1}[\max\{\text{POST}(t(i)), \text{POST}(t(j))\} = 1]$ to identify pairs that include one film produced after the policy change, we estimate:

$$\text{SIMILARITY}_{ij} = \alpha + \beta\text{EXPOSURE}_g \times \text{CROSSPER}_{ij} + \mu_g + \epsilon,$$

[41] We applied three filters. First, we removed punctuation and "stop words," which are common words that usually do not add meaning to the text (e.g., "a," "and," "is," "the"). Second, we applied a "stemming" procedure, aimed at distilling a base from derivatives and inflectional forms. For example, after the stemming procedure is applied, "cars," "car's" and "cars' " are all be replaced by "car." Lastly, we removed personal names. In preliminary runs, we found that personal names artificially inflated measured similarity between films. We removed them using the list of each film's characters and a library of personal names.

[42] Three of the ten most similar pairs correspond to sequels, while among most of the remaining, both films develop the same, distinctive topic (e.g., rodeo, vampires, witch hunts).

[43] For example, the storylines of many thrillers revolve around a crime, while many comedies are inspired on stories about college.

where μ_g represents a cluster-specific fixed effect. A stand-alone EXPOSURE variable is omitted because μ_g makes it redundant. As in our earlier analysis, EXPOSURE is computed based on cluster-specific averages. Recall that, since all pairs in the constructed sample have been matched on X, including these as independent variables does not add much information.[44] With this, β picks the potential impacts of the policy change on storylines. In particular, this parameter captures the potential role of EXPOSURE as mediator of cross-period differences in SIMILARITY. An estimate $\hat{\beta} < 0$ would indicate that storylines were altered a consequence of the policy change.

Panel A of Table 9 presents the results obtained from the plots data. These suggest that storylines were not impacted. Results obtained from the synopsis data (Panel B) yield the same conclusion. This finding is supported by further results presented in Appendix, which rely on a more sophisticated natural-language processing technique, "Term Frequency-Inverse Document Frequency" (TF-IDF), which gives more weight to words that are more frequent in a particular document, and less weight to those are more frequent in the entire collection of texts in the sample (Leskovec et al., 2014). Thus, we conclude that the policy change did not alter the nature of storylines developed by films targeting the Chinese market. In turn, this conclusion supports interpreting of the LSS as the result of a direct control of casting guidelines.

6.4 A Superstar Shift?

The distribution of actors' skin colors in the sample (Figure 2) shows that the majority of star actors have a generally light skin tone. Based on this observation, one may wonder whether the LSS was motivated by the more frequent casting of more popular/famous actors (superstars), but who happen to be light-skinned. If this "superstar shift" hypothesis was true, we should conclude that the LSS was not motivated by the strive to cast more pale-skinned stars, but instead, more popular ones.

Before turning to our analysis, we note that, although entirely plausible, the "superstar shift" could be tempered by supply- and demand-side factors. Namely, by definition, there is a limited number of superstars, who can appear in a limited number of films. From a demand perspective, superstars may also be "too expensive" to be used as a "wink" to Chinese audiences, particularly given that entry into that market cannot be assured.

To address this question we retrieved IMDb's "Starmeter" data, which has been previously used as a proxy for actors' fame/popularity (Mathys et al., 2016; Karniouchina, 2011). IMDb describes this metric as a measure of "what people are interested in, based not on small statistical samplings, but on the actual behavior of millions of IMDb users." The website constructs it using proprietary, undisclosed methods, based on the in-site

[44]Specifically, results do not change when we also include $||X_{ij}^G||$ and $||X_{ij}^C||$.

search behavior of the over 250 million monthly IMDb users. The resulting metric enables a ranking, with lower values indicating higher popularity.

We obtained complete "Starmeter" series for each actor in the sample. Values are reported at the weekly level. Reflecting promotional activities and viewership, these series exhibit considerable short-time variation, which we mute by averaging within years. For each actor/film combination, we consider this value during the year prior to the film's release. For an actor a featuring in film i, we label this variable as FAME_{ai}. Given the definition of the Starmeter metric, lower values of FAME indicate higher popularity.

To investigate the validity of FAME, we first look at the number of roles played by each actor. Here, we would expect more famous actors to play a larger number of roles. Data confirm this prediction.[45] Next, we consider films' budget levels. In this case, we would expect higher-budget films to cast more famous actors. The data also confirm this prediction, by showing that films with top 25% and top 10% budgets are much more likely to cast the 10% and 1% more popular stars. The supporting table is presented in the Appendix.

Our test relies on the variable $\text{SHARELIGHT\_FAME90}_i$, which is computed as the share of star actors in a film i, who both (i) belong the lightest-skin category ($k = 1$), and (ii) belong to the 90% less popular actors within the set of all those starring in films released on a given year. That is, SHARELIGHT_FAME90 reflects the participation of light-skinned, non-superstar actors in each film. We then regress SHARELIGHT_FAME90 on specification (1). An estimate $\hat{\beta} > 0$ would lead us to conclude that the LSS was not driven by a shift towards light-skinned superstars, and thus reject the "superstar shift" hypothesis.

Panel A of Table 10 shows the estimated coefficients. These point to $\hat{\beta} > 0$, prompting the rejection of the "superstar shift" hypothesis. In fact, these estimates are somewhat larger than those obtained from the baseline SHARELIGHT variable, suggesting that the LSS may have represented a stronger turn towards not-so-popular light-skinned actors. Panel B shows analog results for SHARELIGHT_FAME99, which is constructed analogously, but focusing on the 99% less popular actors. These results also suggest that the hypothesis should be rejected.

Although Chinese people have for some time now been able access to IMDb, it is possible that Starmeter data do not accurately reflect popularity in China. To evaluate whether this caveat changes our results, we would ideally employ a variable like the IMDb's Starmeter, but which is constructed based on the browsing behavior of Chinese people only. Having been unable to identify such data, we leverage the available information

[45] Actors in the decile of highest popularity play an average of 4.8 roles; those in the next decile, 3.1; those in the decile that follows, 2.3. This monotonic progression continues to hold as we move down in the rank of measured fame. The average number of roles played by actors in the lowest-popularity decile is 1.1. These results also hold when we consider actors' mean FAME in the sample.

from IMDb.

Following the results of Mathys et al. (2016) —which suggest that an actor's previous number of roles increases consumer interest in his/her personal brand— we adjust FAME by each actor's number starring roles in China releases previous to the release of film i ($\text{PREVCHINAROLES}_{ai}$).[46] We construct CHINAFAME_{ai} as:

$$\text{CHINAFAME}_{ai} = \frac{\text{FAME}_{ai}}{1 + \text{PREVCHINAROLES}_{ai}},$$

Notice that, because lower values of FAME represent higher popularity, a larger value of PREVCHINAROLES increases the popularity in China that is measured by CHINAFAME.[47] Using this variable, we construct and analyze SHARELIGHT_CHINAFAME90 and SHARELIGHT_CHINAFAME99 as before. Panels C and D of Table 10 present the results. Although estimates are somewhat smaller and estimated less precisely, they still evidence the LSS. We conclude that, although Chinese popularity seems to have sustained part of the LSS, it did not drive the effect.

6.5 Composition

Our review of Coloristic preferences in China suggested that these may manifest more like a "light skin premium" than a "dark skin penalty," and that the pale-skin standard may be more stringent for female than male beauty. In this section we investigate the whether the composition of the LSS supports these insights.

To implement this analysis we construct a dataset with observations defined at the role ($r = 1, .., R$) level. That is, observations correspond to unique film/actor combinations. We then divide this dataset in two subsamples, one corresponding to all the roles played by females; the other, to those played by males. As in our main analysis, we drop voice roles and Animation films. In the resulting sample, there are 2,779 male actors, who play 5,683 roles. There are 1,863 actresses, who play 3,541 roles. In the Appendix we show that the policy change did not alter proportion of roles played by each.

For the lightest ($k = 1$) and darkest ($k = 4$) skin-color categories, we define the indicator $y_r^k = \mathbf{1}[\text{The individual who plays role } r \text{ belongs to skin color category } k]$, and estimate the following equation separately on each subsample:

$$y_r^k = f(\alpha + \beta \text{EXPOSURE}_{i(r)} \times \text{POST}_{t(i(r))} + \Theta X_{i(r)} + \lambda_{t(i(r))}),$$

where f corresponds to the Probit functional form. When estimated using y_r^1 on the sub-

[46]To construct PREVCHINAROLES we use casting data for films released 2005-2015, and which satisfy the same popularity criteria as in our main sample.

[47]We add one in the numerator because PREVCHINAROLES equals zero in the majority of cases.

sample of female roles, an estimate $\hat{\beta} > 0$ would suggest that the policy change increased the probability that female roles are played by light-skinned actresses. Analogously, if we were to obtain $\hat{\beta} < 0$ on the same subsample, but instead using y_r^4 as dependent variable, we should conclude that the policy change led to a decrease in the probability of female roles being played by actresses that belong to the darkest skin-color category. While both these results would be consistent with the LSS, only the former would support the "light-skin premium" view.

Results are shown by Table 11. Panels A (for y_r^1) and B (for y_r^4) refer to the subsample of female roles; Panels C and D, to that of male roles.[48] Consistent with the more important role of pale skin for female beauty standards, patterns of statistical significance suggest that the LSS primarily unfolded through female roles. Furthermore, this effect is shown to have manifested through the more frequent casting of pale-skinned actresses rather than the less frequent casting of those with dark-most skin colors. The estimate of Column 1 in Panel A implies that, evaluated at the mean, the probability that a female role of a DEXPOSURE=1 film is played by one of the former, increased by about 12% as a consequence of the policy change.

6.6 Were Gatekeepers Responsible for the Light-Skin Shift?

We have previously emphasized that China entry outcomes depend on the composition of preferences of Chinese audiences and gatekeeping authorities. Even though the stated ideological mandate faced by the latter does not offer grounds to suspect that they may impose a coloristic bias on entry or censorship outcomes, one may still wonder if, at least in part, their preferences may drive the documented LSS. Here we test this hypothesis. Results suggest that gatekeepers do not impose a coloristic bias, which leads us to conclude that the LSS was fundamentally driven by the preferences of Chinese audiences.

Our test compares patterns of film entry into China against those into two "control" markets, Taiwan and Hong Kong. Due to their deep cultural connections, these three markets are likely to have similar preferences, particularly coloristic ones. Moreover, because Taiwan and Hong Kong do not restrict film imports, a comparison thereof can shed some light regarding the potential biases imposed by the gatekeepers of the Chinese market.

Despite the existence of many non-trivial differences across these markets, the existence of a relatively large degree of cultural similarity can be argued on the grounds of a shared history. First and foremost, recall that the degree of political independence that Hong Kong and Taiwan today have has been achieved in relatively recent historic

[48] The number of observations in Panel B (D) is smaller than that in Panel A (C) because, due to the lower frequency of actors in the darkest skin-color category, the dependent variable is sometimes perfectly predicted by some combinations of X.

times, over two millennia after coloristic preferences were forged (Dikötter, 1992). Today, populations remain ethnically similar and overwhelmingly homogeneous across the three markets: over 90% of the population of each is ethnic Han.[49] Furthermore, skin-whitening cosmetics are about equally important for all of them (World Health Organization, 2011; Tan, 2012). Importantly, however, in contrast to the Chinese case, film imports into Taiwan and Hong Kong are largely unregulated. Thus, entry outcomes into these markets are fundamentally driven by the preferences of their audiences.[50]

In our data, the openness of the Hong Kong and Taiwanese markets is reflected by the total entry statistics, which appear at the bottom of the Table 12.[51] Despite that the Chinese market is about 25 times larger than each of the other two, the number films that enter them is at least double that of films that enter China. Statistics in this table also insinuate a high degree of preference overlap. For example, out of all the films that entered either Hong Kong or Taiwan, over 55% entered both of these markets. In addition, only 30 out of the 248 films that entered China entered neither Hong Kong nor Taiwan. As a comparison, when we instead consider two different Asian markets, Vietnam and Thailand, this number is more than double.

To investigate the hypothesis of interest we first consider the China entry Probit model of subsection 5.2. We enrich its specification by adding SHARELIGHT and an indicator activated for films that enter either Taiwan or Hong Kong (TWHKENTRY).[52] By including the latter, the specification controls for the common appeal of the films' overall design across markets. Because stars' skin color is part of this overall design, and Hong Kong and Taiwan do not have gatekeepers, the coefficient of SHARELIGHT captures the potential coloristic bias of Chinese gatekeepers. Table 13 presents the results, for three variants of the used specification. Because the coefficients for SHARELIGHT and its interaction with POST are statistically insignificant, neither set of results supports the alleged bias of Chinese authorities.

