Exposure to foreign media and changes in cultural traits: Evidence from naming patterns in France *

Anne-Célia Disdier[†] Keith Head[‡] Thierry Mayer[§]

January 13, 2009

Abstract

Free trade in audio-visual services has faced opposition on the grounds that foreign media undermine domestic culture, and ultimately, global diversity. Using a long panel of French birth registries, we assess the media-culture link using name frequencies as a measure of tastes. Controlling for the number of people who currently have a name and unobserved name effects, our regressions show that media influences choices via selective imitation. Parents are much more likely to adopt media names that they associate with youth. Using estimated parameters, we simulate our model of name choice to reveal that, absent foreign media, less than 5% of French babies would have been named differently. Our simulations also suggest a positive effect of foreign media on the welfare of parents.

JEL classification: F15, D19, Z10
Keywords: Endogenous Tastes, Cultural transmission, Television, Cinema, Popular Music

*The research was initiated while Head was visiting Paris-Jourdan Sciences Economiques (PSE). We thank David Figlio, Tito Boeri and participants at ERWIT 2005, the LSE-EOPP seminar, the UBC SBE seminar, the CORE Economic Theory seminar, and the IZA/PSE Workshop on Cultural Economics for helpful advice. We especially appreciate the suggestions of Jonathan Eaton and an anonymous referee.

[†]INRA-INAPG, UMR Economie Publique INRA-AgroParistech, anne-celia.disdier@agroparistech.fr
[‡]Corresponding Author: Sauder School of Business, University of British Columbia, keith.head@ubc.ca
[§]Paris School of Economics (Univ. Paris 1), CEPII and CEPR, tmayer@univ-paris1.fr

"Nearly every country in the world is grappling with the question of how to maintain its cultural identity at a time when 'global culture' is washing over the earth." Sheila Copps, 1997, as Minister of Canadian Heritage

1 Introduction

Following the GATT's success in reducing trade barriers on industrial goods, emphasis in multilateral negotiations has shifted to areas, like agriculture and services, where future progress faces severe political obstacles. One of the most contentious issues relates to liberalization of trade in cultural goods and services. On the one hand, countries such as the United States would like to see television programs and films subjected to the same requirements for national treatment and non-discrimination as standard commodities. Opposing this, countries such as France and Canada have advocated a "cultural exception." For example, with strong French and Canadian support, but against US opposition, a 2005 UNESCO conference overwhelmingly approved a new Convention on cultural diversity that asserted the right of a nation to provide public financial assistance to protect cultural diversity within its territory.[1] Article 8 of the United Nations *Universal Declaration on Cultural Diversity* upholds the Franco-Canadian view: "cultural goods and services which, as vectors of identity, values and meaning, must not be treated as mere commodities or consumer goods."[2]

Cultural exceptions might be dismissed as just another form of protectionism. However, as pointed out by Mas-Colell (1999), cultural goods seem to have some distinguishing attributes. Unlike typical goods, individuals not only know what they prefer, they also have preferences over the preferences of others. Bisin and Verdier (2001) emphasize that parents exert effort to pass their own cultural traits on to their children. A growing literature finds that the standard presumption for free trade may not apply for cultural goods. Francois and van Ypersele (2002) show that losses from trade can occur in a model where the cultural good is characterized by fixed costs and heterogenous valuations. Bala and Van Long (2005) model the evolution of preferences using replicator dynamics and show that a large country's preferences can extinguish the preferences of its smaller trading partner.

Three recent papers explore the relationship between culture and trade in models where individuals derive utility from adhering to a cultural identity. Janeba (2007) shows that, because cultural identity is like a network externality, it is possible for trade liberalization to lower welfare. Rauch and Trindade (forthcoming) extend the consumption externalities approach to consider innovation in cultural goods. They argue that "by preserving cultural diversity, protection of cultural goods production can generate dynamic welfare gains that offset the static welfare losses it causes." Olivier et al. (2008) consider the dynamic evolution of cultural identity and find that the opportunity to trade cultural goods leads each country to move towards different mono-cultures. Their model is not designed to generate aggregate

[1] See Article 6 at `http://portal.unesco.org/culture/en/ev.php-URL_ID=28182&URL_DO=DO_TOPIC&URL_SECTION=201.html`. Other examples of cultural exceptions include (1) France's insistence during the Uruguay Round negotiations that the WTO should not apply its trade rules to audio-visual services (`http://www.culture.gouv.fr/culture/actualites/politique/diversite/wto-en.htm`), (2) a proposed EU constitution that explicitly authorized subsidies and protection schemes for cultural industries, and (3) Canada's pursuit of a "general exemption for culture" in its trade agreements (`http://www.international.gc.ca/trade-agreements-accords-commerciaux/agr-acc/ftaa-zlea/C-PandP.aspx`).

[2] `http://unesdoc.unesco.org/images/0012/001271/127160m.pdf`, p. 12.

welfare conclusions, but it does point towards a tension within societies as trade leads to the displacement of one of the autarky cultural identities.

The emerging theoretical literature on culture and trade motivates the need for empirical evidence on this relationship. While Guiso et al. (2006) present a range of evidence that cultural similarity stimulates economic exchange, there is almost no econometric evidence that international trade affects culture.[3] Instead, the notion that foreign cultural goods transform domestic tastes, thereby undermining cultural diversity, seems to be based entirely on casual observation.

This paper brings statistical evidence to the culture and trade debate by examining whether media exposure—of which imports of audio-visual services constitute a growing share—change parental choices for the names of their babies. We estimate that the popularity of a first name in France increases by about 10% when a performer or character with that name appears on one of three main media (movies, television, and songs). Foreign media exert an uneven influence on naming patterns in France. In counterfactual simulations that completely remove foreign media, over 95% of children receive the same names. However, because media exposure is estimated to have stronger impacts on names that have only recently come into use, a subset of names receive a substantial boost in the simulations.

Names have some useful advantages as measures of cultural traits. First, they are consistently and carefully measured (being recorded for virtually everyone by birth registries) over time. Other traits, such as clothing styles or religious beliefs, tend to be difficult to quantify or poorly measured. Second, names are freely available and firms have no profit motive to influence their popularity. This contrasts with, for example, toys, where makers consciously attempt to raise demand via pricing and advertising strategies. Most importantly, there is evidence that names given to children are expressions of cultural identity. For example, Fryer and Levitt (2004) observe that the rapid growth in the use of distinctively Black names might be attributable to a desire by Blacks to "accentuate and affirm Black culture." They invoke the Akerlof and Kranton (2000) model where following identity-appropriate norms of behaviour raises utility.[4]

Our paper proceeds as follows. The next section describes the name data, French regulation of name choice, and trends in naming practices. Section 3 proposes a model of name selection. Section 4 presents our econometric results. Section 5 simulates a counterfactual name distribution in the absence of foreign media. We conclude by reconsidering the merits of a "cultural exception" for trade in audio-visual services in light of our results.

2 Naming regulations and regularities

We start by describing the nature and characteristics of our naming data, which guide our choices on the construction of the dependent variable, and other general issues, such as sample duration, and various approaches to endogeneity concerns.

[3] A very recent draft by Maystre et al. (2008) shows that bilateral trade in goods affects the similarity in responses to 12 questions related to intergenerational transmissions of values from parents to children in the World Values Survey.

[4] The choice of a distinctively Black name appears to be costly: Bertrand and Mullainathan (2004) find that employers are less likely to respond positively to (fake) job applicants whose resumes use Black names. Figlio (2005) finds that teachers are less likely to refer Black-named students to a gifted program.

The data on name frequencies were collected by the French national statistical agency, INSEE, using birth registries. The data set provides the number of babies born in France by name, sex, and year from 1900 to 2002. The panel includes several thousand names each year—every name that was given at least three times. INSEE codes names given to two or fewer children as "rare." The variable we wish to explain is the share of children of a given sex who receive name k in year t. We use the subscript k to denote a name-sex combination, implying that "Camille" is considered a different name when given to a boy from when it is given to a girl. Furthermore, the data set defines names as distinct spellings (not sounds), meaning that "Camille" and "Camylle" are treated as different names.

Until 1993 French parents chose names for their children subject to regulations that date back to 1803.[5] Napoleonic legislation permitted names drawn from the following set: Saints in French calendars, historical figures from ancient Greece and Rome, and Biblical names. The civil registrars charged with enforcing the law were given the discretion to allow some regional and foreign names as well as some spelling variations. If the registrars refused to register a name, parents would have to appeal this decision in court. A ministerial directive in 1966 urged registrars to show greater tolerance for new names, including foreign names. Using "prudence," the officials might accept some diminutives (Ginette for Geneviève), contractions (Marianne for Marie-Anne), and spelling variations (Magdeleine for Madeleine). Legislation on January 8, 1993 dramatically shifted the rules. Now parents can choose any name and register it immediately. If the civil registrars deem a name to be contrary to the interest of the child, they can challenge it in court.

In our regression analysis and simulations, we use only the period where regulations did not strictly constrain the choice of names. We consider both 1967–2002 and 1993–2002 time spans. The former has the advantage of length and therefore more variation in media exposure. The latter permits an analysis with almost no government-imposed constraints on the choices.

We now turn to distinctive patterns of our data that help guide our analysis, in particular by pointing out trends and determinants in naming behavior, and potential endogeneity issues. Figure 1 shows the decline of traditional names, the steady rise of "rare" names, and the rise in American names starting in the 1970s. To define the set of traditional names in France, we made use of the Napoleonic legislation, which explicitly authorized the typical French spellings of the names of Saints from official calendars. Parents have been gradually moving away from Saint names. In 1946 almost three quarters of children received Saint names (down from 86% in 1900). The Saint share had a post-War revival and reached a local maximum in 1964, three years before the ministerial directive that loosened restrictions on names. By 2002, the Saint share had declined to 41%. The share of "rare" names (those given to fewer than three children in a year) has risen steadily from less than one percent in 1946 to six percent in 2002.

The pattern observed for French usage of common names in the US defies simple explanations. In 1946 almost 60% of French babies received names that were also among the top 1000 US names. This reflected names that have long been widely used in both countries such as Daniel, Robert, Marie, and Nicole. Even more stereotypical French names, like Pierre, are included in the US top 1000. However, all the names cited above experience dramatic declines in the post-war period. The rise in French usage of top-1000 US names beginning in 1971 draws mainly from a new set of names (Kevin, Thomas, and Laura are examples of top-ranked names in France during this period).

[5] See `http://www.babyfrance.com/prenoms/legislation.php` for more detail (in French).

