Identifying the Presence and
Cause of Fashion Cycles in Data

Hema Yoganarasimhan [*]

University of Washington

Abstract

Fashions and conspicuous consumption play an important role in marketing. In this paper, we present a three-pronged framework to analyze fashion cycles in data – a) algorithmic methods for identifying cycles, b) statistical framework for identifying cycles, and c) methods for examining the drivers of such cycles. In the first module, we identify cycles based on pattern-matching the amplitude and length of cycles observed to a user-specified definition. In the second module, we define the Conditional Monotonicity Property, derive conditions under which a data generating process satisfies it, and demonstrate its role in generating cycles. A key challenge that we face in estimating this model is the presence of endogenous lagged dependent variables, which we address using system GMM estimators. Third, we present methods that exploit the longitudinal and geographic variations in agents' economic and cultural capital to examine the different theories of fashion. We apply our framework to data on given names for infants, show the presence of large amplitude cycles both algorithmically and statistically, and confirm that the adoption patterns are consistent with Bourdieu's theory of fashion as a signal of cultural capital.

1 Introduction

1.1 Fashion

Fashion, as a phenomenon, has existed and flourished since Roman times across a wide variety of conspicuously consumed products. The impact of fashion can be seen on all aspects of society and culture – clothing, painting, sculpture, music, drama, dancing, architecture, arts, and entertainment. According to the prominent sociologist Blumer (1969), fashion appears even in redoubtable fields such as sciences, medicine, business management, and mortuary practices.

Fashion plays an important role in the marketing of many commercial products. For example, the American apparel and footwear industry follows a seasonal fashion cycle, in the form of spring/summer and fall/winter collections. According to industry experts, a large chunk of the $300 billions that Americans spend on apparel and footwear annually is fashion rather than need driven. Fashion also influences the success of other conspicuously consumed products such as electronic gadgets, furniture, and cars. For instance, 1950's saw the rise and fall of tailfin craze in car designs. Even though tailfins were completely non-utilitarian, they contributed to the phenomenal success of Cadillacs and other cars sporting fins (Gammage and Jones, 1974).

Given the widespread impact of fashion and its economic importance, it is essential that we develop frameworks to help managers and researchers reliably identify fashion cycles in data and examine their drivers. However, to date we do not have an empirical framework to study fashion cycles. Further, no research examines whether the cycles observed in data are consistent with any of the proposed theories of fashion. In fact, apart from a few early descriptive works by Richardson and Kroeber (1940) and Robinson (1975), there is hardly any empirical work on fashion.[1] In this paper, we bridge this gap in the literature.

---

[1] Richardson and Kroeber (1940) document cyclical fluctuations in the dimensions of women's evening dresses advertised in fashion plates from 1789 to 1936. Robinson (1975) counts pictures of men with facial hair in *The Illustrated London News* from 1842 to 1972 and finds that facial hair grew in popularity from 1850-1880 before falling out of fashion. However, neither of these studies use choice data or provide an empirical framework to analyze fashion cycles in data, like we do.

1.2 Our Framework for Analyzing Fashion Cycles

We present a three-pronged framework to analyze fashion cycles in data – a) algorithmic methods for identifying cycles, b) statistical framework for identifying cycles, and c) methods for examining the drivers of fashion cycles.

The first module consists of an algorithmic framework for identifying fashion cycles based on pattern-matching the amplitude and length of cycles observed in the data to a user-specified definition of a cycle as satisfying certain minimum requirements on those dimensions. We also use algorithmic methods to characterize and identify recurring cycles, where each cycle is separated by a dormancy period that is allowed to be a function of the amplitude of the cycle. Taken together, these techniques allow us to characterize different types of cyclical patterns in data.

While algorithmic identification of cycles is sufficient for many purposes, it suffers from user subjectivity. So in the second module, we develop a statistical method to identify the presence of cycles. We define the *Conditional Monotonicity Property* and derive the conditions under which a data generating process satisfies this property. Specifically, an $AR(p)$ process is conditionally monotonic if it is non-stationary and continues to increase (decrease) in expectation if it was on an increasing (decreasing) trend in the last $(p-1)$ periods. We then demonstrate that conditional monotonicity is necessary and sufficient to give rise to cycles.