A potential weakness of this test is that it implicitly assumes that all China entry outcomes in the sample contain useful information. If not all films attempted to enter

[49] In China, this is about 92% (People's Republic of China, 2010); in Taiwan, more than 95% (Department of Information Services, Executive Yuan, 2014); in Hong Kong, about 94% (Population Census Office, 2011). For comparison with other countries in which the Han ethnicity has a significant presence, these percentages are 76% in Singapore (Department of Information Services, Executive Yuan, 2014) and 23% in Malaysia (Department of Statistics Malaysia, 2016).

[50] Taiwan first lifted the importation quota and relaxed print control during the 1980s (Wang, 2003), while other restrictions were dropped when Taiwan joined the WTO in 2001. In 2011, 97% of box-office revenues in this market were generated by foreign movies (Jaffe, 2011). Hong Kong, on the other hand, has historically displayed large degrees of trade openness and become a theatrical hub for the region. (Current US official reports do not mention restrictions for entry into this market. See http://2016.export.gov/hongkong/eg_hk_027494.asp). In 2016, over 80% of all films screened in Hong Kong were foreign (HKTDC Research, 2017).

[51] We identified films released in Hong Kong and Taiwan as previously done for China. That is, we excluded limited and festival releases, and others that do not entail broad access by general audiences.

[52] Because the focus of this analysis is entry (as opposed to design), we also include non-fully controllable design characteristics (MPAA ratings, awards), and define POST as $\mathbf{1}[t \geq 2013]$.

China, the influence of the hypothesized bias could be muted by the variation along other design dimensions.[53] To address this issue, we restrict the attention to the 916 films released in at least one of three markets. Note that, when comparing a film i's entry outcome into China and into a market m (Hong Kong, Taiwan), there are three possible outcomes, described by the definition of the variable z:

$$z_{im} = \begin{cases} 3 & \text{if } i \text{ enters China but not } m \\ 2 & \text{if } i \text{ enters } m \text{ and China} \\ 1 & \text{if } i \text{ enters } m \text{ but not China} \end{cases}$$

Under the hypothesized bias, higher SHARELIGHT values would foster the transition from the first to the second outcome and, in extreme cases, to the third.

Our second test rationalizes the variation of z with the following specification:

$$z_{im} = f(\alpha + \beta \text{SHARELIGHT}_i + \Theta X_i + \lambda_{t(i)}),$$

where f corresponds to Ordered Probit functional form, X represents the set observable design characteristics, and λ are release-year fixed effects. Parameter β captures the role of SHARELIGHT as mediator of differential entry outcomes. The hypothesized bias would be supported by a positive estimate of this parameter.[54] Table 14 presents the obtained results. Again here, the lack of statistical significance means that results support neither a baseline coloristic bias of Chinese gatekeepers, nor an accentuation thereof following the enactment of the new policy.

We conclude by noting that the relevance of coloristic preferences among Chinese audiences could be investigated in a more direct way using revenue data, and guided by the question "do films with higher SHARELIGHT values produce superior competitive outcomes in China relative to in the US?" Such strategy would face the usual hurdles of demand estimation, which in this case would include endogenous allocation of promotional resources and release configurations (e.g., number and types of screens in which the film is shown, and the amount of time it is shown for). It would also face the doubts cast by allegations of manipulations of box-office figures in China (Lang, 2017). We are assembling a dataset to pursue this approach in future research.

[53] For example, the coefficient of SHARELIGHT in the above specification is forced to rationalize China entry outcomes for low-budget horror films, which tend to not enter China for reasons other than the potential coloristic bias of authorities (Martinsen, 2010).

[54] As before, because the focus in entry (as opposed to design), we define POST=$\mathbf{1}[t \geq 2013]$.

7 Conclusion

Discussing the China release of "Transformers: Age of Extintion" (2014), the Sony Pictures executive Nigel Clark wrote:

> The seamless integration of the Chinese elements mentioned above, without the appearance of tokenism, is perhaps the most important key to the appeal of the film to the mainstream Chinese audience.[55]

In this paper we have shown that, beyond anecdote, many large Hollywood productions have accommodated their design to elements of the Chinese culture. Whereas in the above statement, Clark primarily refers to a kind of childhood memories that are deeply held by many Chinese adults, we illustrate the point through another distinctive, although more generalized and socially sensitive element of their culture: a marked aesthetic preference for light skin.

The influence of Chinese culture is taken here as an example of a wider-ranged phenomenon: the economic rise and integration of emerging markets has created an influx of "newly endowed" consumers into the global demand for products and services. Because these consumers are known to considerably differ from their counterparts of developed markets in many cultural and behavioral respects, their "global activation" may also imply a shift of aggregate preferences in the total addressable market for international products and services, and thus condition their optimal design.

Although marketing scholars have widely acknowledged and explored the nature and implications of these differences, these have not yet been shown to have the type of first-order, unexpected global impact that we demonstrate. Like Chandy and Narasimhan (2015) —who state that "the changes that are happening in emerging markets today are unprecedented –in scale, scope, and speed– in human history"— we interpret this finding as a testament to their newly acquired global influence, and ensuing relevance for the practice of international marketing.

Our main result —a *light-skin shift* among Hollywood starring casts— is supported by various pieces of evidence and extensive robustness checks. Yet, an important question remains. To which extent did this *light-skin shift* pay-off for Hollywood studios? Considering that box-office performance in China continues to be a hit-or-miss for Hollywood productions,[56] there is a real possibility that studios may not yet fully grasp the deter-

[55] See `https://wikileaks.org/sony/emails/emailid/62426`.

[56] The are several examples of films that have dramatically over- or under-performed in China relative to pre-release expectations. For example, Mendelson (2017b) claims that "Arthur: Legend of the Sword" (2017) and "Power Rangers" (2017) "got their butts kicked" in their opening weekends in China, further stating that the case of "Power Rangers" was a "real tragedy." In contrast, the unexpected China success of "Pacific Rim" (2013) has been pointed by Lee (2016) as the main factor behind its upcoming sequel. The same article states that in 2013, this seemed an "unlikely proposition."

minants of success in that market. Data limitations and multiple layers of endogeneity make this a difficult empirical question, however. We must therefore leave this question for future research.

References

Abrams, R. (February 20, 2012). China's film quota cracked. Variety. *Available at* http://variety.com/2012/film/news/china-s-film-quota-cracked-1118050508/.

Acemoglu, D., D. H. Autor, and D. Lyle (2004). Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of Political Economy 112*(3), 497–551.

Andress, J. (April 8, 2016). What chinese movie audiences want to see. Inverse. *Available at* https://www.inverse.com/article/13967-what-chinese-movie-audiences-want-to-see.

Azoulay, E., A. Demian, and D. Frioux (2009). *100,000 Years of Beauty*, Volume 4. Gallimard, Éditions Babylone.

Bailey, J. (2017). Health insurance and the supply of entrepreneurs: New evidence from the affordable care act's dependent coverage mandate. *Small Business Economics*.

Bloch, P. H. (1995). Seeking the ideal form: Product design and consumer response. *The Journal of Marketing*, 16–29.

Blumberg, L. J., S. Corlette, and K. Lucia (2014). The affordable care act: Improving incentives for entrepreneurship and self-employment. *Public Policy & Aging Report 24*(4), 162–167.

Burgess, S. M. and J.-B. E. Steenkamp (2006). Marketing renaissance: How research in emerging markets advances marketing science and practice. *International Journal of Research in Marketing 23*(4), 337–356.

Burgess, S. M. and J.-B. E. Steenkamp (2013). introduction to the special issue on marketing in emerging markets. *International journal of research in marketing 1*(30), 1–3.

Chandy, R. and O. Narasimhan (2015). Millions of opportunities: An agenda for research in emerging markets. *Customer Needs and Solutions 2*(4), 251–263.

Child, B. (December 7, 2015). Chinese poster for star wars: The force awakens minimises role of black actor. The Guardian. *Available at* https://www.theguardian.com/film/2015/dec/07/fans-cry-racism-as-chinese-star-wars-poster-shrinks-black-actor.

Child, B. (March 23, 2017). China denies despicable me 2 ban. The Guardian. *Available at* https://www.theguardian.com/film/2013/aug/06/china-denies-despicable-me-2-ban.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment 6*(4), 284.

Cieply, M. (February 19, 2012). In china movie pact, more 3-d, less reality. New York Times. *Available at* http://www.nytimes.com/2012/02/20/business/media/more-3-d-less-reality-in-us-china-movie-pact.html.

Craig, C. S., W. H. Greene, and S. P. Douglas (2005). Culture matters: consumer acceptance of us films in foreign markets. *Journal of International Marketing 13*(4), 80–103.

Department of Information Services, Executive Yuan (2014). The republic of china yearbook. *Available at* http://www.ey.gov.tw/Upload/UserFiles/YB%202014%20all%20100dpi.pdf.

Department of Statistics Malaysia (2016). Current population estimates, malaysia, 2014 - 2016. *Available at* https://www.dosm.gov.

Deuchert, E., K. Adjamah, and F. Pauly (2005). For oscar glory or oscar money? *Journal of Cultural Economics 29*(3), 159–176.

Dikötter, F. (1992). *The Discourse of Race in Modern China*. Hong Kong University Press.

Dranove, D., C. Garthwaite, and M. Hermosilla (2017). Market conditions and the nature of innovation: Evidence from the biotechnology sector. *Working Paper*.

Eliashberg, J., S. K. Hui, and Z. J. Zhang (2007). From story line to box office: A new approach for green-lighting movie scripts. *Management Science 53*(6), 881–893.

Eliashberg, J., S. K. Hui, and Z. J. Zhang (2014). Assessing box office performance using movie scripts: A kernel-based approach. *IEEE Transactions on Knowledge and Data Engineering 26*(11), 2639–2648.

Financial Times (February 16, 2016). Silver screen immune to china's consumer slowdown. FT Confidential Research. *Available at* https://www.ft.com/content/939f2c78-2e44-11e6-bf8d-26294ad519fc.

Finkelstein, A. (2007). The aggregate effects of health insurance: Evidence from the introduction of medicare. *The Quarterly Journal of Economics 122*(1), 1–37.

Ginsburgh, V. (2003). Awards, success and aesthetic quality in the arts. *The Journal of Economic Perspectives 17*(2), 99–111.

Grimm, J. (2015). The import of hollywood films in china: Censorship and quotas. *Syracuse Journal of International Law & Commerce 43*, 155.

Hall, R. E. (2012). *The Melanin Millennium: Skin color as 21st Century International discourse*. Springer Science & Business Media.

Hamilton, D., A. H. Goldsmith, and W. Darity (2009). Shedding "light" on marriage: The influence of skin shade on marriage for black females. *Journal of Economic Behavior & Organization 72*(1), 30–50.

Henrich, J., S. J. Heine, and A. Norenzayan (2010). Most people are not weird. *Nature 466*(7302), 29–29.

HKTDC Research (2017). Film entertainment industry in hong kong. *Available at* http://hong-kong-economy-research.hktdc.com/business-news/article/Hong-Kong-Industry-Profiles/Film-Entertainment-Industry-in-Hong-Kong/hkip/en/1/1X000000/1X0018PN.htm.

Hoad, P. (November 30, 2015). The martian shows hollywood's chinese connection has lift-off. The Guardian. *Available at* https://www.theguardian.com/film/filmblog/2015/nov/30/global-box-office-the-martian-the-good-dinosaur-creed.

Hunter, M. (2007). The persistent problem of colorism: Skin tone, status, and inequality. *Sociology Compass 1*(1), 237–254.

Hunter, M. L. (2013). *Race Gender and the Politics of Skin Tone*. Routledge.

Jaffe, P. (March 24, 2011). Will the great film quota wall of china come down? The Guardian. *Available at* https://www.theguardian.com/business/2011/mar/24/china-film-quota.

Jenq, C., J. Pan, and W. Theseira (2015). Beauty, weight, and skin color in charitable giving. *Journal of Economic Behavior & Organization 119*, 234–253.

Kang, C., K. Thompson, and D. Harwell (December 23, 2014). Hollywoods race problem: An insular industry struggles to change. The Washington Post. *Available at* https://www.washingtonpost.com/business/economy/hollywoods-race-problem-an-insular-industry-struggles-to-change/2014/12/19/d870df04-8625-11e4-9534-f79a23c40e6c_story.html?utm_term=.56f2f291139e.