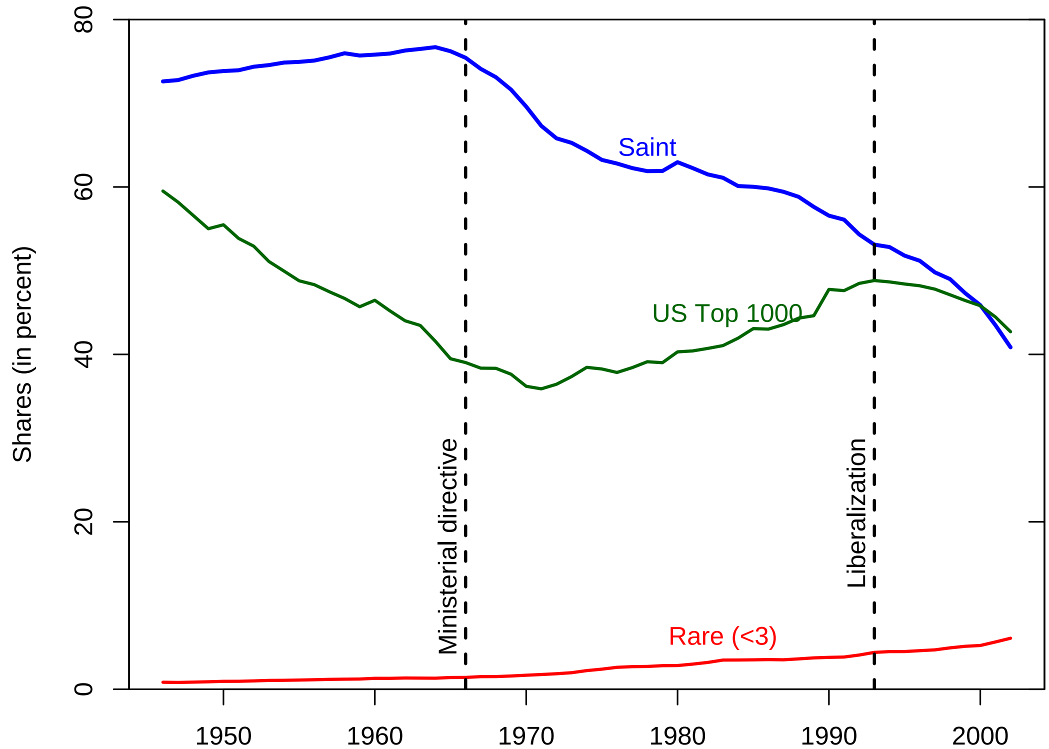

Figure 1: The decline of traditional names

Since the decline of traditional names (Saints) and the rise of alternative sets (US and rare) coincides with broader penetration of television and foreign media more generally, it is tempting to link these trends. It hardly needs to be pointed out that other social factors could contribute to these trends such as declining church attendance, non-Catholic immigration, rising tourism, and foreign-language education. Since identification from aggregate time-series data is doomed to be unconvincing, our approach exploits the name-level variation in media exposure.

Even using this additional dimension of the data, positive associations between media exposure and contemporaneous name popularity could arise for non-causal reasons. One issue is what French sociologists Besnard and Desplanques (2004) refer to as the "illusions of coincidences." For example, Brigitte was the number one name in 1959 (ending Marie's reign of at least 58 years), three years after the release of *And God Created Woman* starring Brigitte Bardot. Kevin was the number one name for French boys in 1990, the same year as Kevin Costner starred in the Oscar-winning *Dances With Wolves*. Many assume that Bardot and Costner were responsible for the popularity of the names Brigitte and Kevin in France. However, our data show that use of these names began to rise *before* the actors in question had released any movies. We respond to the concern over coincidences by using a large panel of names and years in which only a minority of the names were exposed to media in any given year. This allows us to test whether the media-treated names were significantly more popular than the control set.

The examples above relate to actors whose names were chose by their parents almost two decades before their screen careers began. Lieberson (2000) points out that the writers

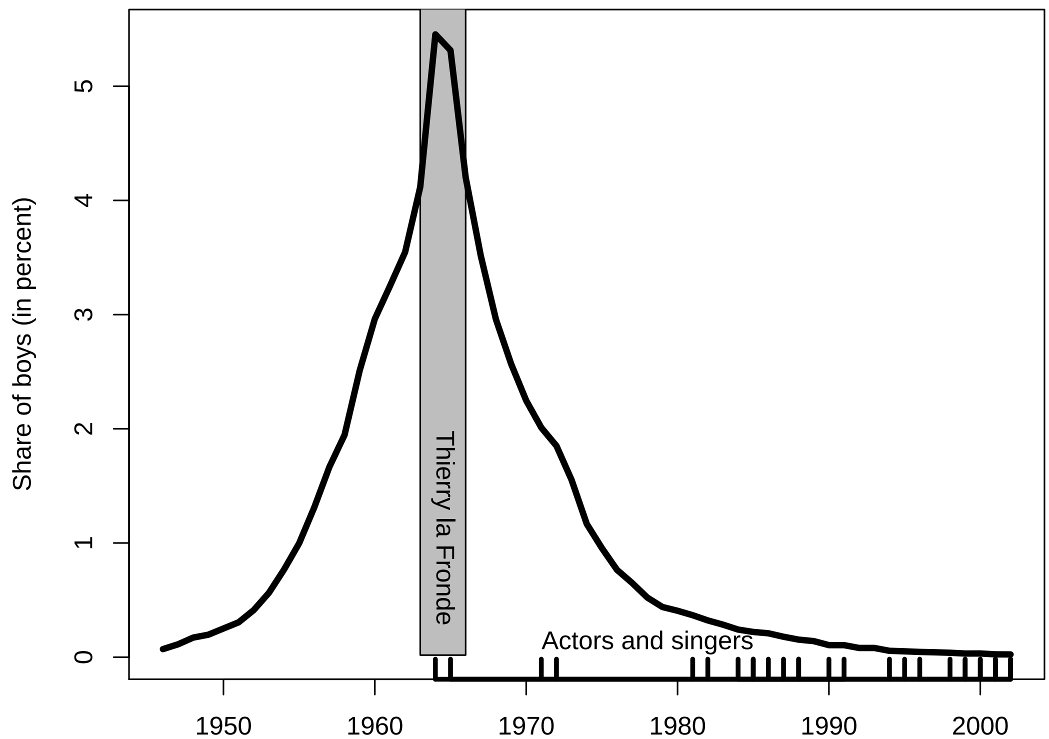

Figure 2: The rise and fall of "Thierry"

creating character names and the actors adopting stage names select names based in large part on their perceived associations. This implies that associations in the public mind can determine media name exposure, rather than vice-versa. Put more generally, media name exposure is endogenous and responds to shocks that affect popularity of names with parents, leading to inconsistent OLS estimates of the causal effect of media exposure.

We use a case study of the name Thierry to illustrate the potential for endogenous media exposure. Many French people attribute the rise of the name "Thierry" to the show *Thierry la Fronde*. As shown in Figure 2, the name peaked in popularity while the show was being broadcast (on the sole French station at the time, ORTF). The figure makes it clear, however, that Thierry became a popular name well before the TV show was broadcast. Thus, it may have been a common shock to tastes affecting *both parents and writers* that lead to the name being chosen for the protagonist of the show.

Figure 2 also illustrates the endogeneity of actor and singer names. The tick marks along the bottom of the figure show years in which we observe a Thierry performing in one of the three media (cinema, TV, or radio). Given the popularity of Thierry as a baby name in the 1960s, it is not surprising that actors with that name become common in the 1980s and 1990s.

Reverse causation wherein popular birth names affect the set of actor names usually occurs with long lags. On the other hand, there can be feedback in the short-run from a name being seen as desirable by parents for their children to the name being seen as appealing for script-writers for their characters. An intermediate case occurs when a performer adopts a "stage" name. For example, the singer Catherine Bodet changed her surname to Lara at some point prior to her 1972 debut album. She enters our media data set at least a decade after changing her name when two of her songs reach the Top 100 in France in 1984 and 1986.

We have two types of solutions to these endogeneity concerns. The first involves a set of controls for characteristics of names (some of which are observable, while others are not) that simultaneously determine their attractiveness to parents and writers. To remove the feedback from past shocks in name popularity to current media exposure, our regressions control for what we call "social exposure," an estimate of the size of the French population with a given name. In addition, our preferred specifications use name-level fixed effects to control for unobserved characteristics of names that remain constant over the estimation period. This specification identifies media effects via *within* co-variation in name popularity and media exposure.

Our second approach to endogeneity issues is to identify sets of names for which simultaneity is likely to be minimal. While we cannot rule out (even with the set of controls described) the influence of contemporaneous shocks affecting writers and parents in general, we argue that this simultaneity issue is much less of a concern with respect to actors than roles. This is because actors and singers generally retain the same stage names throughout their careers and many actors (Brigitte Bardot and Catherine Deneuve, for example) use their birth names. Thus, if one can control for past popularity of a name, the current appearance in the media of an actor with that name should have a causal effect on parent choices. The simultaneity bias therefore predicts that role names should have larger estimated coefficients than actor names (after controlling for social exposure).

The difference between foreign and domestic media can also be useful in this context. While both domestic and foreign screen writers can invent new character names, the actors in domestic productions are much more likely to have traditional French names. Also, if authors choose names for their characters based on current popularity, they should do so based on the frequency of a name in their domestic market. With foreign media, therefore, the simultaneity bias between writer and parent name choices is expected to be of minimal importance. Our regressions will therefore distinguish media effects from performers (as opposed to characters), and foreign media (as opposed to domestic). In both cases, and after controlling for social exposure, we expect the endogeneity bias to be small.

Figure 3 illustrates the type of relationship that one would expect if media indeed has a true causal effect on naming patterns. The figure considers the influence of an American television show that was very popular in France, *Beverly Hills 90210.* This show ran in the US from 1990–2000. Of the four main characters, Brandon, Brenda, and Dylan rose in popularity immediately after the show was released in France in 1993. In contrast, the frequency of Kelly hardly changed. Kelly had already grown before—part of her rise seems attributable to the release of an earlier show, *Santa Barbara*, in 1985. Names such as Brandon or Dylan sound very American to French ears and have been typical examples presented by people arguing that the influence of foreign media on French culture was becoming excessive.[6] Indeed, Dylan climbed up to sixth position in 1996.

We find these illustrations of the possibility of media-enhanced name diffusion intriguing, but hardly convincing. Even if a media figure were found that appeared with exactly the right timing to explain the surge in a particular name's popularity, this could arise because of non-random selection, or "data-mining." This is why we need more rigorous regression analysis,

[6] These names were rising just after American-sounding "Kevin" became the number one name in France. Although French people tend to view these names as American, Dylan, Kelly, and Kevin are actually traditional Welsh and Irish names.

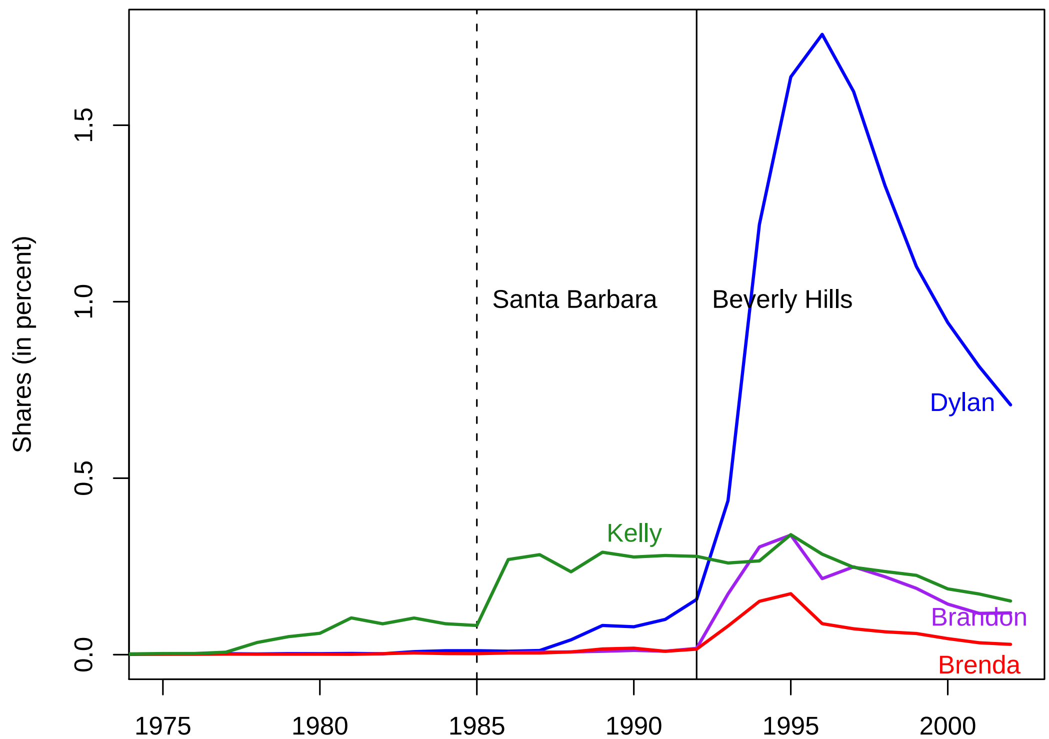

Figure 3: The rise of names originating in the TV show “Beverly Hills, 90210”

using the full sample of names given in the country over a large number of years, combined with information on which of those names where actually media-exposed and when. The next section presents the framework that guides our estimation.