A key challenge that we face in estimating this model and establishing conditional monotonicity is the presence of potentially endogenous lagged dependent variables. In such cases, the two commonly used estimators – random-effects estimator and fixed-effects estimator – cannot be used (Nickell, 1981). While theoretically we can solve this by finding external instruments for the endogenous variables, it is difficult to find variables that affect lagged popularity of a fashion product, but not its current popularity. We address this issue using system GMM estimators that exploit the lags and lagged differences of explanatory variables as instruments (Blundell and Bond, 1998; Shriver, 2015).

Finally, in the third part, we expand our framework to examine the drivers of fashion. While different drivers of fashion have been proposed, two signaling theories have gained prominence due to their ability to provide internally consistent reasoning for the rise and fall of fashions – wealth

signaling theory (Veblen, 1899) and cultural capital signaling theory (Bourdieu, 1984).[2] While existing analytical models of fashion assume one of these social signaling theories and examine the role of firms in fashion markets, they do not test the empirical validity of either of these theories (Pesendorfer, 1995; Amaldoss and Jain, 2005; Yoganarasimhan, 2012a). In contrast, we present empirical tests to infer whether the patterns observed in data are consistent with one of these theories. We use aggregate data on the metrics of wealth and cultural capital of parents in conjunction with state-level name popularity data. We exploit the geographical and longitudinal variation in these two metrics to correlate name adoption to the predictions of the two theories.

1.3 Name Choice Context

We apply our framework to characterizing fashions in the choice of given names, *i.e.*, names given to newborn infants. We choose this as our context for four reasons. First, the choice of a child's name is an important conspicuous decision that parents make. So it is a good area to examine fashion and conspicuous choices. Second, to establish the existence of cycles in a product category, we need data on a large panel of products for a significantly long period. Our context satisfies this data requirement: Social Security Administration (SSA) is an excellent source of data on given names at both the national and state-level starting from 1880. Third, it is a setting where we observe large cycles of popularity, which makes it ideal for this study. Figure 1 depicts the rise and fall in popularity of the most popular male and female baby names from 1980. Note that at their peaks, these names were adopted by over 80,000 parents on a yearly basis, which hints at the presence of cycles of large amplitude in this data. Fourth, to examine the impact of social drivers of fashion cycles, we need both time and geographic variation in agents' status in the society, which is available in the form of metrics on economic and cultural capital through Census data. Together, these factors make it an ideal context

---

[2]Some older research views fashion as reflecting broader external changes within a society (Banner, 1983; Frings, 1991). However, such external theories are limited in their ability to explain and predict fashion cycles because they rely on post hoc correlations between exogenous societal changes and shifts in fashion. Further, a small set of exogenous factors cannot be responsible for all fashion phenomena given the diversity in the domains of fashion. For example, while the rise of skirt hems over the last century can be viewed as a result of the sexual revolution, the alternating fashion cycles of skinny and baggy jeans cannot be explained by external factors. Hence, most fashion theorists have focused on signaling theories that can provide self-consistent explanations of fashion. We refer readers to Sproles (1981); Miller et al. (1993); Davis (1994) for detailed discussions of these ideas.
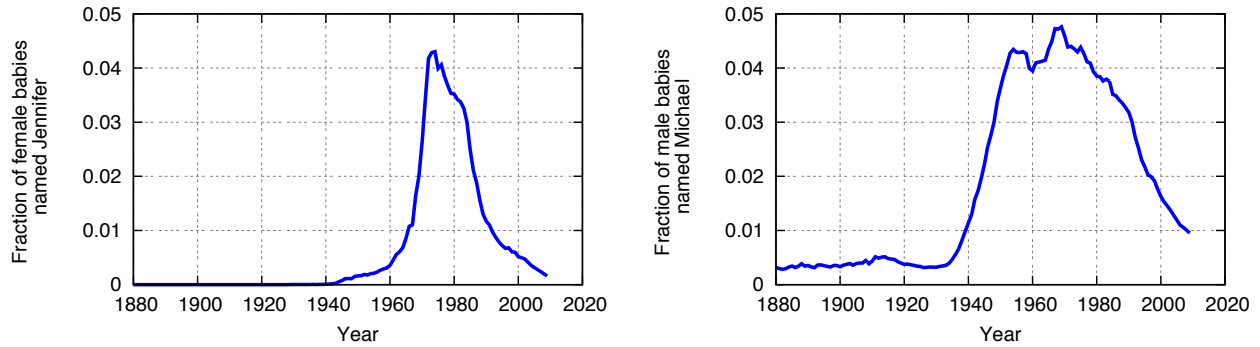
to study fashion cycles.