Karniouchina, E. V. (2011). Impact of star and movie buzz on motion picture distribution and box office revenue. *International Journal of Research in Marketing 28*(1), 62–74.

King, R. D. and B. D. Johnson (2016). A punishing look: Skin tone and afrocentric features in the halls of justice. *American Journal of Sociology 122*(1), 90–124.

Kokas, A. (2017). *Hollywood Made in China*. Univ of California Press.

Kreisman, D. and M. A. Rangel (2015). On the blurring of the color line: Wages and employment for black males of different skin tones. *Review of Economics and Statistics 97*(1), 1–13.

Lakdawalla, D., N. Sood, and Q. Gu (2013). Pharmaceutical advertising and medicare part d. *Journal of Health Economics 32*(6), 1356–1367.

Lander, M. and E. Wong (February 14, 2012). With edge, u.s. greets china's heir apparent. The New York Times. *Available at* http://www.nytimes.com/2012/02/15/world/asia/us-seeks-to-size-up-chinas-heir-apparent-during-visit.html?mcubz=0.

Landreth, J. (October 29, 2012). Hollywood film summit draws chinese movie moguls. China File. *Available at* http://www.chinafile.com/hollywood-film-summit-draws-chinese-movie-moguls.

Lang, B. (2017). China box office audit: Why hollywood needed to get tough. Variety. *Available at* http://variety.com/2017/film/box-office/china-box-office-audit-mpaa-get-tough-fraud-1202479488/.

Lee, B. (June 14, 2016). Eastern promise: The hollywood films making their money in china. The Guardian. *Available at* https://www.theguardian.com/film/filmblog/2016/jun/14/hollywood-films-in-china-asia-market-warcraft-the-beginning.

Leong, S. (2006). Who's the fairest of them all? television ads for skin-whitening cosmetics in hong kong. *Asian Ethnicity 7*(2), 167–181.

Leskovec, J., A. Rajaraman, and J. D. Ullman (2014). *Mining of massive datasets*. Cambridge university press.

Levin, D. (2012). Beach essentials in china: Flip-flops, a towel and a ski mask. The New York Times. *Available at* http://www.nytimes.com/2012/08/04/world/asia/in-china-sun-protection-can-include-a-mask.html?mcubz=0.

Lewis, K. M., S. Harris, C. Camp, W. Kalala, W. Jones, K. L. Ellick, J. Huff, and S. Younge (2013). The historical and cultural influences of skin bleaching in tanzania. In *The Melanin Millennium*, pp. 19–38. Springer.

Lin, L. (November 14, 2016). China set to top u.s. in number of movie screens. The Wall Street Journal. *Available at* http://blogs.wsj.com/chinarealtime/2016/11/14/china-set-to-overtake-u-s-as-worlds-largest-cinema-market/.

Loury, L. D. (2009). Am i still too black for you?: Schooling and secular change in skin tone effects. *Economics of Education Review 28*(4), 428–433.

Luchs, M. and K. S. Swan (2011). Perspective: The emergence of product design as a field of marketing inquiry. *Journal of Product Innovation Management 28*(3), 327–345.

Lynch, E. (December 10, 2016). Chinas draft film industry promotion law: Whatdoes it mean for u.s. studios? China Law & Policy. *Available at* http://chinalawandpolicy.com/tag/quota/.

Martinsen, J. (September 10, 2010). Chinese horror films: It was all just a dream. Danwei. *Available at* http://www.danwei.org/film/yang_jian_horror_films.php.

Mathys, J., A. B. Burmester, and M. Clement (2016). What drives the market popularity of celebrities? a longitudinal analysis of consumer interest in film stars. *International Journal of Research in Marketing 33*(2), 428–448.

McMorrow, S., G. M. Kenney, S. K. Long, and N. Anderson (2015). Uninsurance among young adults continues to decline, particularly in medicaid expansion states. *Health Affairs 34*(4), 616–620.

Mendelson, S. (February 23, 2017a). Box office: 'the great wall' targeted american and chinese audiences but pleased neither. Forbes. *Available at* https://www.forbes.com/sites/scottmendelson/2017/02/23/box-office-the-great-wall-targeted-american-and-chinese-audiences-but-pleased-ne #628764c72e23.

Mendelson, S. (May 12, 2017b). Box office: We're (probably) not getting a "power rangers" sequel. Forbes. *Available at* https://www.forbes.com/sites/scottmendelson/2017/05/12/box-office-were-probably-not-getting-a-power-rangers-sequel/#3ec65a3252e4.

Motion Picture Association of America (2015). Theatrical market statistics. *Available at* http://www.mpaa.org/wp-content/uploads/2016/04/MPAA-Theatrical-Market-Statistics-2015_Final.pdf.

Moul, C. C. (2005). *A concise handbook of movie industry economics*. Cambridge University Press.

Narasimhan, L., K. Srinivasan, and K. Sudhir (2015). Marketing science in emerging markets. *Marketing Science 34*(4), 473–479.

Nelson, R., M. Donihue, D. M. Waldman, and C. Wheaton (2001). What's an oscar worth? *Economic Inquiry 39*(1), 1–6.

O'Connor, S. and N. Armstrong (2015). Directed by hollywood, edited by china: How china's censorship and influence affect films worldwide. *US-China Economic and Security Review Commission 10*.

Page, J. (February 18, 2012). Xi wraps up trip; success seen in what didn't occur. The Wall Street Journal. *https://search.proquest.com/docview/922486080?accountid=11752*.

Palsson, C., J. Price, and J. Shores (2013). Ratings and revenues: Evidence from movie ratings. *Contemporary economic policy 31*(1), 13–21.

People's Republic of China (1996). Regulations on administration of films. *Available at* http://www.asianlii.org/cn/legis/cen/laws/roaof382/#0.

People's Republic of China (2010). Tabulation on the 2010 population census. *Available at* http://www.stats.gov.cn/english/statisticaldata/censusdata.

Perretti, F. and G. Negro (2007). Mixing genres and matching people: a study in innovation and team composition in hollywood. *Journal of Organizational Behavior 28*(5), 563–586.

Pokorny, M. and J. Sedgwick (2010). Profitability trends in hollywood, 1929 to 1999: somebody must know something. *The Economic History Review 63*(1), 56–84.

Pomerantz, D. (December 8, 2010). Why movie ratings matter (and why they shouldn't). Forbes. *Available at* https://www.forbes.com/sites/dorothypomerantz/2010/12/08/why-movie-ratings-matter-and-why-they-shouldnt/2/#3aaac82942b1.

Population Census Office (2011). Population census summary results. *Available at* http://www.census2011.gov.hk/pdf/summary-results.pdf.

Qin, J. Y. (2011). Pushing the limits of global governance: Trading rights, censorship and wto jurisprudencea commentary on the china–publications case. *Chinese Journal of International Law 10*(2), 271–322.

Quartz (March 15, 2016). In china, people are spending $1000 on a camera that takes surgery-enhanced selfies. *Available at* https://qz.com/637550/in-china-people-are-spending-1000-on-a-camera-that-takes-surgery-enhanced-selfie

Rankin, J. and J. Kaiman (July 11, 2014). Hollywood zooms in on china's film market. The Guardian. *Available at* https://www.theguardian.com/world/2014/jul/11/hollywood-zooms-in-on-china-film-market.

Rosen, S. (2015). Hollywood in china: Selling out or cashing in? *The Diplomat*.

Ryabov, I. (2016). Colorism and educational outcomes of asian americans: Evidence from the national longitudinal study of adolescent health. *Social Psychology of Education 19*(2), 303–324.

Sagakian, A. (April, 2016). The audience speaks. Qriously Inc. *Available at* http://www.qriously.com/wp-content/uploads/2016/04/MoviegoerSurvey.pdf.

Sala-i Martin, X. (2006). The world distribution of income: Falling poverty and convergence, period. *The Quarterly Journal of Economics 121*(2), 351–397.

Sautman, B. (1994). Anti-black racism in post-mao china. *The China Quarterly 138*, 413–437.

Sheth, J. N. (2011). Impact of emerging markets on marketing: Rethinking existing perspectives and practices. *Journal of Marketing 75*(4), 166–182.

Sin, B. (March 16, 2016). Why a skin-whitening selfie camera is a hit in china. Forbes. *Available at* https://www.forbes.com/sites/bensin/2016/03/16/a-skin-whitening-selfie-camera-is-a-hit-in-china-because-asia-has-a-racist-perce #78c05353e213.

Singhal, A. (2001). Modern information retrieval: A brief overview. *IEEE Data Eng. Bull. 24*(4), 35–43.

Smith, C. (May 29, 1996). Kodak, fuji face off in neutral territory: China's vast market. The Wall Street Journal.

Squire, J. E. (2004). *The Movie Business Book*. Simon and Schuster.

Statista (2017). Retail sales of skin care products in the us in 2012. Asian Scientist. *Available at* https://www.statista.com/statistics/276827/skin-care-products-sales-in-the-us-by-channel/.

Sudhir, K., J. Priester, M. Shum, D. Atkin, A. Foster, G. Iyer, G. Jin, D. Keniston, S. Kitayama, M. Mobarak, et al. (2015). Research opportunities in emerging markets: An inter-disciplinary perspective from marketing, economics, and psychology. *Customer Needs and Solutions 2*(4), 264–276.

Tan, D. (September 18, 2012). Who's the fairest of them all? Asian Scientist. *Available at* https://www.asianscientist.com/2012/09/features/skin-whitening-products-asia-2012/.

The Economist (December 21, 2013). China's film industry: The red carpet. *Available at* http://www.economist.com/news/christmas-specials/21591741-red-carpet.

The Economist (November 19, 2016). Who is chinese? the upper han. *Available at* http://www.economist.com/news/briefing/21710264-worlds-rising-superpower-has-particular-vision-ethnicity-and-nationhood

Viglione, J., L. Hannon, and R. DeFina (2011). The impact of light skin on prison time for black female offenders. *The Social Science Journal 48*(1), 250–258.

Vogel, H. L. (2014). *Entertainment Industry Economics: A Guide for Financial Analysis*. Cambridge University Press.

Voon, T. (2009). Chinameasures affecting trading rights and distribution services for certain publications and audiovisual entertainment products. *American Journal of International Law*, 710–716.

Walls, W. D. and J. McKenzie (2012). The changing role of hollywood in the global movie market. *Journal of Media Economics 25*(4), 198–219.

Wang, S. (2003). *Framing piracy: Globalization and film distribution in greater China*. Rowman & Littlefield.

Waxman, S. (February 19, 2012). How hollywood and joe biden got china to drop a 20-year movie quota. Reuters. *Available at* http://www.reuters.com/article/idUS22096264620120220.

Wei, L., W. Xuemin, L. Wei, L. Li, Z. Ping, W. Yanyu, L. Ying, L. Yan, T. Yan, W. Yan, et al. (2007). Skin color measurement in chinese female population: analysis of 407 cases from 4 major cities of china. *International journal of dermatology 46*(8), 835–839.

Wong, E. (January 3, 2012). China's president lashes out at western culture. The New York Times. *Available at* http://www.nytimes.com/2012/01/04/world/asia/chinas-president-pushes-back-against-western-culture.html.

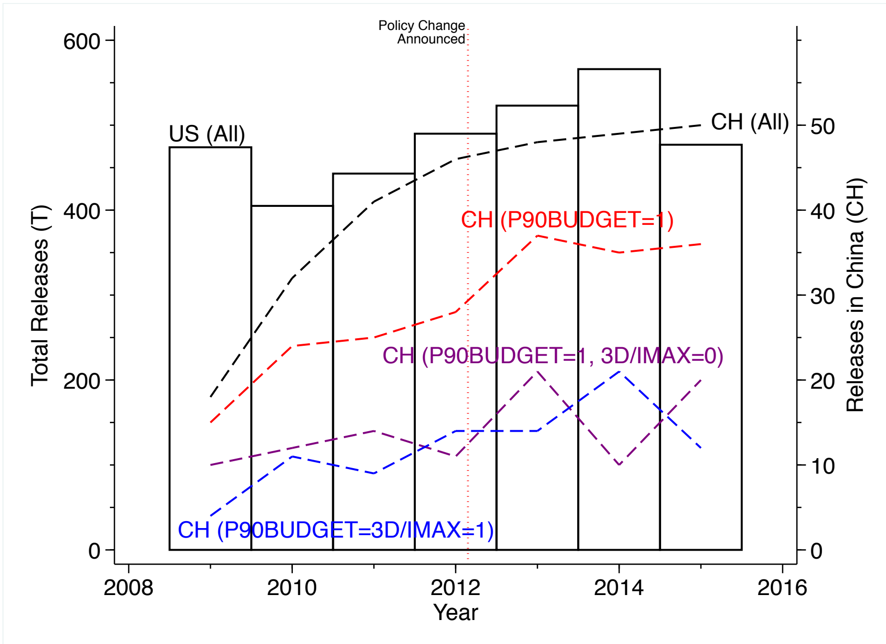
World Health Organization (2011). Mercury in skin lightening products. Preventing Disease Through Healthy Environments. *Available at* http://www.who.int/ipcs/assessment/public_health/mercury_flyer.pdf.