3 Empirical model of name choice

This section develops an empirical framework that incorporates media effects within a broader model of name choice. It is designed to permit estimation of name-choice parameters, so that the model can then be used to simulate counterfactuals.

A continuum of parents, denoted i, select names for babies born in year t from a common choice set $\mathcal{C}_t$. Utility from name k depends additively on commonly perceived attributes v_{kt} and on an idiosyncratic parent-preference term denoted $\epsilon_{kt}(i)$:

$$U_{kt}(i) = v_{kt} + \epsilon_{kt}(i). \tag{1}$$

To obtain a closed-form share formula, it is necessary to assume that $\epsilon_{kt}(i)$ is distributed as a type-I extreme value. For generality, we specify the distribution as $F(x) = \exp(-\exp(-(x - \mu)/\sigma))$, where σ is a scale parameter and μ is a location parameter. The probability a parent chooses name k in year t, $\mathbb{P}_{kt}$, is given by the logit formula:

$$\mathbb{P}_{kt} = \frac{\exp(v_{kt}/\sigma)}{\sum_{j \in \mathcal{C}_t} \exp(v_{jt}/\sigma)}. \tag{2}$$

Because the number of births per year is large enough to use the continuum as an approximation, the probability can be measured as $\mathbb{P}_{kt} = n_{kt}/n_t$, where n_{kt} is the number of name-k births in year t and n_t is total births in that year.

For the econometric specification and welfare analysis, it is useful to introduce a variable referred to in the literature as the inclusive value or "log-sum term":

$$V_t \equiv \ln\left[\sum_{j\in\mathcal{C}_t} \exp(v_{jt}/\sigma)\right]. \tag{3}$$

This notation and the continuum assumption allow us to express name shares as

$$n_{kt}/n_t = \mathbb{P}_{kt} = \exp(v_{kt}/\sigma - V_t) \tag{4}$$

Welfare of the representative parent is given by the expected value of the maximum utility ($U_{kt}(i)$ of the preferred choice). Using different notation, Anderson et al. (1992, pp. 58–61) show that

$$\mathbb{E}[\max_k\{U_{kt}(i)\}] = \sigma V_t + m, \tag{5}$$

where $m \equiv \mathbb{E}[\epsilon_{kt}(i)] = \sigma\gamma + \mu$ is the mean of the parent-specific idiosyncratic utility for a name.[7] Given that we do not know σ, the scale for utility, we cannot express welfare in meaningful units. Also, without arbitrarily restricting $m = 0$, we cannot even calculate percentage changes in welfare. However, since $\sigma > 0$, we can infer the sign of welfare changes induced by policy experiments from the sign of changes in V_t. Moreover, we can compare the magnitudes of welfare changes across experiments.

Taking logs of equation 4, the log share of children given name k is given by

$$\ln(n_{kt}/n_t) = v_{kt}/\sigma - V_t. \tag{6}$$

The next step is to specify the determinants of v_{kt}/σ, the component of the utility of each name that derives from common attributes of name k in year t. The specification should be as simple as possible—to facilitate interpretation and the simulation of counterfactuals—but it should also capture the principal influences on naming decisions.[8] The common utility of a name is a function of three observables (discussed below) and unobserved name attractiveness encompassing both fixed (u_k) and time-varying (e_{kt}) components:

$$v_{kt}/\sigma = f(M_{kt}, S_{kt}, A_{kt}) + u_k + e_{kt}. \tag{7}$$

The first two determinants of name attractiveness are media and social "exposure" of the name in the current year. Media exposure, M_{kt} counts the number of instance in which name k appears on widely released television shows, movies, and songs. M_{kt} comprises counts of appearances of names in 180 major movies, 927 broadcast TV shows, and 4845 popular songs. We consider the names of the actors and roles for the top three roles in each show or movie. Song exposures occur when a name appears as a word in a Top 100 song title that year or as

[7] γ is Euler's number (≈ 0.577).

[8] The discussion paper (Disdier et al., 2006) considers a wider set of determinants of name choice, drawing hypotheses from the Lieberson (2000). Here we focus a more parsimonious specification that keeps the simulations manageable, while nevertheless capturing the main results of interest.

part of the performer's name.[9] In our regressions and simulations we distinguish foreign media exposures, denoted M^F_{kt}, and performer exposures (names of actors and singers, as opposed to names of characters), denoted M^P_{kt}. To shorten expressions, we suppress this distinction in this section.

Social exposure, S_{kt} is an estimate of the number of living individuals in France in year t who have name k. It is obtained by cumulating births by name since 1900 and applying a death rate to remove probabilistically the names of the deceased. Thus, social exposure evolves according to the following stock-flow formula: $S_{kt} = (1-\delta_{kt})S_{k,t-1}+n_{k,t-1}$. The death rate, δ_{kt}, depends on the name and year in order to allow for higher death rates for names that, on average, pertain to older individuals. More detail on the construction of media and social exposure variables is provided in the Data Appendix.

Media and social exposures enter the utility function in much the same way as Becker and Murphy (1993) model the influence of advertising on product demand. That paper stipulates that advertisements "give favorable notice" to other goods. Similarly, we view media and social exposure as enhancing attractiveness of a name. The most straightforward mechanism through which this would work would be a pure desire to imitate. The specification can also be thought of as a reduced form for more complex processes in which media and social exposure raise awareness of names or associate them with desired characteristics. As the focus of this paper is to estimate the impact of media exposure, while controlling for social exposure, we will not attempt to disentangle the mechanisms through which exposures increase name attractiveness. Salganik et al. (2006) provide laboratory evidence that individual choices of cultural goods are strongly influenced by choices of strangers.[10] Econometric evidence that social exposure influences name choices can be found in Head and Mayer's (2008) finding that geographically and socio-economically proximate districts in France have greater similarity in naming patterns.

Social exposure tends to have a conservative influence on naming patterns. If parents based naming decisions only on social exposures, the distribution of names would tend to remain stable over time. This is inconsistent with the rise of non-traditional names shown in figure 1 and the patterns described in Lieberson (2000), who views names as examples of fashion-motivated behavior. The notion of fashion involves a taste for things that are "current." By selecting against things that were popular in the past, parents signal that they are not "old-fashioned." We formalize this motive by assuming that parents avoid names that are "dated," i.e. statistically linked to age. We therefore associate each name in year t with an estimated age, A_{kt}.[11]

The age of a name is given by the difference between the current year(t) and the weighted average birth year of people given that name in the past (b_{kt}), i.e. $A_{kt} = t - b_{kt}$. For example, the age associated with the name Thierry in 1962 was 4 years ($1962-1958$). Forty years later, the age of Thierry had risen to 37 ($2002-1965$). In the same year "Neo" is an example of a

[9] In the discussion paper, Disdier et al. (2006) we estimated different effects for movies, shows, and songs but did not find systematically important differences.

[10] They study decisions to download songs of unknown bands after listening to samples and observing downloading behavior of other participants.

[11] Carter (2004) reports that a consumer marketing company calculates the likely age of a person with a given name using a system it calls "Monica." They use the age classification for direct marketing purposes since first name information is often available when true age is not. While their algorithm is not publicly available, the description in the article makes it look similar to the approach described below.

"young" name ($A = 1.1$).[12]

In addition to the independent effects of exposure and age, we estimate specifications that include interaction terms between these variables. These interactions allow for *selective imitation* in which individuals are more likely to imitate current (and therefore fashionable) names than dated names. In particular, we expect the marginal effect of media and stock exposures to be decreasing in name age. This hypothesis is analogous to an effect observed in the literature on advertising effects on sales: advertising elasticities are higher for new brands and "decrease during the product life cycle."[13]

We can now specify $f(M_{kt}, S_{kt}, A_{kt})$ so as to obtain an estimable regression equation.

$$\begin{aligned} f(M_{kt}, S_{kt}, A_{kt}) &= \beta_1 \ln(1 + M_{kt}) + \beta_2 \ln(1 + S_{kt}) + \beta_3 \ln(1 + A_{kt}) \\ &\quad +\beta_4 \ln(1 + M_{kt}) \times \ln(1 + A_{kt}) + \beta_5 \ln(1 + S_{kt}) \times \ln(1 + A_{kt}). \end{aligned} \tag{8}$$

The first row comprises the direct effects of media and social exposures and name age. The "log of one plus" functional form was selected because each of these variables are right-skewed (logs) and frequently take on zeros (one plus). The second row contains the interaction terms motivated by our hypothesis that the impacts of media and social exposures are decreasing in the age of the name (we predict $\beta_4 < 0$, $\beta_5 < 0$).

During the 1967–1992 period, naming regulations continued to favor a subset of names considered traditional. Given the decline in usage of Saint names observed since the 1960s, it is not clear whether the old rules were being consistently enforced. Nevertheless, we allow for lingering effects of the French naming rules by including an indicator for Saint names. Substituting equation (8) into (7) and the result into (6) and including the rules indicators, we obtain

$$\begin{aligned} \ln(n_{kt}/n_t) &= \beta_1 \ln(1 + M_{kt}) + \beta_2 \ln(1 + S_{kt}) + \beta_3 \ln(1 + A_{kt}) \\ &\quad +\beta_4 \ln(1 + M_{kt}) \times \ln(1 + A_{kt}) + \beta_5 \ln(1 + S_{kt}) \times \ln(1 + A_{kt}) \\ &\quad \beta_6 \text{Saint}_{kt} - V_t + u_k + e_{kt} \end{aligned} \tag{9}$$

The last three terms in the third row are treated as year effects ($-V_t$), name fixed effects (u_k), and an error term (e_{kt}). Regression standard errors are clustered at the name level to make them robust to correlations between e_{kt} and $e_{kt'}$ for name k. By using year dummies, we do not impose any relationship between our estimates of V_t and the underlying determinants contained in the log sum of $\exp(v_{kt}/\sigma)$ shown in equation (2). There are two reasons why we do not use a non-linear least squares approach to constrain the V_t term to depend on the vector of βs. First, to estimate a fixed effects model with 18,947 name fixed effects, u_k, we need to use the within transformation. This requires us to keep the specification linear in the parameters. Second, v_{kt}/σ depends on the unobserved attributes of names captured in $u_k + e_{kt}$. Non-linear least squares estimation of (9) would not incorporate the unobserved name attributes in the $-V_t$.

[12] It first appeared as a name in France in 2000, the year after the release of the movie *The Matrix*, featuring a protagonist with that name.

[13] See Vakratsas and Ambler's (1999) survey for references.