Figure 1: Popularity Curves of the Top Female and Male Baby Name from 1980.

1.4 Findings

Using our framework, we provide a series of substantive results. First, we establish the existence of large magnitude cycles in the names data using algorithmic methods. We show that more than 80% of the 361 names in Top50 have seen at least one cycle of popularity, and a significant fraction (about 30%) of these has gone through two or more cycles. In datasets with less popular names, the fraction of names with cycles is lower, but still significant. In fact, over 75% of the 1468 female names in Top500 have gone through at least one cycle. We also find that a significant fraction of names have gone through at least two cycles of popularity. For instance, 13.6% of names in Top100 have gone through a /_/\\ pattern, while 6.54% have gone through a _/_/ pattern.

Second, we apply our statistical framework to the name choice data and show that it follows an $AR(2)$ process that satisfies the Conditional Monotonicity Property. We show that the names data – a) exhibits non-stationarity, *i.e.*, has a unit root and b) in expectation moves in the direction of the movement from the last period, thereby satisfying the two conditions for Conditional Monotonicity. These results are robust across different types of data and model specifications. Thus, we have statistical evidence that the data generating process satisfies properties that lead to cycles when sampled over significantly long periods of time.

Third, we exploit the longitudinal and geographical variations in cultural and economic capital to show that these cycles are consistent with Bourdieu's cultural capital signaling theory. We present

4

three findings in this context. First, we show that states that have higher average cultural capital are the first to adopt names that eventually become fashionable; they are then followed by the not-so-cultured states. Similarly, the states with higher average cultural capital are the first to abandon increasingly popular names. In other words, the rate of adoption is higher among the cultured states at the beginning of the cycle, while the opposite is true at the end of cycle. Second, we find that adoption among the cultured states has a positive impact on the adoption of the general population, while adoption among not-so-cultured states has a negative impact on the overall adoption. Third, we do not find any such parallel results for economic capital. Taken together these results provide support for Bourdieu's theory in the name choice context.

Our results have implications for a broad range of fashion firms. First, our empirical framework allows firms to test for the presence of fashion cycles in their context. Second, it allows them to uncover the social signaling needs of their consumers, which in turn would allow them to target the right consumers at different stages of the fashion cycle. For example, if a firm finds that its products serve as signals of cultural capital, it can initially seed information with cultured consumers, and then release information to the larger population in a controlled manner so as to maximize profits.

2 Related Literature

Our paper relates to three broad streams of literature in marketing, sociology, finance, and economics.

First, it relates to the theoretical literature on conspicuous consumption and fashion cycles. Karni and Schmeidler (1990) present one of the earliest models of fashion with two social groups, high and low. Agents in both groups value products used by high types but not those used by low types. In this setting, they show that fashion cycles can arise in equilibrium. Similarly, Corneo and Jeanne (1994) show that fashion cycles may arise out of information asymmetry. On a related front, Amaldoss and Jain (2005) study the pricing of conspicuous goods. Pesendorfer (1995) adopts the view that fashion is a signal of wealth (Veblen, 1899), adds a firm to the mix and goes on to show that a monopolist produces fashion in cycles to allow high types to signal their wealth. In contrast, Yoganarasimhan (2012a) shows that firms may want to cloak information on their most fashionable products based on

a model where agents want to signal being 'in the know' or access to information.