Xi, Z. (September 23, 2011). A lighter shade of pale. China Daily. *Available at* http://usa.chinadaily.com.cn/weekly/2011-09/23/content_13775846.htm.

Yao, S. (2000). How important is agriculture in china's economic growth? *Oxford Development Studies 28*(1), 33–49.

Figures

Figure 1: Number of Films: Sample Total (US) and Releases in China (CH)



Total releases depicted by bars are computed based on films' US release date. Releases in China are computed based on films' date of release in that market.

Figure 2: Distribution of Color Codes Among the Sample of 5,442 Coded Starring Actors

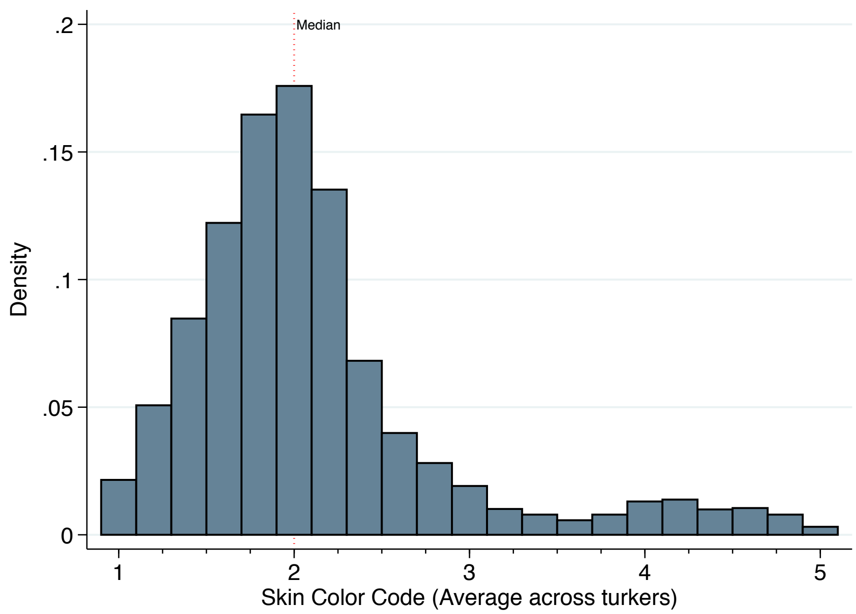
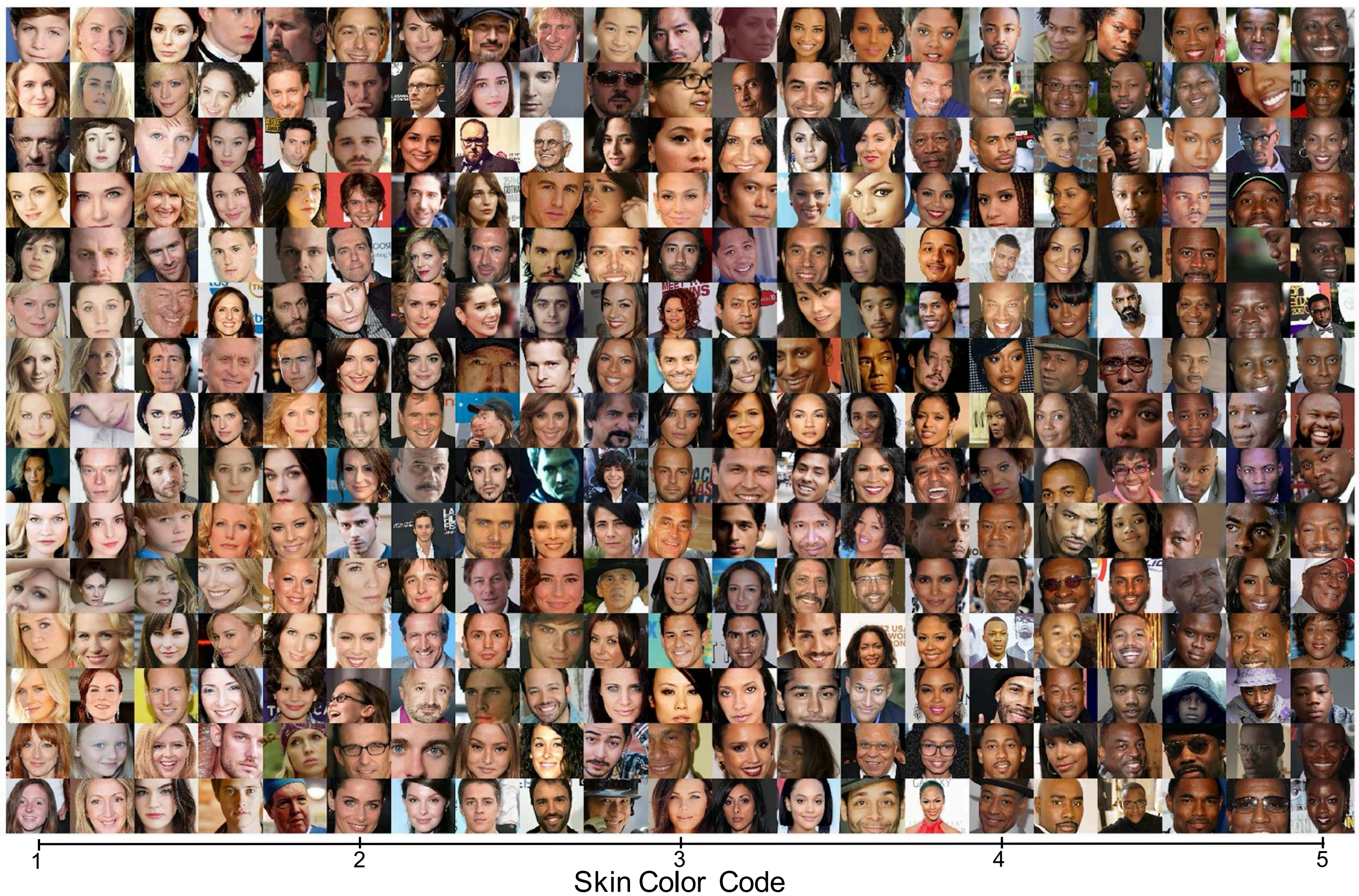
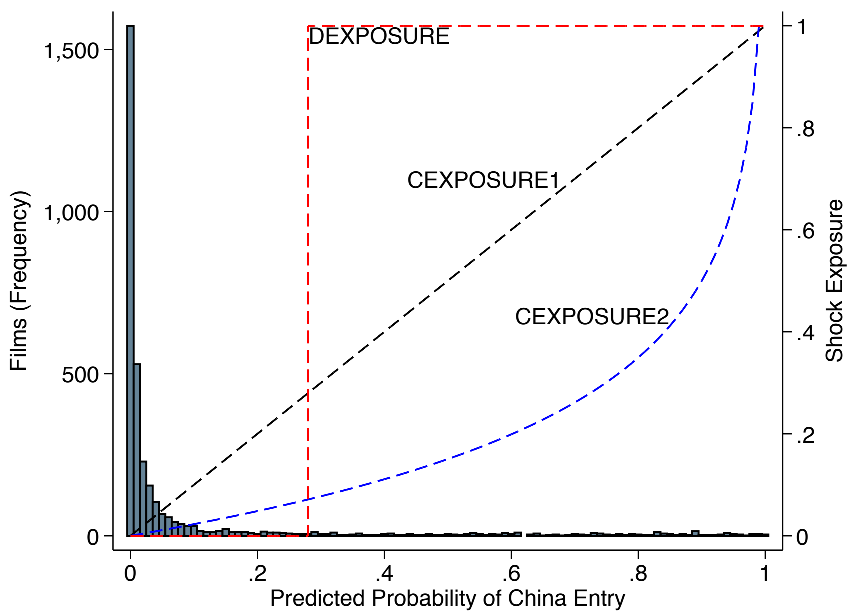


Figure 3: Random Sample of Actors and Actresses by Skin Color Code



Images obtained from IMDb.

Figure 4: Predicted Probabilities of China Entry and Shock Exposure Metrics



Predicted entry probabilities and EXPOSURE metrics are computed based on the estimates of Table 2, Column 1.

Figure 5: Average Shares of Light-Skinned Star Actors by Degree of Exposure to the Policy Change.

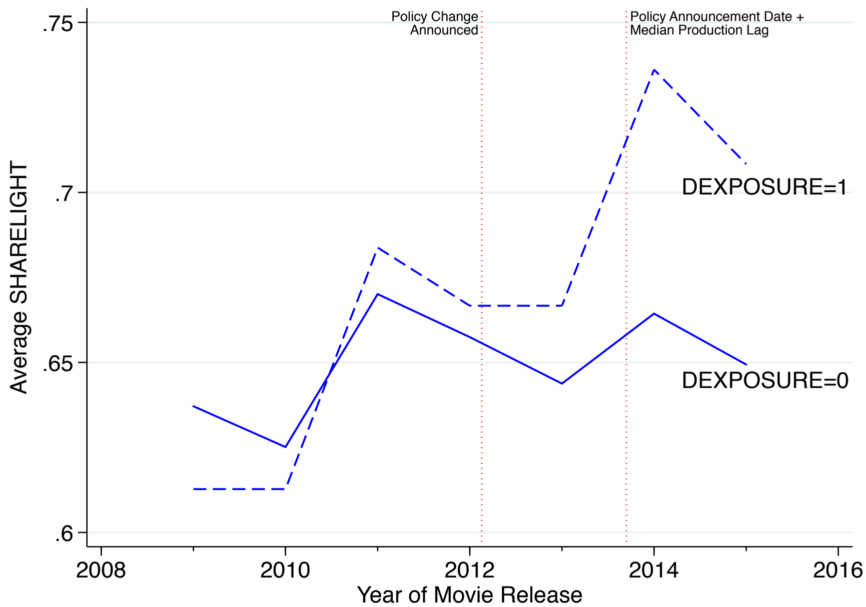


Figure 6: Average Fraction of All-Light-Skinned-Stars Films.