4 Regression results

This section reports estimates of the parameters of equation (9) for two different sets of names. The first sample comprises all of the names given in France to three or more children in a given year. While this set is the most complete possible given the data, it has the disadvantage of the sample being determined endogenously by the choices made by parents in a given year. To investigate the potential effects of selection bias on the results, we also estimate the model using an exogenous set of names: the contemporaneous 1000 most popular names in the United States. With this choice set, we are able to take into account names given to less than three children using Tobit estimation.

All reported specifications use a pooled sample of male and female names. Thus, the βs we estimate should be seen as averages of the two sexes. The rationale for this is that the sex-specific estimates did not differ from each other in ways that were interesting and therefore did not warrant the additional reporting space. The theory dictates that the V_t be sex-specific (since log-sum term capturing all the alternatives differs for boys and girls) so we estimate the models with interacted sex-year dummies.

4.1 Sample of all non-rare names in France

We estimate the name choice model for two time periods: 1967–2002 and 1993–2002. Names given prior to 1967 were subject to closer regulation and thus may have diverged from the unconstrained maximization assumed in our model. After 1992, name choices appear to be essentially unregulated. Using the information from 1967–1992 has the potential to help estimate the model more precisely but we want to make sure it does not give results that are inconsistent with the final period that is clearly unconstrained by regulation.

Tables 1 and 2 show the results in six different specifications. In each table, we start with the most simple model in which name popularity only depends on media exposure (and the unreported sex-year dummies included in all our regressions). Specifications (2)–(6) distinguish between foreign and domestic-source media as well as between the names of performers (actors, singers) and characters (roles in TV and movies, people named in song titles). Column (3) adds the indicator for whether name k is a Saint name (a proxy for compliance with tradition). Column (4) adds the impact of social exposure (stocks) and fashion (age of a name). Column (5) introduces name-level fixed effects, while column (6) adds the interaction terms intended to capture the selective imitation behavior.

Column (1) shows that the correlation between media exposure and name use is not just a matter of anecdotes and data-mining. Names that are currently exposed on media are systematically more popular than other names. To express the impact of media in a way that is comparable across specifications, all tables report the Media Multiplier (MM) corresponding to the coefficients in that column. The MM is defined as the ratio of the name probability with a single media exposure over the probability with no exposures. Thus for column (1), it is $\exp(2.362 \times \ln(1 + 1)) = 5.14$, suggesting that the first media exposure raises name use by a factor of five. The corresponding estimate for the 1993–2002 sample predicts a smaller, but still massive, four-fold increase.

Column (2) distinguishes foreign and performer media exposures. It reveals that names appearing on foreign media are *prima facie* less influential than those appearing on domestic media. This naive specification does not include any controls. What this result means in this

Table 1: Media effects on name shares: 1967–2002 sample

Specification :	Dependent Variable: ln share of babies named k in year t					
	(1)	(2)	(3)	(4)	(5)	(6)
ln $(1+M_{kt})$	2.362^a	3.580^a	2.921^a	0.442^a	0.107^c	0.741^a
all media	(0.080)	(0.144)	(0.138)	(0.101)	(0.059)	(0.098)
ln $(1+M^F_{kt})$		-1.559^a	-1.095^a	0.210^b	0.069	-0.005
foreign media		(0.144)	(0.139)	(0.098)	(0.062)	(0.046)
ln $(1+M^P_{kt})$		-0.301^b	-0.201	-0.277^a	-0.078	-0.041
media performers		(0.136)	(0.131)	(0.084)	(0.053)	(0.041)
Saint name			1.038^a	0.151^a		
			(0.052)	(0.028)		
ln $(1+S_{kt})$				0.664^a	0.627^a	0.808^a
name stock				(0.006)	(0.007)	(0.007)
ln $(1+A_{kt})$				-0.911^a	-0.850^a	-0.504^a
name age				(0.010)	(0.012)	(0.010)
ln $(1+S_{kt}) \times$ ln $(1+A_{kt})$						-0.201^a
						(0.004)
ln $(1+M_{kt}) \times$ ln $(1+A_{kt})$						-0.201^a
						(0.029)
Media Multiplier (MM):	5.14	3.293	3.084	1.296	1.07	1.102
Fixed effects	none				name-sex (k)	
Observations	227219	227219	227219	227219	227219	227219
R^2	0.084	0.091	0.149	0.730	0.412	0.500

Note: Standard errors (name-sex clustered) in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10%. Sex-year dummies included in all specifications. MM is the ratio of $\mathbb{P}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 1$ to $\tilde{\mathbb{P}}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 0$. For specification (6) MM sets $A_{kt} = 14.8$ (sample mean).

Table 2: Media effects on name shares: 1993–2002 sample

Specification :	Dependent Variable: ln share of babies named k in year t (1)	(2)	(3)	(4)	(5)	(6)
ln $(1+M_{kt})$	2.132^a	3.537^a	2.857^a	0.414^a	0.004	0.400^a
all media	(0.082)	(0.190)	(0.186)	(0.148)	(0.041)	(0.092)
ln $(1+M^F_{kt})$		-1.510^a	-1.001^a	0.291^b	0.090^b	0.039
foreign media		(0.185)	(0.181)	(0.142)	(0.040)	(0.038)
ln $(1+M^P_{kt})$		-0.450^a	-0.325^b	-0.340^a	-0.001	0.019
media performers		(0.168)	(0.165)	(0.114)	(0.036)	(0.032)
Saint name			0.927^a	0.120^a		
			(0.056)	(0.036)		
ln $(1+S_{kt})$				0.629^a	0.391^a	0.491^a
name stock				(0.007)	(0.009)	(0.009)
ln $(1+A_{kt})$				-0.849^a	-0.486^a	-0.315^a
name age				(0.013)	(0.015)	(0.014)
ln $(1+S_{kt}) \times$ ln $(1+A_{kt})$						-0.149^a
						(0.006)
ln $(1+M_{kt}) \times$ ln $(1+A_{kt})$						-0.125^a
						(0.025)
Media Multiplier (MM):	4.384	2.984	2.89	1.288	1.066	1.076
Fixed effects	none				name-sex (k)	
Observations	80990	80990	80990	80990	80990	80990
R^2	0.096	0.105	0.146	0.698	0.161	0.184

Note: Standard errors (name-sex clustered) in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10%. Sex-year dummies included in all specifications. MM is the ratio of $\mathbb{P}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 1$ to $\tilde{\mathbb{P}}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 0$. For specification (6) MM sets $A_{kt} = 15.6$ (sample mean).

specification is probably that the endogeneity problem raised above is much less severe for foreign media. Scriptwriters outside France do not choose names for their characters so as to confirm to current tastes in French naming patterns. This interpretation is reinforced by the results on media exposure of persons as opposed to characters. The latter's names are much more likely to be chosen precisely to match parent's tastes than are the names of performers, and indeed the effect of media exposure is lower for person names. For this specification and all others that distinguish between types of media exposure, we calculate the MM for foreign performers, since we have argued that these exposures are less likely to be influenced by endogenous media names and are therefore closer to a causal effect.

Column (3) shows that a small part of the association between name popularity and media exposure arises because both draw from a common set of traditional names. Once taking into account the positive effect of Saint names, media exposure has a lower influence. Column (4) shows a much more important drop in estimated media effects after accounting for social exposure and fashion motives (name stocks and age). The multiplier falls to about 1.3 in both samples. Column (5) completes the set of controls by taking into account unobserved characteristics of a name through the inclusion of name-sex fixed effects. The media multipliers shrink to 1.07 in both samples and media significance levels become marginal. By contrast, social exposure and fashion motives retain a very high level of significance in column (5). For the 1967–2002 sample, a 10% increase in the age of a name translates into an 8.5% fall in popularity.

Column (6) allows for selective imitation, and reveals that the effects of media and social exposures are highly dependent on the age of the name being exposed. A name with the 1967–2002 sample mean age (14.8 years) has a Media Multiplier of $\exp[(.741 - .005 - .041 - .201 \ln[1 + 14.8]) \ln(1 + 1)] = 1.102$. That is, *a single foreign performer average-age exposure boosts name popularity by 10%*. The impact of new name ($A_{kt} = 0$) is considerably larger: 62%. On the other hand, there is no media stimulus for a 31-year old foreign performer name ($\exp[(0.741-0.004-.041)/0.201]-1 = 31$). The corresponding calculations for the 1993–2002 sample give an average-age multiplier of 1.076 and threshold age of 38. Social exposure exhibits a similar pattern, with positive effects disappearing for names aged $\exp(0.808/0.201) - 1 = 55$ years in the long sample, and $\exp(0.491/0.149) - 1 = 26$ years in the short one. Therefore, both kinds of exposure are strongly affected by fashion, with exposure of "middle-aged" names having small or even negative impacts on popularity.

The preferred specifications of Tables 1 and 2 suggest that media exposure has an effect on tastes that is similar to social exposure in terms of magnitude and sensitivity to fashion. With the full set of controls and the age interactions, one cannot reject the hypothesis that all media exposures have the same impact, regardless of whether they are domestic or foreign, performer or character. This gives us some confidence that our controls have purged the media counts of the endogeneity that was so visible in specifications (2) and (3). The media effects for the average-aged name are not notably lower (8% versus 10%) in the recent sample, which is consistent with the view that name choice was not strongly constrained in the 1967–1993 period.

4.2 Sample of top 1000 names in the US

The sample we have used in the estimations above was selected based on a minimum threshold of popularity. It comprises all names given in France—as long as the name was given more than twice in that year. This is the most comprehensive data available, including around five thousand names for each sex per year. However there are thousands of other possible names (especially when one considers possible alternate spellings) that were not used at all or were given to just one or two babies. For example, the name "Arwen" was rare (or non-existent) in France prior to 2002. The estimations in the previous subsection do not take into account that the name transitioned from rare to non-rare the year after the movie *Fellowship of the Ring* was released (featuring a character named Arwen). Similarly, the estimation is not influenced by names like "Chuck" that appeared repeatedly in media (Berry, the 1950s singer) but were never non-rare in France.

It seems worthwhile to pursue an alternate sample selection procedure that is not predicated on the use of the name in France. In light of our interest in media as a mode of international transmission of cultural traits, we use a sample based on popularity in the United States. This provides a natural way to relate our empirical method to the public policy concern over "invasion" of national culture by American cultural traits, transmitted by what is widely perceived as the world's dominant media industry.

In each year, the sample comprises the 2000 names in the top 1000 for boys and girls in United States.[14] Prior to 1990, the top 1000 rankings in the US were constructed on a decadal basis. Hence, for the 1967–2002 estimation period, the set of names remains constant within each decade. For the 1993–2002 estimation period, we use annual top 1000 rankings from the US to determine the set of names. The data depicted in figure 1 reveal that 36–49% of all French babies were given names in the US top 1000 during the 1967–2002 period. Using this sample frame, Brandon is included in every year because, in the US, Brandon has been a top-1000 boy's name since the 1950s. In the national sample, Brandon only entered the sample in 1986, the first year in which three or more Brandons appear in France. In contrast, Arwen is excluded from this sample in *every year*—even in 2002, the year it was actually non-rare in France—because Arwen never attained a top-1000 ranking in the US.