Second, our paper relates to the macroeconomic literature on identification of business cycles from data pioneered by Burns and Mitchell (1946). Recent developments in this area advocate the use of band-pass filters to separate cycles from short-term fluctuations as well as long-term trends under the assumption that cycles indeed exist and that cycle length falls under certain limits (Hodrick and Prescott; Baxter and King, 1999). They are designed to work with a small number of time series that exhibit similar behaviors. Also, they do not offer any insights on the factors that give rise to cycles. Our approach differs from these methods in three important ways. First, we dont know whether a given name has gone through a cycle or not, and we don't limit the length of the cycles. Second, we have very large number of names and there is no co-movement or even similarity in the cycles (if any) across names. Third, we are interested in exploring the underlying reasons for fashion cycles and hence need a methodology that can accommodate endogenous explanatory variables.

Third, our paper relates to the finance literature on identifying stock market bubbles using non-stationarity tests (Diba and Grossman, 1988; Evans, 1991; Charemza and Deadman, 1995). The key difference between these papers and ours is that they define a bubble as any long-term deviation from the stable mean of an auto-regressive process. Thus, non-stationarity tests are sufficient to identify them. In contrast, we are interested in fashion cycles, which are defined as long-term deviations characterized by consecutive increases followed by consecutive decreases (or vice-versa), and caused due to social signaling. We show that non-stationarity is necessary but not sufficient to identify cycles, and go on to define conditional monotonicity, and demonstrate its ability to establish the presence of cycles. As with the previous methods, these cannot inform us on the drivers of fashion cycles, a point of interest for us.

Our paper also contributes to the growing literature on the measurement of social effects in marketing. See Tellis et al. (2009); Chintagunta et al. (2010); Nair et al. (2010); Yoganarasimhan (2012b); Sun et al. (2014) and Toubia et al. (2014) for some recent developments in this area. Finally, our paper relates to the literature on name choice, which we discuss in the next section.

3 The Naming Decision

How do parents choose names, and why does the popularity of a name change over time? These are interesting questions that have attracted attention of researchers in various domains. Sociologists were among the first to study names, and early works in this area include Rossi (1965), Taylor (1974), Lieberson and Bell (1992), Lieberson and Lynn (2003), and Lieberson (2000). More recently, in a descriptive study, Hahn and Bentley (2003) show that naming patterns can be described using power-law distributions and random regenerative models, while Gureckis and Goldstone (2009) include the effect of past adoptions to build a predictive model of name choice. Berger and Le Mens (2009) show that the speed of adoption of a name is correlated with its speed of abandonment. Based on a survey of expecting parents, they argue that this stems from negative perceptions of fads. While all these studies give us excellent insight into the sociological aspects of name choice and demonstrate interesting naming patterns, they do not empirically establish the presence of cycles in the data or examine drivers of these cycles – which is the focus of this paper.

We now present a discussion of factors that potentially affect parents' naming decisions. In §7.1.3, we discuss how these are controlled for in the empirical model.

3.1 Name Attributes

The popularity of a name is likely to depend on its attributes. For example, short names are both easy to speak and spell, which makes them attractive to many parents (*e.g.*, John *vs*. Montgomery). Similarly, parents may pick names that symbolize positive imagery and qualities, such as bravery (Richard), charm (Grace), and beauty (Helen, Lily).

3.2 Familial and Religious Reasons

Traditionally, newborns were named after their relatives. For instance, first-born boys were named after their father or paternal grandfather and first-born girls after their paternal grandmother. However, Rossi (1965) finds that this custom has been on a decline due to the rise of nuclear families.

Religious beliefs can also influence name choice. Many long-term popular names such as Joseph and Daniel have Biblical origins. However, Lieberson (2000) finds no correlation between church

attendance and popularity of Biblical names in both U.S and England. Therefore, even though some Biblical names have remained popular (*e.g.*, Samuel, Seth), their choice is likely driven by other considerations since many others have declined in popularity (*e.g.*, Michael, Paul).

3.3   Assimilation and Differentiation

Researchers have shown that names associated with an obvious ethnic or minority population can have a negative impact on a child's future employment and success (Bertrand and Mullainathan, 2004). Recognizing this, minority parents may choose conventional names to avoid discrimination and integrate their children into the mainstream society. Consistent with this theory, Mencken (1963) finds that names such as Lief, Thorvald, and Nils, that were popular among Norwegian immigrants suffered a rapid loss in popularity after their immigration to the United States.