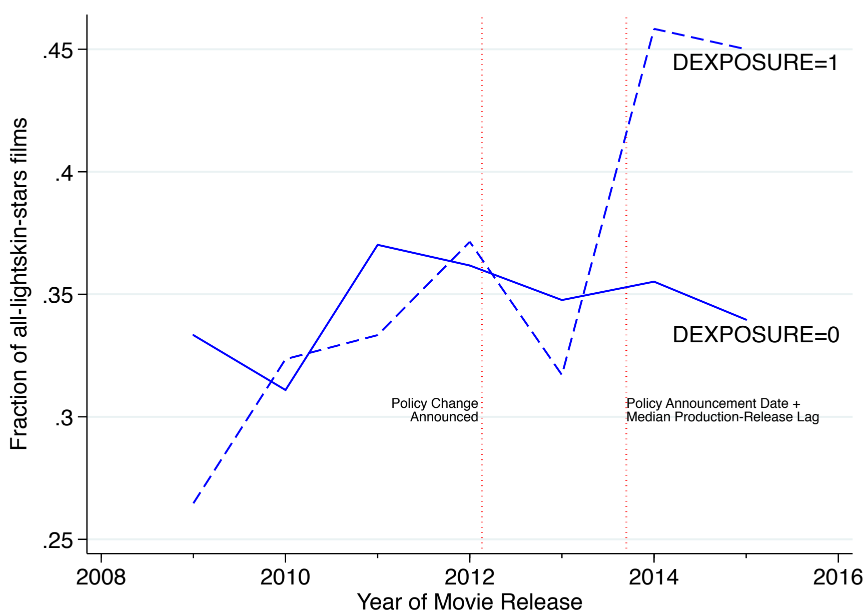
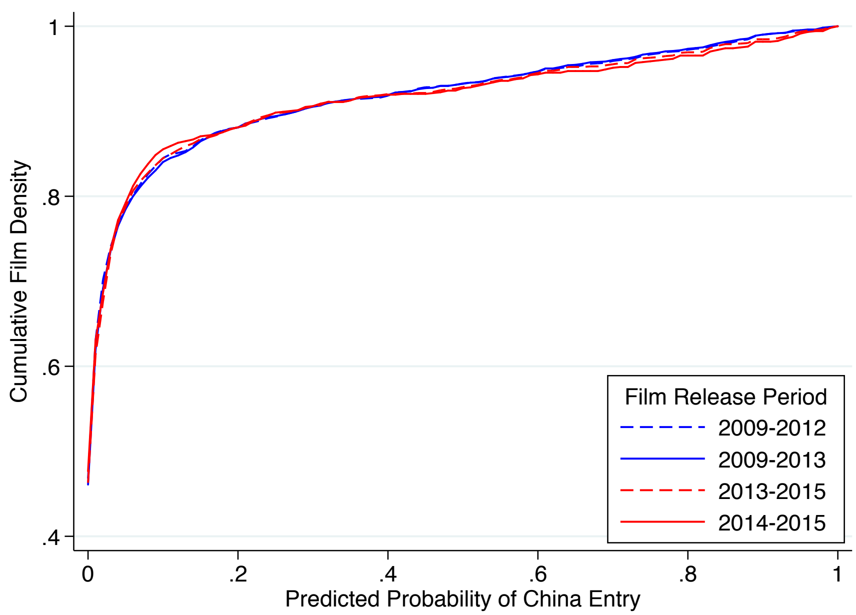


Figure 7: Film Distribution Along the Support of China Entry Probabilities.



Predicted probabilities of China entry are computed from estimates of Table 2, Column 1.

Tables

Table 1: Sample Descriptive Statistics.

	Mean	Std. Dev.		Mean	Std. Dev.
Genre Indicators			Format Indicator		
Action	0.19	0.39	3D/IMAX	0.05	0.21
Adventure	0.11	0.31	Production Indicators		
Animation	0.03	0.18	USONLYFILM	0.77	0.42
Comedy	0.34	0.47	CHINESESTAR	0.00	0.07
Crime	0.14	0.35	CHINESECOPROD	0.01	0.12
Drama	0.48	0.50	BIG6STUDIO	0.37	0.48
Family	0.07	0.25	P75BUDGET	0.26	0.44
Fantasy	0.08	0.27	P90BUDGET	0.10	0.30
Horror	0.21	0.41	Major Awards		
Mystery	0.10	0.30	AWRNOM	0.59	2.61
Romance	0.18	0.38	AWRWON	0.15	0.90
Science Fiction	0.11	0.31	China Entry Indicator		
Thriller	0.34	0.47	CHINAENTRY	0.09	0.28
Other	0.11	0.31	Color Coded Star Actors		
Sensitive Content Indicators			n_1	1.85	0.94
Sex	0.38	0.49	n_2	0.74	0.79
Nudity	0.16	0.37	n_3	0.11	0.37
Violence	0.35	0.48	n_4	0.11	0.37
Drug use	0.19	0.40	N (total)	2.83	0.43
Strong language	0.55	0.50			
MPAA Rating Indicators					
PG	0.07	0.26			
PG-13	0.21	0.40			
R	0.44	0.50			

Table 2: Probit Specifications for CHINAENTRY.

	(1)	(2)	(3)	(4)
	2009-2012		2009-2015	
	Production Indicators			
3D/IMAX	0.94***	0.88***	0.79***	0.74***
	(0.22)	(0.23)	(0.16)	(0.16)
BIG6STUDIO	-0.08	-0.23	-0.01	-0.08
	(0.17)	(0.16)	(0.12)	(0.12)
P75BUDGET	0.51***	0.51***	0.68***	0.63***
	(0.18)	(0.18)	(0.13)	(0.13)
P90BUDGET	1.08***	1.06***	1.02***	0.96***
	(0.18)	(0.19)	(0.13)	(0.13)
CHINESECOPROD	0.96***	1.00***	0.92***	0.95***
	(0.37)	(0.39)	(0.27)	(0.28)
CHINESESTAR	1.36***	1.45***	1.10**	1.19**
	(0.44)	(0.46)	(0.48)	(0.48)
USONLYFILM	-0.55***	-0.54***	-0.59***	-0.56***
	(0.13)	(0.14)	(0.10)	(0.10)
	Genre Indicators			
Action	0.56***	0.61***	0.38***	0.42***
	(0.15)	(0.16)	(0.11)	(0.12)
Adventure	0.35**	0.35**	0.30**	0.27**
	(0.17)	(0.18)	(0.13)	(0.13)
Animation	0.43	0.52*	0.45**	0.51**
	(0.27)	(0.29)	(0.22)	(0.23)
Comedy	-0.69***	-0.77***	-0.47***	-0.47***
	(0.21)	(0.22)	(0.15)	(0.14)
Crime	-0.18	-0.12	-0.15	-0.09
	(0.19)	(0.20)	(0.14)	(0.14)
Drama	-0.15	-0.20	-0.27**	-0.30***
	(0.15)	(0.16)	(0.11)	(0.11)
Family	0.58**	0.75**	0.26	0.37*
	(0.23)	(0.30)	(0.18)	(0.20)
Fantasy	-0.25	-0.29	-0.24*	-0.21
	(0.19)	(0.20)	(0.14)	(0.14)
Horror	-1.15***	-1.01***	-1.40***	-1.31***
	(0.25)	(0.25)	(0.22)	(0.22)
Mystery	0.22	0.19	0.21	0.20
	(0.19)	(0.20)	(0.15)	(0.15)
Romance	0.43***	0.41**	0.15	0.11
	(0.17)	(0.18)	(0.14)	(0.14)
SciFi	-0.02	-0.04	0.24*	0.23*
	(0.18)	(0.20)	(0.13)	(0.13)
Thriller	0.38**	0.44**	0.25**	0.26**
	(0.16)	(0.17)	(0.11)	(0.12)
Other	-0.12	-0.28	0.03	-0.05
	(0.21)	(0.22)	(0.14)	(0.14)

Table 2 (Continued): Probit Specifications for CHINAENTRY.

	(1)	(2)	(3)	(4)
	2009-2012		2009-2015	
	Sensitive Content Indicators			
Sex	0.03	0.03	-0.06	-0.10
	(0.16)	(0.17)	(0.12)	(0.13)
Nudity	-0.34	-0.27	-0.25*	-0.19
	(0.21)	(0.22)	(0.15)	(0.15)
Violence	0.36**	0.33**	0.28***	0.22*
	(0.15)	(0.17)	(0.11)	(0.12)
Drug Use	0.01	0.02	0.03	0.06
	(0.18)	(0.19)	(0.14)	(0.13)
Strong Language	-0.08	-0.09	0.06	0.05
	(0.15)	(0.17)	(0.11)	(0.12)
	MPAA Ratings			
PG		0.38		0.18
		(0.33)		(0.23)
PG-13		0.58**		0.43**
		(0.28)		(0.20)
R		0.19		0.16
		(0.29)		(0.22)
	Major Awards			
AWRNOM		-0.05		-0.00
		(0.04)		(0.02)
AWRWON		0.35***		0.12*
		(0.10)		(0.07)
Constant	-1.95***	-2.25***	-1.85***	-2.01***
	(0.22)	(0.26)	(0.17)	(0.19)
Observations	1,812	1,812	3,378	3,378

Robust standard errors are presented in parentheses. Legend: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Table 3: Pierson correlations for predicted China entry probabilities under different specifications.

Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) 2009-2010	1							
(2) 2009-2010†	0.97	1						
(3) 2009-2012	0.97	0.94	1					
(4) 2009-2012†	0.94	0.96	0.98	1				
(5) 2013-2015	0.91	0.88	**0.95**	**0.93**	1			
(6) 2013-2015†	0.91	0.87	**0.95**	**0.92**	0.99	1		
(7) 2009-2015	0.95	0.92	0.99	0.97	0.98	0.98	1	
(8) 2009-2015†	0.95	0.94	0.98	0.98	0.97	0.97	0.99	1

All correlations are computed with all films in the sample and are significant at a 1% confidence level.
†Model includes award and MPAA ratings variables.

Table 4: Average SHARELIGHT values.

	DEXPOSURE = 0	DEXPOSURE = 1
PRE ($t \leq 2013$)	0.65	0.65
	(0.32)	(0.31)
POST ($t \geq 2014$)	0.66	0.72
	(0.32)	(0.31)

Standard deviations are presented in parentheses.

Table 5: Differences-in-Differences Specifications for SHARELIGHT.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
EXPOSURE×POST	0.08**	0.13**	0.22**
	(0.03)	(0.05)	(0.10)
	[0.02]	[0.01]	[0.02]
Observations	3,268	3,268	3,268

EXPOSURE metrics are computed with the estimates Table 2, Column 1. Estimated models include release-year fixed effects as well as the controls X used in the China entry specification referenced above. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 6: Temporal Unfolding of the "Light-Skin Shift."

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
EXPOSURE $\times\mathbf{1}[t = 2010]$	0.02	0.12	0.19
	(0.06)	(0.09)	(0.22)
	[0.66]	[0.18]	[0.40]
EXPOSURE $\times\mathbf{1}[t = 2011]$	0.03	0.12	0.26
	(0.05)	(0.09)	(0.23)
	[0.55]	[0.21]	[0.25]
EXPOSURE $\times\mathbf{1}[t = 2012]$	0.03	0.13	0.30
	(0.06)	(0.10)	(0.24)
	[0.59]	[0.19]	[0.21]
EXPOSURE $\times\mathbf{1}[t = 2013]$	0.02	0.06	0.00
	(0.05)	(0.09)	(0.28)
	[0.68]	[0.54]	[0.99]
EXPOSURE $\times\mathbf{1}[t = 2014]$	0.09*	0.19**	0.29*
	(0.05)	(0.08)	(0.17)
	[0.10]	[0.02]	[0.08]
EXPOSURE $\times\mathbf{1}[t = 2015]$	0.10*	0.23***	0.47***
	(0.05)	(0.08)	(0.17)
	[0.06]	[0.00]	[0.01]
Observations	3,268	3,268	3,268

EXPOSURE metrics are computed with the estimates of Table 2, Column 1. Estimated models include release-year fixed effects as well as the controls X used in the China entry specification referenced above. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 7: Estimates from the Matched-Samples Procedure.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
EXPOSURE	0.11**	0.15**	0.19**
	(0.06)	(0.07)	(0.08)
	[0.05]	[0.02]	[0.02]
Constant	0.01	0.00	0.00
	(0.02)	(0.02)	(0.02)
	[0.59]	[0.97]	[0.81]
Observations	176	176	176

OLS results. The matching procedure is described in the text. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Table 8: Falsification Tests.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
	A. 2009-2012 Subsample, POST=$\mathbf{1}[t \geq 2011]$		
EXPOSURE×POST	-0.04	0.01	0.21
	(0.04)	(0.07)	(0.18)
	[0.33]	[0.86]	[0.24]
Observations	1,745	1,745	1,745
	B. Animation Films		
EXPOSURE×POST	0.08	-0.10	-0.33
	(0.12)	(0.21)	(0.63)
	[0.51]	[0.62]	[0.61]
Observations	110	110	110
	C. Voice role played by light-skin ($k = 1$) actor (Probit)		
EXPOSURE×POST	0.21	-0.39	-1.22
	(0.35)	(0.55)	(1.43)
	[0.55]	[0.48]	[0.39]
Observations	324	324	324

Panels A and B: OLS specifications. Panel C: Probit specification. EXPOSURE metrics are computed with the estimates Table 2, Column 1. Estimated models include release-year fixed effects as well as the controls X used in the China entry specification referenced above. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Table 9: Storyline Similarity.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
	A. Summary Plots		
EXPOSURE × CROSSPER	-0.001	-0.003	-0.003
	(0.00)	(0.00)	(0.00)
	[0.61]	[0.36]	[0.42]
Constant	0.051***	0.051***	0.051***
	(0.00)	(0.00)	(0.00)
	[0.00]	[0.00]	[0.00]
Observations	23,177	23,177	23,177
	B. Synopsis		
EXPOSURE × CROSSPER	0.004	0.013	0.013
	(0.01)	(0.01)	(0.01)
	[0.56]	[0.13]	[0.26]
Constant	0.086***	0.085***	0.086***
	(0.00)	(0.00)	(0.00)
	[0.00]	[0.00]	[0.00]
Observations	5,201	5,201	5,201