Our regression specifications follow the same sequence as in Table 1, but now take into account the fact that many of the most popular names in the US were not chosen at all by French parents. More precisely, when we do not observe the name k in the set of non-rare names, it means that this name has been chosen two or fewer times in year t. For any name that is rare ($n_{kt} \leq 2$) in France, we recode $n_{kt} = 2$ and estimate using Tobit to account for censoring. Just over 56% of the observations in this specification are censored in the long sample but censoring falls to 48% in the 1993–2002 sample. Tobit methods were not feasible in the previous sample design since we had no way of selecting a finite set of censored names. Note that we change our dependent variable to the log number of births ($\ln n_{kt}$) in these specifications, instead of shares ($\ln n_{kt}/n_t$), since the statistical censoring occurs on births rather than shares. We maintain consistency with our theoretical framework by having sex-year dummies on the right-hand-side of the equation, which now account for $\ln n_t - V_t$, the total number of births for each year and sex and the inclusive value. Another modification to the prior econometric specification is that we have to incorporate the unobserved name effects, u_k, as random effects in columns (5) and (6) of the Tobit specifications.[15]

A comparison of columns (1) and (2) in Table 3 seems to tell the same story as in Table 1, even though the multiplier of media appearance is much larger with this estimation method.[16]

[14] See data appendix for details on sources.

[15] Tobit does not allow for the within transformation needed to estimate large numbers of fixed effects.

[16] For the top-1000 US name set, we report only Tobit results for the 1967–2002 period. Tables in the

Table 3: Media effects on French use of popular American names: 1967–2002 sample

Specification :	Tobit on censored number of babies named k in year t (1)	(2)	(3)	(4)	(5)	(6)
$\ln(1+M_{kt})$	3.836^a	6.251^a	3.663^a	-0.362^a	0.177^a	0.325^a
all media	(0.045)	(0.126)	(0.116)	(0.047)	(0.035)	(0.056)
$\ln(1+M^F_{kt})$		-2.542^a	-0.865^a	0.843^a	-0.013	0.003
foreign media		(0.109)	(0.100)	(0.040)	(0.032)	(0.030)
$\ln(1+M^P_{kt})$		-0.730^a	-0.423^a	-0.370^a	-0.124^a	-0.071^a
media performers		(0.108)	(0.098)	(0.039)	(0.028)	(0.025)
Saint name			3.835^a	-0.384^a	-0.141^b	2.732^a
			(0.040)	(0.018)	(0.069)	(0.126)
$\ln(1+S_{kt})$				0.948^a	0.931^a	1.165^a
name stock				(0.003)	(0.004)	(0.006)
$\ln(1+A_{kt})$				-0.812^a	-0.848^a	-0.252^a
name age				(0.007)	(0.010)	(0.013)
$\ln(1+S_{kt}) \times \ln(1+A_{kt})$						-0.233^a
						(0.004)
$\ln(1+M_{kt}) \times \ln(1+A_{kt})$						-0.043^a
						(0.015)
Media Multiplier (MM):	14.276	7.883	5.189	1.08	1.028	1.105
Random effects	none				name-sex (k)	
Observations	72457	72457	72457	72457	72457	72457

Note: Standard errors (name-sex clustered) in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10%. Sex-year dummies included in all specifications. MM is the ratio of $\mathbb{P}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 1$ to $\tilde{\mathbb{P}}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 0$. For specification (6) MM sets $A_{kt} = 12.5$ (sample mean).

The naive estimations yield very large media effects, with lower impacts for exposure of foreign media and performers. Interestingly, adding controls lowers the media effect drastically to bring them in line with the ones obtained using linear regressions on the French name sample in Table 1. The final column reveals a surprisingly similar estimate of the media multiplier (for a single foreign character of average age) of 1.105, against 1.102 in Table 1. An important difference is the cutoff age when media exposure ceases to have a positive effect, which is now almost 400 years, implying that media impacts are positive for the whole sample range of ages. This result pertains only to the 1967–2002 sample where the estimated interaction term is very small in absolute value. For the 1993–2002 period, Table 9 in the appendix reveals a cutoff point of $\exp[(0.384 - 0.041 - 0.006)/0.088] - 1 = 45$ years, which is remarkably close to the 38 years obtained for the full set of names in the 1993–2002 sample.

An additional natural control which we now introduce to this sample is the popularity of the name in the United States. This variable is intended to capture transmission of American names through means other than the media we measure—such as tourism or magazines. The results shown in Table 4 suggest an important role for non-media interactions. While the set of American names appearing on foreign media has a large impact on naming patterns in France, the effect comes in part from the popularity of those names in the USA. This can be seen by comparing the first three columns of Tables 3 and 4. The impact of a name's popularity in the US almost translates one-for-one into popularity in France, and the estimated media multiplier is cut dramatically when controlling for $\ln n_{kt}^{\text{US}}$ in specifications (1)–(3). Unsurprisingly, adding the control variables reduces the impact of popularity in the US, although it remains significantly positive, even when name random effects are introduced.[17] In the preferred specification, the MM is hardly changed by controlling for US popularity (1.107 vs 1.105).

The bottom line from the US sample is reassuring: the preferred specification upholds the finding that a single foreign performer exposure boosts name popularity by about 10%. Moreover, the finding of selective imitation whereby media and social exposure have less positive effects as names become "older" seems very robust.

5 Policy experiment: foreign media exclusion

The parameters we have estimated can be plugged back into the logit choice probabilities to determine the share of each name for any setting of the right hand side variables. This allows us to conduct counterfactual exercises in which we manipulate the amount of media exposure. Such a policy change is also realistic: In 1986 the French government introduced quotas for audio-visual services. French law now requires that 60% of the movies and shows on TV be of European origin. Of those, 40% of free-channel programming should be in French. In addition,

appendix show additional results. There we first show the 1993–2002 period, but using just the US names that were non-rare in France, and thus not taking account censoring with Tobit. This allows comparison with Table 2 to see the impact of name universe change holding the regression method unchanged. The main results are quite similar, passing the same significance thresholds. In our preferred regression, column (6), the coefficient on $\ln(1+M_{kt})$ goes from 0.400 to 0.321, and the media-age interaction changes from -0.125 to -0.090. Table 9 then changes estimation to be Tobit and Table 10 adds the popularity in the US as a control. Results diverge very little from the corresponding Tables for 1967–2002 shown in the text. A notable consequence of using Tobit is that the media multiplier becomes much larger—until the controls are introduced.

[17] Table 10 shows that the impact of US popularity is an order of magnitude higher (0.28 vs. 0.03) in the preferred specification for the 1993–2002 sample, suggesting an increase in non-media transmission.

Table 4: Media and non-media effects on French use of popular American names: 1967–2002

Specification :	Tobit on censored number of babies named k in year t (1)	(2)	(3)	(4)	(5)	(6)
$\ln n_{kt}^{\text{US}}$	0.958^a	0.984^a	0.834^a	0.112^a	0.084^a	0.030^a
US popularity	(0.011)	(0.011)	(0.010)	(0.005)	(0.007)	(0.007)
$\ln(1+M_{kt})$	2.628^a	5.560^a	3.375^a	-0.328^a	0.173^a	0.319^a
all media	(0.044)	(0.117)	(0.109)	(0.046)	(0.034)	(0.056)
$\ln(1+M_{kt}^F)$		-3.345^a	-1.747^a	0.686^a	-0.002	0.007
foreign media		(0.102)	(0.094)	(0.040)	(0.032)	(0.030)
$\ln(1+M_{kt}^P)$		-0.653^a	-0.390^a	-0.365^a	-0.120^a	-0.070^a
media performers		(0.100)	(0.092)	(0.039)	(0.028)	(0.025)
Saint name			3.373^a	-0.373^a	-0.153^b	2.665^a
			(0.038)	(0.018)	(0.069)	(0.125)
$\ln(1+S_{kt})$				0.930^a	0.923^a	1.163^a
name stock				(0.003)	(0.004)	(0.006)
$\ln(1+A_{kt})$				-0.799^a	-0.850^a	-0.255^a
name age				(0.007)	(0.010)	(0.013)
$\ln(1+S_{kt}) \times \ln(1+A_{kt})$						-0.231^a
						(0.004)
$\ln(1+M_{kt}) \times \ln(1+A_{kt})$						-0.042^a
						(0.015)
Media Multiplier (MM):	6.181	2.954	2.358	0.995	1.036	1.107
Random effects	none				name-sex (k)	
Observations	72457	72457	72457	72457	72457	72457

Note: Standard errors (name-sex clustered) in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10%. Sex-year dummies included in all specifications. MM is the ratio of $\mathbb{P}_{kt}$ with $M_{kt} = M_{kt}^F = M_{kt}^P = 1$ to $\tilde{\mathbb{P}}_{kt}$ with $M_{kt} = M_{kt}^F = M_{kt}^P = 0$. For specification (6) MM sets $A_{kt} = 12.5$ (sample mean).

the law imposes compulsory investment in the production of European and French-language content. With respect to radio, at least 40% of the songs played should be in French.[18]

Within the context of this study we cannot know what names would have been exposed on French media in the absence of the quota system. We consider a counterfactual move in the opposite direction: the complete removal of foreign media and therefore of all name exposure on foreign-origin movies and shows and foreign-language songs. This experiment is analogous to the autarky policy often studied in the context of trade in goods.

The removal of foreign media has two types of effects in the context of our model. There is a static effect of lowering attractiveness of names that would have been exposed on foreign media and, correspondingly, raising the attractiveness of other names. There is also a dynamic effect because the change in a name's popularity (flows) in year t can affect the stocks and age of names in all subsequent years. The simulation therefore enables rich dynamics: When a young name is exposed in the media, it has an immediate boost in popularity, which may be reinforced over time, because the initial surge raises the stock of people exposing that name socially and lowers the age associated with it, both having positive effects on the desire to adopt this name for one's child.

Our simulation method proceeds as follows.

Step 0: Estimate the coefficients (β), fixed effects (u_k) and residuals (e_{kt}) used in the calculation of name shares using sex-specific versions of the preferred specification (column 6 in Table 1). We parameterize the simulations with the following estimates:

Females:

$$\begin{aligned} f(\cdots) &= .82\ln(1+M_{kt}) + 0.01\ln(1+M^F_{kt}) - 0.14\ln(1+M^P_{kt}) + 0.83\ln(1+S_{kt}) \\ &- 0.50\ln(1+A_{kt}) - 0.21\ln(1+M_{kt})\times\ln(1+A_{kt}) - 0.21\ln(1+S_{kt})\times\ln(1+A_{kt}). \end{aligned} \tag{10}$$

Males:

$$\begin{aligned} f(\cdots) &= 0.69\ln(1+M_{kt}) - 0.01\ln(1+M^F_{kt}) + 0.04\ln(1+M^P_{kt}) + 0.78\ln(1+S_{kt}) \\ &-0.50\ln(1+A_{kt}) - 0.20\ln(1+M_{kt})\times\ln(1+A_{kt}) - 0.19\ln(1+S_{kt})\times\ln(1+A_{kt}). \end{aligned} \tag{11}$$

Step 1: In the first year ($t = 1$ corresponding to 1967 or 1993) of the simulation, we set age and stocks at their actual levels, A_{k1} and S_{k1}. We then determine counterfactual name attractiveness, $\tilde{v}_{k1}/\sigma = f(\tilde{M}^F_{j1} = 0, S_{j1}, A_{j1}) + \hat{u}_j + \hat{e}_{j1}$. This sets the foreign media appearance counts, M^F_{k1}, to zero and also subtracts foreign media counts from all media, M_{k1}, and media performers, M^P_{k1}. The counterfactuals assume that the same fixed effects, $\hat{u}_k$, and residuals $\hat{e}_{kt}$ apply in the absence of foreign media. Calculate the log-sum term, $\tilde{V}_1$, using equation (3). Calculate the counterfactual name shares $n_{kt}/n_t = \tilde{\mathbb{P}}_{k1}$ using the logit formula, equation (4). In $\tilde{\mathbb{P}}_{k1}$, not only the numerator is affected by the zeroing of foreign media. Because the inclusive value, $\tilde{V}_1$, changes in the simulation, the number of predicted births in the counterfactual changes *even for names that did not receive media exposure.* When foreign media are excluded,

[18]See `http://www.csa.fr/infos/controle/controle_intro.php` for more detail on the French quota system. Other countries employ similar quota systems. Canadian content rules require that 60% of broadcast TV programming and 35% of broadcast radio be of Canadian origin. South Korea required movie theaters to show locally-produced firms at least 40% of the year—until the signing of the Korea-US FTA, which lowered the requirement to 20%.

our simulations show that $\tilde{V}_t < V_t$. This increases the share of children given names that did not receive media exposure in the baseline. Taking total births as given, we calculate the flows for each name as $\tilde{n}_{k1} = \tilde{\mathbb{P}}_{k1} n_1$.