In contrast, some minority parents may try to differentiate from the majority by choosing names that highlight their distinctive ethnic background. Fryer and Levitt (2004) find that African-American parents chose increasingly distinctive names in the 1970's, often with African roots, to emphasize their "Blackness". Of course, neither of these effects are at play for non-Black or non-ethnic names.

3.4   Celebrity names

Popular entertainers, sports stars, and celebrities are often mentioned in the mass media, and this exposure can influence parents' name choices. However, past research refutes the idea that fashion cycles in names are caused by celebrities. First, many stars adopt names that are currently popular, which in fact implies reverse causality. For example, Marilyn was already a popular name before Norma Jean Baker adopted it as her stage name, and Marilyn actually declined in popularity in the following years. Second, not all stars' names become popular and not all names that become popular are those of celebrities. Third, in the few cases where a name became popular around the same time as a rising celebrity, the resulting increase in its popularity has been minor compared to the magnitude of the usual cycles that we observe in the data. Finally, if popularity cycles in names are caused by celebrities, then, empirically, we should not find any difference in the rate of adoption among different classes of people at different stages of the fashion cycle. For example, a celebrity theory cannot give

rise to an adoption pattern where wealthy or the cultured parents are first to both adopt and abandon a name. See Lieberson (2000) for a detailed discussion of this idea.

3.5 Signaling Theories

Finally, parents may choose names to signal their (and their children's) high status in the society. Two kinds of signaling mechanisms can be at work.

Signal of wealth: Parents may choose certain names to signal their affluence. The wealth signaling theory would predict name cycles as follows – (1) wealthy parents first adopt certain names, which makes them signals of wealth; (2) the not-so-wealthy imitate these names, which dilutes their signaling values; (3) the wealthy abandon them because they are no longer exclusive signals of wealth, (4) when the wealthy abandon these names, their signaling value decreases even more and which leads to abandonment by the not-so-wealthy. This entire process constitutes a fashion cycle.

There is some support for this theory in the literature. Some sociologists have argued that the use of middle names by the English middle class is an imitation of the British aristocratic practice (Withycombe, 1977). Others have provided correlational evidence that suggests that names popular among the wealthy were later adopted by the not-so-wealthy (Taylor, 1974; Lieberson, 2000). However, the evidence in these studies is suggestive, not conclusive.

Signal of cultural capital: Parents' may choose names to signal their cultural capital and artistic temperament; and such an incentive on parents' part can also give rise to cycles in the popularity of names (following the same reasoning as that used in the context of wealth-based fashion cycles). In fact, Kisbye (1981) provides some evidence for this theory. In his study of English names in 19th century Aarhus (Denmark), he finds an increase in the use of English names in the earlier part of the century (with the introduction of English books by Shakespeare, Dickens *etc.*), followed by a decrease towards the end of the century. Kisbye argues that English names were first adopted by the cultured or well-read Danes. However, towards the end of the 19th century, the not-so-cultured residents obtained access to these previously obscure texts and started adopting English names, which in turn diluted their signaling value and led to their eventual decline. While Kisbye does not provide concrete evidence to substantiate this speculation, his study suggests that names can be used as a

vehicle to signal cultural capital. Similarly, Lieberson and Bell (1992) and Levitt and Dubner (2005) also provide some correlational evidence for the cultural capital signaling theory.[3]

By definition, signaling theories require an action to be not only costly, but also require it to be differentially costly across types for it to serve as a credible signal of the sender's type. Given that names are free, we may not expect either of these signaling theories to work. However, this is a naive inference because the cost of gathering information on the set of names popular among the high types (wealthy or culture) would vary with the parent's own wealth and cultural capital. There is considerable evidence on network homophily (McPherson et al., 2001). Researchers have found that social networks are strongly homophilous on both wealth and cultural capital. For example, Marsden (1990) found that about 30% of personal networks are highly homophilous on education, which is one of the strongest indicators of cultural capital (see §4.2.2). This homophily has powerful implications for people's access to information. If cultured people live in similar neighborhoods, attend similar cultural events, work in similar environments, and overall interact more with each other than with those outside their group, then it is easier for a cultured parent to obtain information on the names that other cultured people have given their children compared to a not-so-cultured parent. Hence, network homophily can give rise to heterogeneity in signaling costs across classes of people and therefore allow names to serve as signals of parents' types.