EXPOSURE metrics are computed with the estimates of Table 2, Column 1. Estimated models include cluster-specific fixed effects. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

Table 10: The "Superstar Shift" Hypothesis.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
	A. SHARELIGHT_FAME90		
EXPOSURE×POST	0.10**	0.18***	0.35***
	(0.04)	(0.06)	(0.13)
	[0.02]	[0.00]	[0.01]
	B. SHARELIGHT_FAME99		
EXPOSURE×POST	0.08*	0.14**	0.29**
	(0.04)	(0.06)	(0.13)
	[0.05]	[0.01]	[0.03]
	C. SHARELIGHT_CHINAFAME90		
EXPOSURE×POST	0.06	0.10*	0.23*
	(0.04)	(0.06)	(0.13)
	[0.14]	[0.07]	[0.07]
	D. SHARELIGHT_CHINAFAME99		
EXPOSURE×POST	0.06	0.11**	0.21*
	(0.04)	(0.06)	(0.13)
	[0.13]	[0.05]	[0.09]
Observations	3,268	3,268	3,268

OLS estimates for specification (1). Dependent variables are defined in text and listed as each panel's title. EXPOSURE metrics are computed with the estimates of Table 2, Column 1. Estimated models include release-year fixed effects as well as the controls X used in the China entry specification referenced above. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

Table 11: Composition of the Light-Skin Shift.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
	A. Female Role played by light-skinned ($k = 1$) actress		
EXPOSURE×POST	0.41*	0.73**	1.14*
	(0.21)	(0.31)	(0.68)
	[0.06]	[0.02]	[0.09]
Observations	3,541	3,541	3,541
	B. Female Role played by dark-skinned ($k = 4$) actress		
EXPOSURE×POST	-0.20	-0.07	1.49
	(0.41)	(0.78)	(1.39)
	[0.63]	[0.93]	[0.28]
Observations	3,156	3,156	3,156
	C. Male Role played by light-skinned ($k = 1$) actor		
EXPOSURE×POST	0.16	0.25	0.59
	(0.12)	(0.18)	(0.42)
	[0.17]	[0.16]	[0.16]
Observations	5,683	5,683	5,683
	D. Male Role played by dark-skinned ($k = 4$) actor		
EXPOSURE×POST	-0.10	-0.24	-0.90
	(0.20)	(0.32)	(0.90)
	[0.62]	[0.45]	[0.32]
Observations	5,654	5,654	5,654

Probit specifications. The dependent variables are defined as $\mathbf{1}$[Female (male) role played by female (male) actor in skin color category k]. EXPOSURE metrics are computed with the estimates of Table 2, Column 1. Estimated models include release-year fixed effects as well as the controls X used in the China entry specification referenced above. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table 12: Descriptive Statistics for Entry into China, Hong-Kong and Taiwan (2009-2015)

Markets	Number of Films	Percent
China only	30	3.3%
Hong Kong only	200	21.8%
Taiwan only	135	14.7%
China and Hong Kong only	35	3.8%
China and Taiwan only	16	1.8%
Hong Kong and Taiwan only	333	36.4%
China, Hong Kong, and Taiwan	167	18.2%
Total China, Hong Kong, or Taiwan	916	100%
Total China	248	
Total Hong Kong	735	
Total Taiwan	651	

Table 13: Probability of China Entry as a function of SHARELIGHT and Entry into Taiwan or Hong Kong.

	(1)	(2)	(3)
HKTWENTRY	0.71***	0.69***	0.69***
	(0.15)	(0.15)	(0.15)
	[0.00]	[0.00]	[0.00]
SHARELIGHT	-0.10	-0.11	-0.23
	(0.16)	(0.16)	(0.20)
	[0.54]	[0.50]	[0.26]
SHARELIGHT × POST			0.26
			(0.31)
			[0.40]
Release Year Fixed Effects	No	Yes	Yes
Observations	3,268	3,268	3,268

Probit Estimates. Models are estimated on the full sample of film and include the controls X used in the china entry specifications of odd-number columns in Table 2. POST is implemented as $\mathbf{1} = [t \geq 2013]$. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 14: Differential Probabilities of Entry into China, Hong Kong, and Taiwan.

	(1)	(2)	(3)	(4)
	China and Hong Kong		China and Taiwan	
SHARELIGHT	-0.24	-0.29	-0.28	-0.30
	(0.19)	(0.22)	(0.19)	(0.26)
	[0.20]	[0.19]	[0.14]	[0.24]
SHARELIGHT × POST		0.11		0.04
		(0.36)		(0.34)
		[0.77]		[0.91]
Observations	781	781	716	716

Ordered Probit Specifications. Estimated models include release-year fixed effects as well as the controls X used in the China entry specifications of odd-number columns in Table 2. POST is implemented as $\mathbf{1} = [t \geq 2013]$. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

APPENDIX
Can Emerging Markets Impact Tilt Global Product Design? Impacts of Chinese Colorism on Hollywood Castings

October 2017

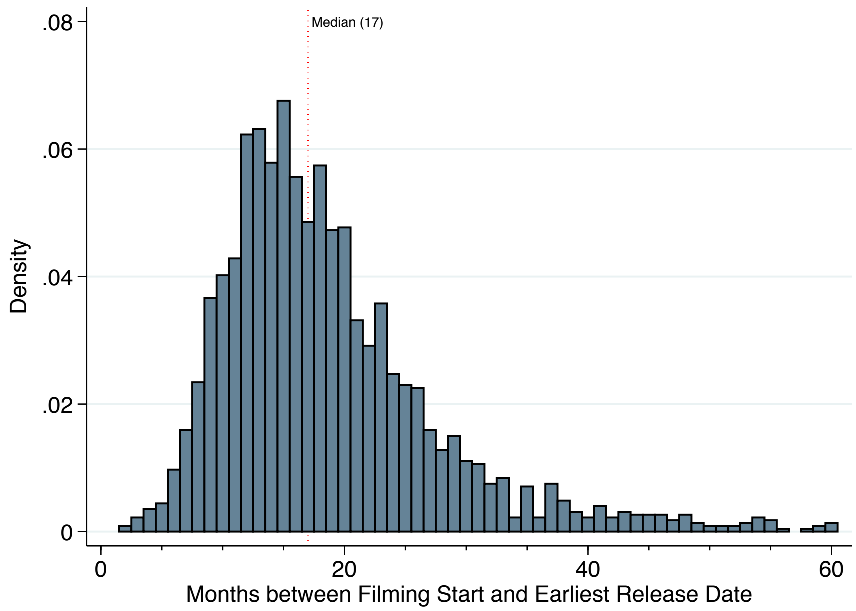
A. Film Popularity and Availability of Budget Information

Table A1: Availability of Film Budget Information and IMDb Popularity Votes.

Budget Information	Number of Films	Percentile of Vote Distribution				
		10	25	50	75	90
Available	1,816	858	2,262	19,886	84,908	207,315
Missing	1,452	622	881	1,869	5,537	15,012
Total	3,268	702	1220	4,573	34,766	119,119

B. Distribution of Production Lags

Figure A1: Distribution of Production Lags

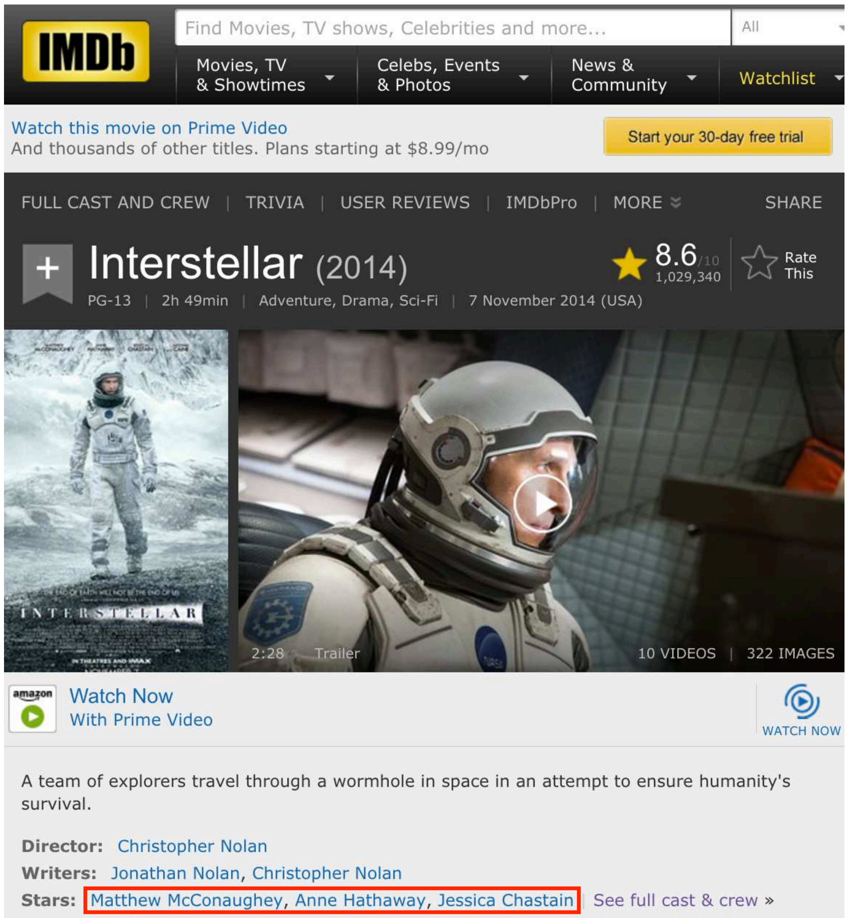


Lags are computed as the number of months between the earliest observed release and the reported filming start date. Computed with data for 2,377 films for which filming dates are observed.

C. Retrieval of Films' Starring Casts from IMDb

Figure A2 illustrates the retrieval of starring casts for each film, using as an example the film "Interstellar" (2014). Retrieved starring casts are highlighted by the red rectangle. Figure A3 presents a promotional poster of the same film, which highlights the actors retrieved from the IMDb page.

Figure A2: Sample Identification of Starring Casts on IMDb: "Interstellar" (2014)



Source: http://www.imdb.com/title/tt0816692

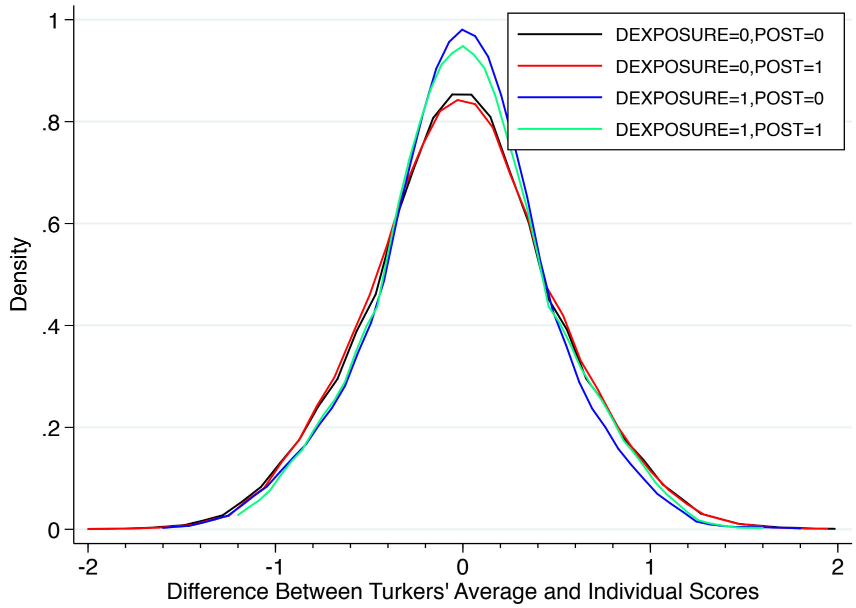
Figure A3: Promotional Poster: “Interstellar” (2014)

Source: http://camartin.deviantart.com/art/Interstellar-Poster-Comp-496032882

D. Robustness for the "Light-Skin Shift"

D.1 Color-Coding Discrepancies

Figure A4: Distributions of Within-Actor Coding Discrepancies. Considers Starring Actors Within each set of Films Separately.