Step 2: The first variable to be adjusted in the simulation based on the counterfactual flow ($\tilde{n}_{k1}$) is the age of a name in year 2. Counterfactual birth years are calculated recursively as

$$\tilde{b}_{k,t+1} = t(\tilde{n}_{k,t}/\tilde{N}_{k,t+1}) + \tilde{b}_{k,t}(\tilde{N}_{k,t}/\tilde{N}_{k,t+1}),$$

where N is our notation for (un-depreciated) cumulative births, $\tilde{N}_{k,t} = \tilde{N}_{k,t-1} + \tilde{n}_{k,t-1}$, and $\tilde{b}_{k,1} = b_{k,1}$ for the initial year of the simulation (1967 or 1993). Subtracting from the current year, $t+1$, we obtain age for each name as $\tilde{A}_{kt} = t - \tilde{b}_{kt}$.

Step 3: The counterfactual age calculation for each name and year implies a different set of death rates. The new death rates are given by $\tilde{\delta}_{kt} = d(\tilde{A}_{kt})$ (see Appendix A.1). Counterfactual social exposures are obtained by adding on the simulated births, $\tilde{n}_{kt}$, to the depreciated stock of each name: $\tilde{S}_{k,t+1} = (1 - \tilde{\delta}_{kt})\tilde{S}_{k,t} + \tilde{n}_{k,t}$, where $\tilde{S}_{k1} = S_{k1}$.

Step 4: Calculate (using the formulas above) the next year values of $\tilde{v}_{k,t+1}/\sigma = f(\tilde{M}^F_{kt} = 0, \tilde{S}_{k,t+1}, \tilde{A}_{k,t+1}) + \hat{u}_k + \hat{e}_{k,t+1}$, $\tilde{V}_{t+1}$, $\tilde{\mathbb{P}}_{k,t+1}$, and $\tilde{n}_{k,t+1}$.

Step 5: Repeat steps 1–4 year-by-year until 2002.

We refer to the procedure including steps 1 to 4 as the dynamic version of the simulation. We also conduct a static version of the simulation for comparison purposes. In that case we skip steps 2 and 3. Step 4 in the static simulation calculates $\tilde{v}_{k,t+1}/\sigma = f(\tilde{M}^F_{kt} = 0, S_{k,t+1}, A_{k,t+1}) + \hat{u}_k + \hat{e}_{k,t+1}$. Thus, the static version leaves stocks and age unaffected by the policy changes.

We use two measures to quantify the aggregate effects of media on French parents. The "positive" measure is a calculation of the share of parents in a given year who changed the name of their baby because of media exposure. Define a "stayer" as a child that retains the same name under the baseline and the counterfactual. We calculate this as the lesser of N_{kT} and $\tilde{N}_{kT}$, where N_{kT} and $\tilde{N}_{kT}$ are the cumulated name-k births in the baseline and simulation up until year T (2002) and $N_T = \sum_k N_{kT}$. The change share is therefore given by

$$\% \text{ renamed} = 1 - \frac{\sum_k \min\{N_{kT}, \tilde{N}_{kT}\}}{N_T}.$$

The normative measure is the change in expected utility implied by our counterfactual removal of media exposure. For ease of interpretation we express this as the gain (or loss if negative) in expected utility attributable to media exposure. Equation (5) implies that the media-induced change in expected utility in a given year is given by $\sigma(V_t - \tilde{V}_t)$. Consequently, we measure the total welfare change due to media by accumulating the differences in the inclusive values over all periods included in the simulation:

$$\Delta\text{welfare} = \sum_t (V_t - \tilde{V}_t).$$

Although we cannot interpret the units of this measure (because σ is unknown), it does indicate the sign of the welfare change and the units are comparable across policy experiments for a given gender (σ may differ across sexes).

Table 5: Simulated effects of media on renamings and welfare, 1967–2002

Sex	Media	Dynamics	# renamed	% renamed	Δ welfare
females					
	foreign	no	150634	1.13	.51
		yes	386131	2.89	.27
	all	no	242501	1.81	.88
		yes	595513	4.45	.50
males					
	foreign	no	262906	1.84	.77
		yes	656422	4.61	.45
	all	no	375224	2.63	1.28
		yes	982895	6.9	.68

Table 5 provides the simulated percentage of babies renamed when removing foreign media in the static and dynamic versions of our simulations. Over the 1967–2002 period, about 28 million babies were born in France. Our dynamic simulation predicts that, among those, over a million (386,131 girls and 656,422 boys) would have had a different name without the influence of foreign media. This represents 2.89% of baby girls and 4.61% of baby boys. Table 11 in the appendix presents results when running the same simulations on the 1993–2002 period. The corresponding percentages of babies renamed are 1.68% and 2.36%, reflecting a shorter period over which the dynamic part of the model can produce its effects.

The last column of Table 5 shows the change in welfare ($V_t - \tilde{V}_t$) that the simulations attribute to foreign media. To the extent that media exposure raises the name-level terms in V, i.e. the v_{kt}, it will tend to raise welfare. The coefficients reported in equations (10) and (11) show that media exposure raises v_{kt} so long as the age of name, A_{kt}, is below a critical value—49 years for girls and 31 years for men.[19] Since almost 90% of the foreign-exposed female name-years and over 75% of the male name-years are younger than the critical values, media is mainly enhancing name attractiveness rather than detracting from it. Parents of male children obtain higher welfare benefits despite the lower critical value and this appears to be the result of a higher rate of exposure: 5.9% of the male name-year combinations had positive foreign media exposure whereas only 2.7% of the female name-years did. One puzzling aspect of the results is that the dynamic welfare gains from media are smaller than the static gains, even though the dynamics lead to more name changes. The explanation seems to be that media exposure for a small set of names with relatively large stocks lowered the attractiveness of a larger number of unexposed names that on average had relatively small stocks and young ages. The adjustment in stocks and ages that resulted lowered the subsequent attractiveness of those names, partially offsetting some of the welfare benefits from media.

For the sake of comparison, we also run the simulations for the unlikely policy experiment where all media would be shut down. The dynamic effects entail names changes for 4.45% of

[19] The calculations for these critical values are $\exp(0.82/.21) - 1$ and $\exp(.69/.2) - 1$, respectively. For media exposure of performers, the figures are $\exp(0.68/.21) - 1 = 24$ years for girls and $\exp(.73/.2) - 1 = 37$ years for boys.

girls and 6.9% of boys. In this case, almost 1.6 million babies would be renamed. The higher magnitudes were to be expected since many French names appear only on domestic media. In the average year, 4.3% of female names and 9.3% of male names received media exposure. This probably explains why removing all media would have lowered welfare more than just removing foreign media.

Our first measure of the global impact of foreign media exposure on cultural patterns gives figures ranging between 1 and 5%. Although not negligible, the overall effects are small, even when allowing for 36 years of dynamic media effects through changes in the stocks and age of names. However, for some names the effects revealed by our simulations are really large. We focus here on one specific case: the removal of all foreign media exposure, accounting for dynamic effects, in Table 6. This table lists the 20 names that the simulations assign the greatest positive impact of media exposure. Out of those 20 names, 13 appear to have their popularity more than double. Most of those names sound like typical examples of the "foreign cultural invasion" claims. They are not traditional French names (Britney and Jason have a respective age of 2 and 8 years in 2002), sound American to French ears, and have been heavily media exposed.

The other side of the spectrum is also revealing: the names that actually *suffer* from foreign-media exposure. Table 7 shows that some of the most harmed names have very low media exposure while others are heavily exposed. The highly exposed names that are harmed by media seem to share a common feature—their names convey age. For male names over 31 and female names over 49, our estimates indicate that media exposure has a negative impact on attractiveness. This can explain why names like Paul and Charles—with average ages of 56 and 57 in the 1967–2002 sample—are harmed by foreign media. We expect names that were rarely or never exposed to experience declines in popularity due to heavy media exposure for other names. Thus, it is not surprising that names like Margot, Hugo, Sebastien, and Elodie would have done better in the absence of foreign media. The question is why those names lose so much from the existence of media, while others that received the same low amount of exposure suffer much less. In other words, what accounts for the unequal distribution of dynamics losses? This inequality does not arise in the static simulation where we find that the losses for unexposed names range from two to four percent.

Some investigation revealed a common thread to the unexposed names that were hardest hit in the dynamic simulation. Consider the starting year of the simulation, 1967, when we shut down foreign media. For all U names that are un-exposed in 1967, there is an identical positive percentage rise in popularity, $\tilde{\mathbb{P}}^U_{1967}$, that comes from a fall in the attractiveness of media-exposed names, captured by a decrease in the inclusive values: $\tilde{V}_{1967} < V_{1967}$. In the static version of our simulation, the effect remains identical for those names that are never exposed over the duration of the simulation. For the dynamic case however, things are more complicated: In 1968, stocks of U names are adjusted to account for $\tilde{\mathbb{P}}^U_{1967} > \mathbb{P}^U_{1967}$.[20] Names with a high initial popularity in 1967 experience the same percentage increase as the others, but a higher *absolute* increase in the number of babies born with that name in 1967, which translates into a higher absolute stock increase in 1968. Since stocks enter positively in utility, this will feed into a rise in popularity in 1968, with the resulting percent increase in $\tilde{\mathbb{P}}^U_{1968}$ being all the higher if the stock in 1968 was low. To summarize, unexposed names have a big response

[20]The age of those names also adjusts and the effect goes in the same direction, as those names become a little younger which is good for future popularity. For clarity, we keep our explanation focused on changes in stocks.

Table 6: Names most helped by foreign media, 1967–2002

name	male	years non-rare	years exposed	simulated # births media off	actual # births media on	(on-off)/off
Tia	0	4	4	32	105	224.41
Laura	0	36	17	44100	117913	167.37
Lisa	0	36	23	13206	34098	158.21
Tom	1	35	30	9594	23881	148.92
Jennifer	0	36	28	25721	63831	148.17
David	1	36	35	116501	287812	147.05
Jonathan	1	36	20	40278	98436	144.39
Michael	1	36	35	36247	85245	135.18
Britney	0	4	3	59	136	132.21
Theo	1	36	16	24707	54213	119.43
Shakira	0	1	1	23	49	114.21
Xena	0	6	5	29	61	110.6
Calista	0	4	4	140	292	107.96
Shannen	0	10	10	72	141	96.87
Tasha	0	3	3	5	10	91.83
Alan	1	36	30	7217	13620	88.73
Jason	1	34	23	8121	15314	88.57
Rowan	1	10	10	72	135	86.71
Anastacia	0	3	3	11	20	86.66
Anthony	1	36	25	85452	157938	84.83

Note: Actual and simulated # births are cumulated between 1967 and 2002. The number of births has been rounded to the nearest unit, while the (On-Off)/Off percentage is calculated before the rounding. “Years exposed” counts the number of years that the name was non-rare and appeared on media.

Table 7: Names most harmed by foreign media, 1967–2002

name	male	years non-rare	years exposed	simulated # births media off	actual # births media on	(on-off)/off
Lara	0	36	9	5863	4436	-24.34
Paul	1	36	34	77286	60771	-21.37
Arthur	1	36	24	47887	38102	-20.43
Anais	0	36	1	105454	87381	-17.14
Valentin	1	36	0	79677	66341	-16.74
Sebastien	1	36	2	353398	294439	-16.68
Romain	1	36	3	181912	155590	-14.47
Enzo	1	35	7	28234	24186	-14.34
Corentin	1	35	0	42580	36726	-13.75
Hugo	1	36	1	70404	60759	-13.7
Charles	1	36	33	43184	37304	-13.62
Guillaume	1	36	6	224498	194493	-13.37
Andrea	0	36	12	17207	14920	-13.29
Eva	0	36	20	34753	30222	-13.04
Manon	0	36	3	106110	92527	-12.8
Celeste	0	36	14	1519	1328	-12.57
Margot	0	32	0	23504	20592	-12.39
Lucas	1	36	23	79265	69584	-12.21
Elodie	0	36	1	172341	151465	-12.11
Julie	0	36	30	188168	166214	-11.67

Note: Actual and simulated # births are cumulated between 1967 and 2002. The number of births has been rounded to the nearest unit, while the (On-Off)/Off percentage is calculated before the rounding. "Years exposed" counts the number of years that the name was non-rare and appeared on media.

to the shutting down of media if they have a high flow to stock ratio, that is if they experience a popularity boom over the simulation period. This is the case for Sebastien, which starts as the 89th most popular name in 1967 and reaches number one from 1975 to 1979.[21] As a contrasting case, take Gustave, which has only 16 babies born in 1967, with rank 591, but stays in this range for the whole period, never rising above rank 504. The negative effect of foreign media on Gustave is only -0.98% (as 13 other boys' names), when it is -16.68% for Sebastien.

Last, we also run the same simulations but remove *all* media, rather than only the foreign exposures. While this scenario is unlikely as a real policy, it is nonetheless interesting to see if the list of names most helped and most harmed by media changes a lot or not. The names shown in Tables 12 and 13 in the appendix overlap considerably with those shown in Tables 6 and 7. Foreign media is therefore an important component of the overall effect of media, at least for the extreme gainers and losers. Note also that names like Laura or David—which are frequently used on both French and foreign media—would be considerably more harmed by total media removal simulation (all media generated a 252% gain for Laura while foreign media contributed just 167%).

6 Conclusion

We investigate whether exposure to media in general and foreign-origin media in particular affect naming patterns in France. The names chosen for babies are emblematic characteristics of national cultural traditions. Changes in practices on this subject have been interpreted as one manifestation of globalization, possibly endangering cultural diversity. France has been at the forefront of political activity, arguing for a cultural exception that would allow for government intervention to protect domestic culture. The political discussion of protecting culture tends to obscure whether it is the consumer or the producer that requires protection. If it is the producer, then the old arguments of trade policy imply that it is more efficient to promote domestic production via subsidies than to inhibit import consumption via trade barriers. However, if import consumption has adverse external effects, the case for limiting foreign access could make more sense.

In this paper we offer what we believe to be the first systematic evidence of the impact of foreign media on a cultural trait. Our results show that foreign media have a positive, but complex, influence on naming patterns in France. Our "naive" regression analysis finds very big effects of media exposure on a name's popularity, thus seeming to corroborate anecdotal accounts of media influence. The introduction of controls for attributes that currently lend popularity to a name dramatically lowers the estimated media effect. Our preferred specification maintains those controls and allows for *selective* imitation of names that appear on media: parents adopt media names only if they are sufficiently fashionable, i.e. "young." When a brand new name appears on media, its popularity jumps by 62% compared to an unexposed new name. The effect of media falls to 10% for a name with the mean age in the sample and becomes negative for ages over 31.

Our model of name choice allows for counterfactual analysis, which we use to quantify the total positive and normative effects of media on naming patterns in France. The simulations

[21] As other examples: Hugo rises from rank 409 in 1967 to rank 97 in 1987 and ranks fourth in 2002. Elodie starts at rank 298 in 1967 and reaches rank 1 from 1988 to 1990.

also identify the names that were most helped and most harmed. We find that foreign media changed less than five percent of names. The broader implication from this specific result is that reports of the death of local cultural diversity may be exaggerated. Although the aggregate impact appears modest, we find many examples of non-traditional names for which our simulation attributes recent surges in popularity to foreign media. Perhaps these cases explain the strength of the public concern over cultural invasion channeled through foreign media.

Even if we had found stronger overall foreign media effects, it would not have provided a sufficient justification for barriers to trade in audio-visual services. Just as we normally presume that imports of goods benefit the consumer, parents may benefit from choosing media-exposed names. Our simulations point to welfare gains from both domestic and foreign media. This is because the attractiveness of the most exposed names is estimated to be enhanced by media exposure, leading to a higher expected utility for the name actually chosen. Because we take the choice set as exogenous, our simulations do not capture welfare gains from the introduction of new names to France. Since the logit model builds in a love of variety, it seems likely that endogenizing the choice set would lead to larger welfare increases from foreign media.

References

Akerlof, George A. and Rachel E. Kranton, 2000, "Economics and identity," *Quarterly Journal of Economics* 115, 715–753.

Anderson, S., A. De Palma and J.-F. Thisse, 1992, *Discrete Choice Theory of Product Differentiation* (MIT Press, Cambridge).

Bala, Venkatesh and Ngo Van Long, 2005, "International trade and cultural diversity with preference selection," *European Journal of Political Economy* 21, 143–162.

Becker G. and K. Murphy, 1993, "A Simple Theory of Advertising as a Good or Bad," *The Quarterly Journal of Economics* 108(4), 941–964.

Bertrand, Marianne and Sendhil Mullainathan, 2004, "Are Emily and Greg more Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *The American Economic Review* 94(4), 991–1013.

Besnard, Joséphine and Guy Desplanques, 2004, *La côte des prénoms en 2005* Balland-Paris.

Bisin, Alberto and Thierry Verdier, 2001, "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic Theory* 97, 298–319.

Carter, Helen, 2004, "Names that show their age," *The Guardian* Friday, October 08 Online edition.

Disdier, A.-C., K. Head, and T. Mayer, 2006, "Exposure to foreign media and changes in cultural traits: Evidence from naming patterns in France," CEPR DP 5674.

Figlio, David N., 2005, "Names, Expectations and the Black-White Test Score Gap," NBER Working Paper # 11195.

Francois, Patrick and Tanguy van Ypersele, 2002, "On the protection of cultural goods," *Journal of International Economics* 56, 359-369.

Fryer, Roland G. and Steven D. Levitt, 2004, "The Causes and Consequences of Distinctively Black Names," *Quarterly Journal of Economics* 119(3), 767-805.

Guiso L., P. Sapienza, L. Zingales, 2006, " Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives* 20(2): 23-48.

Head, K. and T. Mayer, 2008, "Detection of local interactions from the spatial pattern of names in France", *Journal of Regional Science* 48(1): 67–95.

Janeba, Eckhard, 2007, "International Trade and Consumption Network Externalities" *European Economic Review* 51(4), 781–803.

Lieberson, Stanley, 2000, *A Matter of Taste: How names, fashions, and culture change*, Yale University Press: New Haven.

Maystre, N., Olivier, J., Thoenig, M. and T. Verdier, 2008, "Product-Based Cultural Change: Is the Village Global?", mimeo.

Mas-Colell, Andreu, 1999, "Should Cultural Goods Be Treated Differently?" *Journal of Cultural Economics* 23, 87–93.

Meslé, France and Jacques Vallin, 2002, "Comment améliorer la précision des tables de mortalité aux grands âges? Le cas de la France", *Population* 57(4-5):603–663.

Olivier, J., M. Thoenig, T. Verdier, 2008, "Globalization and the dynamics of cultural identity," *Journal of International Economics*, 76(2):356–370.

Rauch, James E. and Vitor Trindade, forthcoming, "Neckties in The Tropics: A Model of International Trade and Cultural Diversity" *Canadian Journal of Economics*.

Salganik, Matthew J., Peter Sheridan Dodds, and Duncan J. Watts, 2006, "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market" *Science* 311, 854–856.

Vakratsas, D. and T. Ambler, 1999, "How Advertising Works: What Do We Really Know?" *The Journal of Marketing* 63(1): 26–43.

A Data appendix

A.1 French names

The French statistical agency, INSEE, sells a CD-ROM called the *Fichier des Prénoms* that provides national data based on filings of birth certificates at the Civil Registry. The database includes all babies born in all of France (including the overseas departments Reunion, Guadeloupe, Martinique, and Guiana). This gives us our key variable n_{kt}, the number of births of name k in year t. Particular names are shown if they were given to at least three babies for a given sex and year, i.e. n_{kt} is not observed for $n_{kt} \leq 2$. The total number of births with names given to just one or two babies are summed and reported under the name "rare." We use n_{kt} to calculate $\mathbb{P}_{kt} = n_{kt}/n_t$, the proportion of babies given name k in year t. We also use this variable to estimate the age of name based on the distribution of name-frequencies observed up until time t:

$$A_{kt} = t - b_{kt} = t - \sum_{\tau=0}^{t-1} \frac{\tau n_{k\tau}}{\sum_{j=0}^{t-1} n_{kj}}.$$

Using data on the age distribution of the French population from 1900–1998, we calculate the average death rate as function of age.[22] Let $S(a)_t$ denote the population aged a in year t. Then the death rate for age a individuals in year t is $(S(a)_{t-1} - S(a+1)_t)/S(a)_{t-1}$. For each age from 0 to 99 years, we average over all the annual death rates from 1967 (the first year in our estimation) to 1998 (the last year for which we have the age distribution of the population) to obtain a non-parametric relationship between the death rate and age, $d(a)$. We then combine information on the average age of a name, A_{kt} with the age-specific death rate to estimate the stock of individuals with name k in year t, denoted S_{kt}. The formula applied is

$$S_{kt} = [1 - d(A_{kt})]S_{k,t-1} + n_{k,t-1}.$$

We used the website `nominis.cef.fr` to obtain a list of Saints recognized with "fêtes" in France. It uses the typical French spelling (e.g. Jean, not John). Of the 2664 listed Saints, 1101 are direct matches for names used in our data set and 1563 are names of Saints that were never used more than twice in France. We added compound names to the Saint list even if they were not the names of actual Saints if both elements are Saint names (as in Jean-Claude). This adds 910 additional names, giving 2011 Saint names in usage or 10.5% of the "universe" of 19,108 names given at least 3 times for a given gender in a year between 1900 and 2002.

A.2 Media-based names

The presence of names on French Media are measured using data for cinema, television and radio.

Movies

The movies that entered our data set, and their countries of origin, were listed in "Best-sellers du marché franais de 1945 à 2003," published online by the National Center of Cinematography (CNC). While no longer available in the form we downloaded, an updated version can be found in `http://www.cnc.fr/CNC_GALLERY_CONTENT/DOCUMENTS/publications/dossiers_et_bilan/306/ch01.pdf`. Our sample comprises the 180 movies receiving the largest audiences in France since 1945. Using the Internet Movie Database, `imdb.com`, we obtained the given names and sexes of the three principal roles (as ordered by IMDB) and the corresponding actors. Movie exposures "turn on" in the year of release in France and continue for two years thereafter. This extended effect is designed for movies that continued to be shown in various theaters in the year after release and were then distributed on other media (e.g. VHS).