D.2 Classifying Actors Based on the Median (as opposed to Mean) MTurk Scores

Table A2: OLS Specifications for SHARELIGHT (actors categorized based on median MTurk scores).

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
EXPOSURE×POST	0.05*	0.11***	0.23***
	(0.03)	(0.04)	(0.08)
	[0.07]	[0.00]	[0.00]
Observations	3,268	3,268	3,268

Actors are assigned to skin-color categories based on the median score awarded MTurk coders (as opposed to based on the average, as in our main analysis). Estimated models include release-year fixed effects as well as the controls X used in the china entry specification. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: ${}^{*}p < 0.1,{}^{**}p < 0.05,{}^{***}p < 0.01$.

D.3 Computing Exposure Metrics from Alternative Samples and Specifications for the China Entry Probit Model

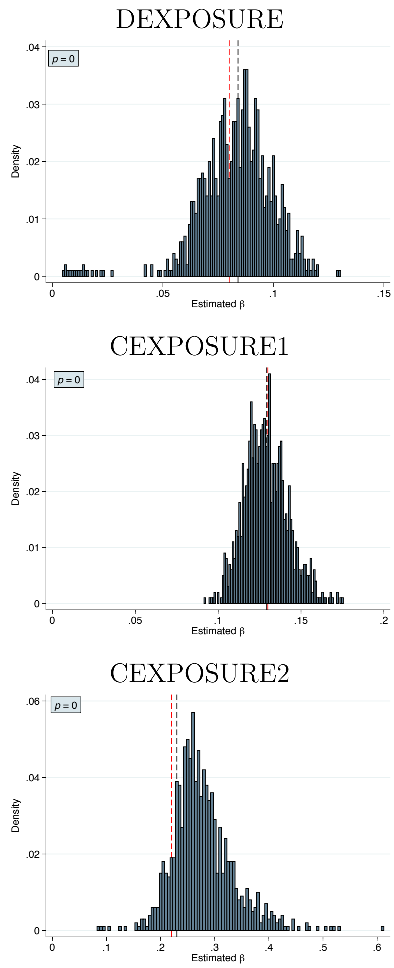
Table A3: Main Results When Exposure Metrics are Computed from Alternative Samples and Specifications for the China Entry Probit Model.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
	(1) 2009-2010		
EXPOSURE×POST	0.07*	0.11*	0.20
	(0.04)	(0.06)	(0.13)
	[0.06]	[0.05]	[0.12]
	(2) 2009-2010, includes non-fully controllable X		
EXPOSURE×POST	0.07*	0.11**	0.37
	(0.03)	(0.05)	(0.27)
	[0.05]	[0.04]	[0.17]
	(3) 2009-2012		
EXPOSURE×POST	0.09**	0.13**	0.22**
	(0.03)	(0.05)	(0.10)
	[0.01]	[0.01]	[0.03]
	(4) 2009-2012, includes non-fully controllable X		
EXPOSURE×POST	0.09**	0.14***	0.54**
	(0.03)	(0.05)	(0.22)
	[0.01]	[0.01]	[0.01]
	(5) 2013-2015		
EXPOSURE×POST	0.06*	0.13***	0.33***
	(0.03)	(0.05)	(0.11)
	[0.07]	[0.01]	[0.00]
	(6) 2013-2015, includes non-fully controllable X		
EXPOSURE×POST	0.05	0.13***	0.31***
	(0.03)	(0.05)	(0.11)
	[0.16]	[0.01]	[0.00]
	(7) 2009-2015		
EXPOSURE×POST	0.08**	0.13**	0.27**
	(0.03)	(0.05)	(0.11)
	[0.02]	[0.01]	[0.02]
	(8) 2009-2015, includes non-fully controllable X		
EXPOSURE×POST	0.09**	0.13**	0.27**
	(0.03)	(0.05)	(0.11)
	[0.01]	[0.01]	[0.02]
Observations	3,268	3,268	3,268

Panels are titled according to the sample and specification used to estimate the China entry probit. (They follow the same order used in the Table of correlations of predicted China entry probabilities.) "Non-fully controllable X" refers to MPAA rating indicators and award variables. Estimated models include release-year fixed effects. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

D.4 Bootstrapping

Figure A5: β Estimates from Bootstrapped Samples.



The Bootstrapping procedure focuses on the estimation of the China entry probability model. We generate 1,000 pseudo samples by sampling 1,812 films released in 2009-2012 (with replacement) and generate the set of exposure metrics from the resulting estimates. For each of these set of estimates, we then estimate specification (1). Black dashed lines mark the median of each distribution; red lines, the respective point estimates from our main results table.

D.5 Details and Robustness for The Matched-Samples Procedure

We apply k-means clustering on the entire vector of observable film characteristics X. Throughout our analysis, we only consider the main sample of non-animation films.

A crucial aspect to implement clustering is to determine the number of clusters. A common way to approach this problem –the "Elbow" method– progressively increases the number of clusters until the marginal reduction of total explained variance stabilizes. Our approach adapts this principle to the overarching empirical design of our research.

In particular, because our inference crucially relies on a "pre/post" comparison, we seek to balance this principle with the goal of producing as many clusters including both "pre" and "post" films, as is possible.

Figure A6: Main elements for determination of the number of film clusters

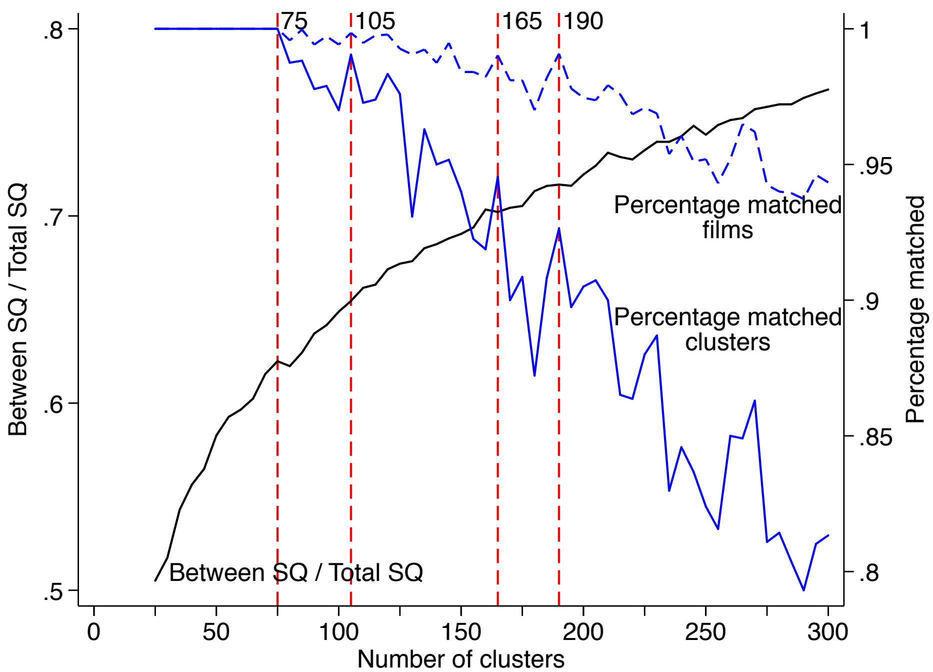


Figure A6 presents the main elements to consider in deciding the number of clusters. The horizontal axis shows the number of clusters considered in each case. We vary these in increments of 5. (Because X contains 25 variables, we consider a minimum of 25.) The black line depicts the percentage of total variance (total sum of squares - Total SQ) accounted for by between-cluster variation (between SQ). The larger this percentage, the higher the share of total variance that is explained by the clustering. The solid blue line shows the percentage of clusters within which there is at least one "pre" and at least one "post" film. We refer to this statistic as the percentage of "matched" clusters. The dashed blue line presents the percentage of all films in the sample included in these clusters.

Although the percentage of variance explained by the clustering increases concavely with the number of clusters, there is no noticeable stabilization of this trend in the considered range. Thus, we primarily base our decision on the number of matched clusters.

As shown by the generally declining blue lines, finer clustering translates into a marked decrease in extent of matching. Up until 75 clusters, there is perfect matching. The percentage of matched films (dashed blue line) remains generally high ($\geq$ 99%) up to 190 clusters, after which it drops without recovering. With this configuration, 93% of clusters (176) are matched and can be used in the main analysis.

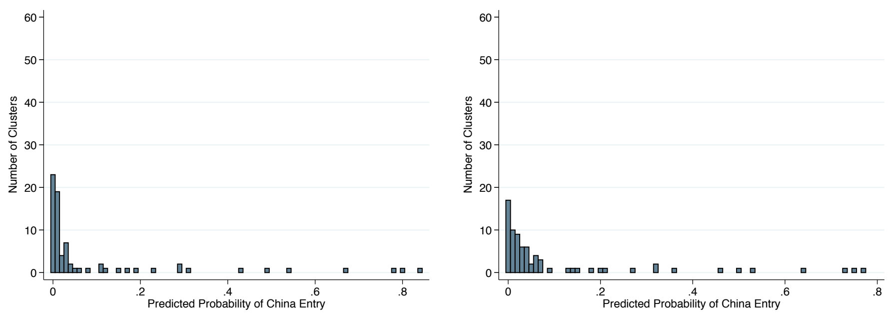
Because it provides the finest categorization of films without implying a significant loss of information, we use the 190-cluster clustering for our main analysis (presented in text). Nevertheless, for robustness, we also consider three alternative, coarser clusterings. These are shown by the dashed vertical lines in the graph. They are selected because they represent a local peak in the number of matched clusters.

Before turning to our main robustness results, in Figure A7 we present the distribution of clusters on the support of predicted probabilities of China entry (which map one-to-one to EXPOSURE metrics). Graphs on the left are computed by predicting entry probabilities after clustering (i.e., computed based on each cluster's average X); those on the right, by predicting these before clustering (predicted probabilities are computed before clustering). Panels A-D reproduce these distributions under the different numbers of considered clusters. Overall, these distributions have similar shapes. More importantly, they all resemble the distribution of films on the same support (see Figure 4 in article).

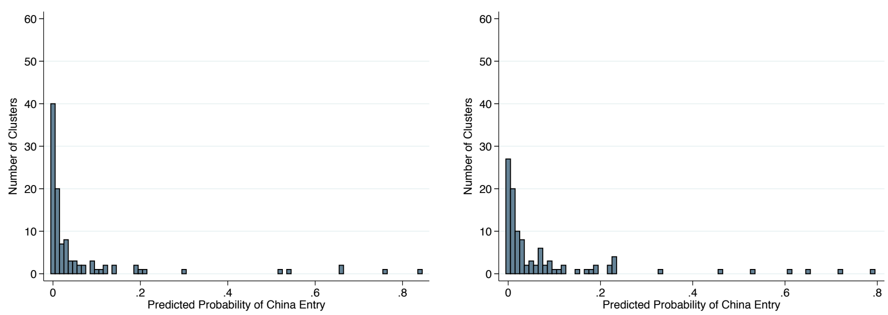
Table A4 reproduces the matched-sample regression presented in the main text, for each case (i.e., number of clusters), when EXPOSURE metrics are based on China entry probabilities predicted after clustering (i.e., using each cluster's average X). Table A5 reproduces the same results, but with EXPOSURE metrics based on China entry probabilities predicted before clustering. Results support the overall robustness of our main estimate.

Figure A7: Main elements for determination of the number of film clusters
(Left: probabilities predicted with average X within each cluster – Right: probabilities predicted with each film's X, then averaged within each cluster)

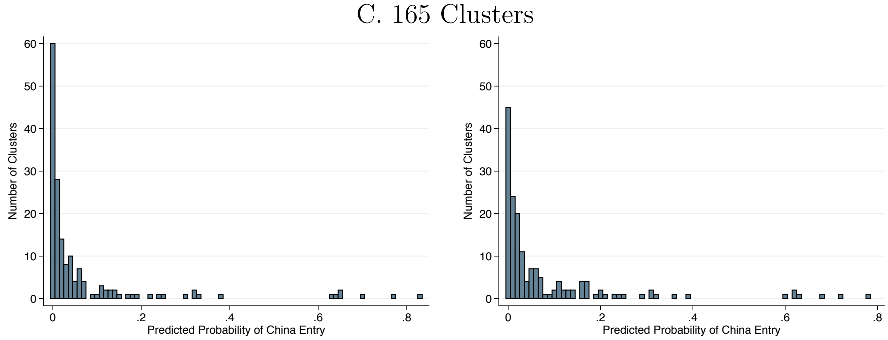
A. 75 Clusters



B. 105 Clusters



C. 165 Clusters



D. 190 Clusters

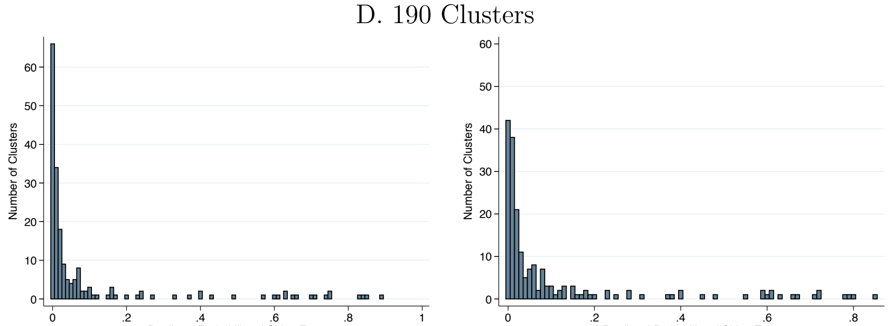


Table A4: Robustness for the Matched-Samples procedure. (China entry probabilities predicted after clustering.)