[22]Those data are available on the website of the French national institute for demographic studies (INED) `http://www.ined.fr/cdrom_vallin_mesle/Donnees-de-base/Donnees-de-base.htm` and are based on the work by Meslé and Vallin (2002).

Television shows

For each of the non-pay channels in France—ORTF, TF1, Antenne2 (now France2), FR3 (France3), La Cinq (La Cinquieme/Arte), M6—we record data for all shows covered on the websites `www.leflt.com/annuseries` and `encyclopedie.snyke.com`. In most cases, we know the release dates in France and the US. In cases where we do not know the French release we set it at two years after the US release (the median gap in the data where both release years are known). We also know the number of seasons and assume that all seasons of the show are exhibited in France. As with movies, the main three role and actor names are taken from IMDB. This creates errors in the cases—mainly in the 1960s—when the French changed the character names in a TV show (e.g. Darrin was renamed Jean-Pierre in the French broadcast of Bewitched). Exposure turns for the duration of the initial run of the show on a non-pay station.

Songs

The website `www.infodisc.fr` provides, for a charge, the annual Top 100 popular song list for France going back to 1955 (note that the lists have less than 100 songs prior to 1959). The rankings aggregate multiple charts and take into account both sales of singles and radio play. We parsed the song title and the name of the performer into their constituent "words." We classified these words as names if they met two criteria: i) actually used as baby names in either France or the US, and ii) not among the most common 500 words in written French or English. Songs were classified as foreign if the title consisted mainly of non-French words. In cases where the title was a ambiguous (e.g. Michelle), we looked at the probable nationality of the performer, or, in a few cases, at websites that provide song lyrics. Exposure turns on only during the years a song is in the Top 100 in France.

A.3 US names

The Social Security Administration tracks given names in the US and makes them available on its website, `www.ssa.gov/OACT/babynames/`. At the time we downloaded the data, it provided the top 1000 names by sex by decade back to 1900 and annual top-1000 names after 1990.

B Additional Regression and Simulation Results

Table 8: Linear regression estimates of media effects of top 1000 American names, 1993–2002

	Dependent Variable: ln share of babies named k in year t					
Specification :	(1)	(2)	(3)	(4)	(5)	(6)
ln $(1+M_{kt})$	1.237^a	2.721^a	1.856^a	0.187	0.043	0.321^a
all media	(0.097)	(0.270)	(0.274)	(0.223)	(0.055)	(0.112)
ln $(1+ M^F_{kt})$		-1.508^a	-0.828^a	0.242	0.002	-0.023
foreign media		(0.258)	(0.258)	(0.221)	(0.051)	(0.049)
ln $(1+ M^P_{kt})$		-0.370^c	-0.281	-0.372^a	0.003	0.031
media performers		(0.194)	(0.190)	(0.122)	(0.039)	(0.035)
Saint name			1.383^a	0.094		
			(0.145)	(0.109)		
ln $(1+S_{kt})$				0.720^a	0.514^a	0.625^a
				(0.015)	(0.027)	(0.026)
ln $(1+ A_{kt})$				-1.124^a	-0.867^a	-0.372^a
				(0.036)	(0.055)	(0.056)
ln $(1+ S_{kt}) \times$ ln $(1+ A_{kt})$						-0.198^a
						(0.013)
ln $(1+ M_{kt}) \times$ ln $(1+ A_{kt})$						-0.090^a
						(0.029)
Media Multiplier (MM):	2.356	1.793	1.678	1.04	1.034	1.036
Fixed effects	none				name-sex (k)	
Observations	10363	10363	10363	10363	10363	10363
R^2	0.094	0.113	0.185	0.723	0.298	0.349

Note: Standard errors (name-sex clustered) in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10%. Sex-year dummies included in all specifications. MM is the ratio of $\mathbb{P}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 1$ to $\tilde{\mathbb{P}}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 0$. For specification (6) MM sets $A_{kt} = 21.1$ (sample mean).

Table 9: Media effects on French use of popular American names, 1993–2002

Specification :	Tobit on censored number of babies named k in year t (1)	(2)	(3)	(4)	(5)	(6)
$\ln(1+M_{kt})$ all media	3.355^a (0.062)	5.717^a (0.190)	3.455^a (0.183)	-0.326^a (0.077)	0.052 (0.046)	0.384^a (0.079)
$\ln(1+M^F_{kt})$ foreign media		-2.221^a (0.175)	-0.595^a (0.167)	0.795^a (0.070)	0.029 (0.044)	-0.041 (0.044)
$\ln(1+M^P_{kt})$ media performers		-0.820^a (0.149)	-0.595^a (0.139)	-0.394^a (0.058)	-0.012 (0.032)	-0.006 (0.033)
Saint name			3.188^a (0.073)	-0.471^a (0.034)	-0.024 (0.085)	0.577^a (0.081)
$\ln(1+S_{kt})$				0.926^a (0.005)	0.764^a (0.009)	0.913^a (0.012)
$\ln(1+A_{kt})$				-0.820^a (0.013)	-0.539^a (0.020)	-0.345^a (0.021)
$\ln(1+S_{kt}) \times \ln(1+A_{kt})$						-0.080^a (0.004)
$\ln(1+M_{kt}) \times \ln(1+A_{kt})$						-0.088^a (0.020)
Media Multiplier (MM):	10.234	6.395	4.809	1.053	1.049	1.072
Random effects	none				name-sex (k)	
Observations	20000	20000	20000	20000	20000	20000

Note: Standard errors (name-sex clustered) in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10%. Sex-year dummies included in all specifications. MM is the ratio of $\mathbb{P}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 1$ to $\tilde{\mathbb{P}}_{kt}$ with $M_{kt} = M^F_{kt} = M^P_{kt} = 0$. For specification (6) MM sets $A_{kt} = 13.7$ (sample mean).

Table 10: Media and non-media effects on French use of popular American names, 1993–2002

	Tobit on censored number of babies named k in year t					
Specification :	(1)	(2)	(3)	(4)	(5)	(6)
$\ln n_{kt}^{\text{US}}$	0.866^{a}	0.896^{a}	0.794^{a}	0.145^{a}	0.305^{a}	0.284^{a}
US popularity	(0.021)	(0.021)	(0.020)	(0.009)	(0.014)	(0.014)
$\ln(1+M_{kt})$	2.416^{a}	5.293^{a}	3.318^{a}	-0.274^{a}	0.027	0.346^{a}
all media	(0.062)	(0.178)	(0.173)	(0.077)	(0.044)	(0.077)
$\ln(1+M_{kt}^{F})$		-2.917^{a}	-1.386^{a}	0.614^{a}	0.016	-0.055
foreign media		(0.165)	(0.159)	(0.070)	(0.043)	(0.043)
$\ln(1+M_{kt}^{P})$		-0.778^{a}	-0.582^{a}	-0.394^{a}	0.006	0.011
media performers		(0.140)	(0.132)	(0.058)	(0.032)	(0.032)
Saint name			2.843^{a}	-0.467^{a}	-0.041	0.553^{a}
			(0.069)	(0.033)	(0.085)	(0.083)
$\ln(1+S_{kt})$				0.903^{a}	0.724^{a}	0.864^{a}
				(0.005)	(0.009)	(0.012)
$\ln(1+A_{kt})$				-0.789^{a}	-0.497^{a}	-0.310^{a}
				(0.013)	(0.019)	(0.021)
$\ln(1+S_{kt}) \times \ln(1+A_{kt})$						-0.076^{a}
						(0.004)
$\ln(1+M_{kt}) \times \ln(1+A_{kt})$						-0.084^{a}
						(0.019)
Media Multiplier (MM):	5.336	3.028	2.548	.964	1.034	1.054
Random effects	none				name-sex (k)	
Observations	20000	20000	20000	20000	20000	20000

Note: Standard errors (name-sex clustered) in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10%. Sex-year dummies included in all specifications. MM is the ratio of $\mathbb{P}_{kt}$ with $M_{kt} = M_{kt}^{F} = M_{kt}^{P} = 1$ to $\tilde{\mathbb{P}}_{kt}$ with $M_{kt} = M_{kt}^{F} = M_{kt}^{P} = 0$. For specification (6) MM sets $A_{kt} = 13.7$ (sample mean).

Table 11: Simulated effects of Media on renamings and welfare, 1993–2002

Sex	Media	Dynamics	# renamed	% renamed	Δ welfare
females	foreign	no	42157.5	1.19	.14
		yes	59641	1.68	.14
	all	no	62801.5	1.77	.22
		yes	86626	2.45	.21
males	foreign	no	63567.5	1.7	.20
		yes	88528	2.36	.20
	all	no	75748	2.02	.05
		yes	103036.5	2.75	.08

Table 12: Names most helped by all media, 1967–2002

name	male	years non-rare	years exposed	simulated # births media off	actual # births media on	(on-off)/off
Laura	0	36	17	33517	117913	251.8
David	1	36	35	86296	287812	233.52
Tia	0	4	4	32	105	229.2
Michael	1	36	35	28475	85245	199.37
Tom	1	35	30	9066	23881	163.4
Lisa	0	36	23	13251	34098	157.33
Jonathan	1	36	20	38448	98436	156.02
Jennifer	0	36	28	25524	63831	150.09
Theo	1	36	16	22753	54213	138.27
Johnny	1	36	32	5645	13446	138.19
Britney	0	4	3	58	136	134.99
Shakira	0	1	1	23	49	116.86
Calista	0	4	4	139	292	110.6
Xena	0	6	5	29	61	109.58
Tasha	0	3	3	5	10	93.35
Shannen	0	10	10	73	141	93.09
Anastacia	0	3	3	11	20	89.34
Jason	1	34	23	8161	15314	87.65
Alan	1	36	30	7270	13620	87.36
Tamera	0	2	1	20	37	86.42

Note: Actual and simulated # births are cumulated between 1967 and 2002. The number of births has been rounded to the nearest unit, while the (On-Off)/Off percentage is calculated before the rounding. "Years exposed" counts the number of years that the name was non-rare and appeared on media.

Table 13: Names most harmed by all media, 1967–2002

name	male	years non-rare	years exposed	simulated # births media off	actual # births media on	(on-off)/off
Sebastien	1	36	2	468649	294439	-37.17
Paul	1	36	34	86430	60771	-29.69
Delphine	0	36	1	164833	119963	-27.22
Guillaume	1	36	6	263542	194493	-26.2
Julien	1	36	30	353505	267636	-24.29
Margot	0	32	0	26467	20592	-22.2
Emma	0	36	14	52572	41087	-21.85
Arthur	1	36	24	48335	38102	-21.17
Romain	1	36	3	197121	155590	-21.07
Charles	1	36	33	46642	37304	-20.02
Camille	0	36	11	140999	113999	-19.15
Anais	0	36	1	107966	87381	-19.07
Jeremie	1	36	0	33919	27695	-18.35
Charlotte	0	36	17	93677	76777	-18.04
Eva	0	36	20	36663	30222	-17.57
Victor	1	36	20	46065	37990	-17.53
Louis	1	36	29	57074	47256	-17.2
Celine	0	36	12	261955	216949	-17.18
Pierre	1	36	36	180658	151868	-15.94
Hugo	1	36	1	72262	60759	-15.92

Note: Actual and simulated # births are cumulated between 1967 and 2002. The number of births has been rounded to the nearest unit, while the (On-Off)/Off percentage is calculated before the rounding. "Years exposed" counts the number of years that the name was non-rare and appeared on media.