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
	A. 75 Clusters		
EXPOSURE	0.02	0.11**	0.11**
	(0.05)	(0.05)	(0.05)
	[0.67]	[0.04]	[0.04]
Constant	0.03	0.02	0.03
	(0.02)	(0.02)	(0.02)
	[0.14]	[0.29]	[0.21]
Observations	75	75	75
	B. 105 Clusters		
EXPOSURE	0.09**	0.12*	0.13*
	(0.04)	(0.06)	(0.07)
	[0.02]	[0.06]	[0.05]
Constant	0.01	0.01	0.01
	(0.02)	(0.02)	(0.02)
	[0.59]	[0.59]	[0.48]
Observations	104	104	104
	C. 165 Clusters		
EXPOSURE	0.10***	0.18***	0.18**
	(0.04)	(0.06)	(0.08)
	[0.01]	[0.01]	[0.02]
Constant	0.01	0.00	0.01
	(0.02)	(0.02)	(0.02)
	[0.58]	[0.79]	[0.59]
Observations	156	156	156
	D. 190 Clusters		
EXPOSURE	0.11**	0.15**	0.19**
	(0.06)	(0.07)	(0.08)
	[0.05]	[0.02]	[0.02]
Constant	0.01	0.00	0.01
	(0.02)	(0.02)	(0.02)
	[0.59]	[0.87]	[0.72]
Observations	176	176	176

OLS results. Exposure metrics are computed from the estimates of Table 2, Column 1, given the average values of the X vector within each cluster. Except for in panel A (where there is 100% "pre/post" matching), the number of observations used in the regression is smaller than the number of clusters because some observations are unmatched. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Table A5: Robustness for the Matched-Samples procedure. (China entry probabilities predicted before clustering.)

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
		A. 75 Clusters	
EXPOSURE	0.02	0.11**	0.10**
	(0.05)	(0.05)	(0.05)
	[0.67]	[0.04]	[0.05]
Constant	0.03	0.02	0.02
	(0.02)	(0.02)	(0.02)
	[0.14]	[0.35]	[0.26]
Observations	75	75	75
		B. 105 Clusters	
EXPOSURE	0.05	0.11*	0.13*
	(0.05)	(0.06)	(0.07)
	[0.34]	[0.08]	[0.06]
Constant	0.01	0.01	0.01
	(0.02)	(0.02)	(0.02)
	[0.42]	[0.66]	[0.55]
Observations	104	104	104
		C. 165 Clusters	
EXPOSURE	0.12***	0.18***	0.19**
	(0.04)	(0.06)	(0.08)
	[0.00]	[0.00]	[0.01]
Constant	0.01	0.00	0.01
	(0.02)	(0.02)	(0.02)
	[0.62]	[0.97]	[0.73]
Observations	156	156	156
		B. 190 Clusters	
EXPOSURE	0.11**	0.15**	0.19**
	(0.06)	(0.07)	(0.08)
	[0.05]	[0.02]	[0.02]
Constant	0.01	0.00	0.00
	(0.02)	(0.02)	(0.02)
	[0.59]	[0.97]	[0.81]
Observations	176	176	176

OLS results. Exposure metrics are computed from the estimates of Table 2, Column 1, for each film, then averaged within each cluster. Except for in panel A (where there is 100% "pre/post" matching), the number of observations used in the regression is smaller than the number of clusters because some observations are unmatched. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

E. Endogenous Design Characteristics

E.1 Observable Characteristics

To further investigate the "redirection hypothesis" we aggregate the number of films released each year, at different shock exposure levels. We label the resulting variable as NFILMS and perform the aggregation in two ways. First we aggregate within each (DEXPOSURE,t) cell, and then within (q, t) cells, where q represents deciles of each of the continuous exposure metrics among the distribution of films released prior to the new policy's announcement. The first aggregation procedure leaves us with 14 observations, whereas the second, with 70. For each approach, respectively, we estimate the following specifications:

$$\begin{aligned} \mathbb{E}[\text{NFILMS}|\text{DEXPOSURE}, t] &= f(\alpha + \beta\text{DEXPOSURE} \times \text{POST}_t \\ &\quad +\gamma\text{DEXPOSURE} + \lambda_t) \\ \\ \mathbb{E}[\text{NFILMS}|q, t] &= f(\alpha + \beta\text{EXPOSURE}_q \times \text{POST}_t + \delta_q + \lambda_t) \end{aligned}$$

Where, as in our main specification, POST=$\mathbf{1}[t \geq 2014]$ and λ are release-year fixed effects. f represents the functional form of the Poisson count-data model.[1] In the second specification, δ represents a fixed effect for each exposure decile. In this specification, we implement EXPOSURE_q as the mean of EXPOSURE metrics within each decile. Qualitative results don't change if we instead use the maximum or minimum.

As before, the parameter of interest is β. A positive estimate thereof would suggest that the 2012 policy change translated into a relative increase in the number of films with characteristics X associated with more likely entry into the Chinese market. Because identification arguments described in the main text apply directly to this specification, they are omitted here. Similar specifications have been used by research exploring the impacts of market expansion on technological innovation activity in the context of natural experiments (e.g., Blume-Kohout and Sood, 2013; Dranove et al., 2017; Hermosilla and Wu, 2016). Estimation results are presented by Table A6. The lack of statistical significance of β estimates invariably supports the graphical result, suggesting that the 2012 policy change did not fuel innovation of the types of films that cater to the Chinese market.

[1] The data do not reject the Poisson assumptions. Results from analog Negative Binomial specifications do not change the qualitative results.

Table A6: Impacts on Innovative Activity.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
EXPOSURE×POST	-0.02	-0.05	-0.07
	(0.05)	(0.09)	(0.25)
	[0.75]	[0.58]	[0.78]
Observations	14	70	70

Procedures for data aggregation are described in text. Estimated models include release-year fixed effects as well fixed effects for exposure categories. Exposure metrics are computed from the estimates of Table 2, Column 1. POST is implemented as $\mathbf{1} = [t \geq 2014]$. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

E.2 Unobservable Characteristics

Table A7 presents the validation results mentioned in text, for both metrics. Table A8 replicates our main results presented in text using the TF-IDF approach.

Table A7: Natural-Language Processing SIMILARITY and Differences in Observable Characteristics.

	(1) Bag-of-Words	(2) TF-IDF
	A. Summary Plots	
$\|\|X^G\|\|$	-0.0074***	-0.0042***
	(0.00)	(0.00)
	[0.00]	[0.00]
$\|\|X^C\|\|$	-0.0002***	-0.0003***
	(0.00)	(0.00)
	[0.00]	[0.00]
Constant	0.0571***	0.0292***
	(0.00)	(0.00)
	[0.00]	[0.00]
Observations	2,748,340	2,748,340
	B. Synopsis	
$\|\|X^G\|\|$	-0.0031***	-0.0023***
	(0.00)	(0.00)
	[0.00]	[0.00]
$\|\|X^C\|\|$	-0.0013***	-0.0006***
	(0.00)	(0.00)
	[0.00]	[0.00]
Constant	0.0814***	0.0226***
	(0.00)	(0.00)
	[0.00]	[0.00]
Observations	774,390	774,390

The dependent variable is SIMILARITY. $||X_{ij}^G||$ and $||X_{ij}^C||$ are Euclidean norms of vectors of genre and sensitive content indicators, respectively. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table A8: Storyline Similarity (TF-IDF SIMILARITY).

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
	A. Summary Plots		
EXPOSURE × CROSSPER	-0.000	-0.001	-0.001
	(0.00)	(0.00)	(0.00)
	[0.96]	[0.66]	[0.76]
Constant	0.027***	0.027***	0.027***
	(0.00)	(0.00)	(0.00)
	[0.00]	[0.00]	[0.00]
Observations	23,177	23,177	23,177
	B. Synopsis		
EXPOSURE × CROSSPER	-0.003	-0.001	-0.001
	(0.01)	(0.01)	(0.01)
	[0.69]	[0.95]	[0.97]
Constant	0.024***	0.024***	0.024***
	(0.00)	(0.00)	(0.00)
	[0.00]	[0.00]	[0.00]
Observations	5,201	5,201	5,201

The dependent variable is SIMILARITY. EXPOSURE metrics are computed with the estimates of Table 2, Column 1. Estimated models include cluster-specific fixed effects. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

F. Starmeter Data

Table A9: Films with larger budgets are more likely to include more famous actors.

Films' budget percentiles	Shares of actors with top 10% popularity	top 1% popularity
0-75	0.04	0.00
	(0.21)	(0.05)
76-90	0.20	0.02
	(0.40)	(0.14)
91-100	0.39	0.07
	(0.49)	(0.27)

Budget levels are computed according to the P75BUDGET and P90BUDGET variables introduced in Section 3. Actors' popularity is measured using IMDB's "Starmeter" variable.

G. Composition

We first provide a falsification test for the Probit results of Table 11 (Panel) A. These suggest that the participation of light-skin actresses increased as a consequence of the policy change. Because once we focus on female roles we are left with few observations from voice roles and animation films, we focus on the test that falsifies the date of the policy change. As in our main analysis, we drop films released after 2012, and assume a that the policy change impacted films released in 2011 and 2012. That is, in this case we implement POST=$\mathbf{1}[t \geq 2011]$. Results are presented in Table A10. These support the causal interpretation of the estimates presented in the text.

Table A10: Female/Male Starring Participation.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
EXPOSURE×POST	0.24	0.48	1.64
	(0.22)	(0.37)	(1.27)
	[0.27]	[0.20]	[0.19]
Observations	1,867	1,867	1,867

Probit results. The dependent variable is $y_r^1 = \mathbf{1}$[Individual playing female role r belongs to skin color category $k = 1$]. Exposure metrics are computed from the estimates of Table 2, Column 1. POST is implemented as $\mathbf{1} = [t \geq 2011]$. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

We now turn to providing evidence suggesting that the 2012 Chinese policy change did not alter the female/male composition among starring roles. To this end, we compute the following variable:

$$\text{SHAREFEMALEROLES}_i = \frac{\text{Number of female actors in non-voice roles in film } i}{\text{Total (female+male) actors in non-voice roles in film } i}$$

We use this as a dependent variable in a regression using the specification of equation (1). Results are presented in Table A11. These suggest that the policy change did not impact the male/female composition among starring roles.

Table A11: Female/Male Starring Participation.

	(1) DEXPOSURE	(2) CEXPOSURE1	(3) CEXPOSURE2
EXPOSURE×POST	-0.01	-0.00	0.04
	(0.03)	(0.04)	(0.09)
	[0.70]	[1.00]	[0.67]
Observations	3,268	3,268	3,268

The dependent variable is SHAREFEMALEROLES. Exposure metrics are computed from the estimates of Table 2, Column 1. POST is implemented as $\mathbf{1} = [t \geq 2014]$. Robust standard errors are presented in parenthesis and p-values in brackets. Legend: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

References

Blume-Kohout, M. E. and N. Sood (2013). Market size and innovation: Effects of medicare part d on pharmaceutical research and development. *Journal of public economics 97*, 327–336.

Dranove, D., C. Garthwaite, and M. Hermosilla (2017). Market conditions and the nature of innovation: Evidence from the biotechnology sector. *Working Paper*.

Hermosilla, M. and Y. Wu (2016). Market size and innovation: The intermediary role of technology licensing. *Working Paper*.