In §7, we examine whether the name cycles are consistent with one of these two signaling theories, after controlling for the alternative explanations discussed earlier.

4 Data

We use two types of data in our study: (1) data on popularity of names, and (2) data on the cultural and economic capital of parents.

---

[3]Their evidence is purely correlational, *i.e.*, they do not control for other factors that could simultaneously drive name choices. In a critical commentary on Lieberson and Bell (1992), Besnard (1995) counters that most of the names popular among the highly educated in the early parts of the cycles studied by Lieberson and Bell (1992) were also popular among the larger population. He also asserts that their findings are unlikely to be meaningful given their short time-frame of 13 years. Our own analysis suggests that name cycles are, on average, much longer than thirteen years.

4.1 Data on Names

Our data on names comes from the Social Security Administration (SSA), the most comprehensive source of given names in the United States. All newborn U.S citizens are eligible for a Social Security Number (SSN) and their parent(s) can easily apply for one while registering the newborn's birth. While getting a SSN for a child is optional, almost all parent(s) choose to do so because a SSN is necessary to declare the child as a dependent in tax returns, open a bank account in the child's name, and obtain health insurance for the child.[4] The SSA therefore has information on the number of children of each sex who were given a specific name, for each year, starting 1880. The SSA was established in 1935 and became fully functional only in 1937. Many people born before 1937 never applied for a SSN, and the data from 1880 to 1937 is a partial sample of the names from that period. Therefore, we restrict our empirical analysis to the data from 1940 to 2009.

This data is available at both national and state levels. At the national level, for each name i, we have information on the number of babies given name i in time period t, which we denote as n_{it}. Since our data of interest starts from 1940, $t=1$ denotes the year 1940. The name identifier i is sex-specific. For example, the name Addison is given to both male and female babies, but we assign different ids to the two Addisons. To preserve privacy, if a name has been given to less than 5 babies in a year, SSA does not release this number for that particular year. In such cases, we treat n_{it} as zero. The state level data is available for all the 50 states. We use n_{ijt} to denote the number of babies given name i in state j in time period t. As in the national dataset, n_{ijt} is also left-truncated at 5, in which case we treat it as zero.

For each name i, we construct the following variables:

- s_i = the sex of name i. $s_i = 1$ if i is a female name and $s_i = 0$ if it is a male name.

- l_i = the number of characters in name i.

- bib_i = the number of times that name i appears in the Bible.

---

[4]There is a small discrepancy between the number of annual registered births and number of SSNs assigned. This may be due to the fact that some infants die before the assignment of SSNs. Alternately, a small set of parent(s) may choose not to participate in the process for personal reasons.

SSA also furnishes data on the total number of SSNs issued to newborns each year both nationally and statewide. We use this data to construct the following variables:

- Γ_{s_it} = total number of babies of sex s_i assigned SSNs in period t, nationally. Thus, Γ_{0t} and Γ_{1t} are the total number of male and female babies born in period t.

- Γ_{s_ijt} = total number of babies of sex s_i assigned SSNs in state j in period t.

- $f_{it} = n_{it}/\Gamma_{s_it}$ is the fraction of babies of sex s_i given name i in period t.

- $f_{ijt} = n_{ijt}/\Gamma_{s_ijt}$ is the fraction of babies of sex s_i given name i in period t within state j.

While there are a total of 56,937 female and 33,745 male names in the data, a small subset of these names account for a large portion of name choices. In order to focus our analysis on a representative sample of names, we work with the following four subsets of data:

- Top50 dataset: For each year starting with 1940, we collect the top 50 male and the top 50 female names given to newborns in the country. We then pool these names and denote the resulting set as the Top50 dataset.

- Top100 dataset: same as above but including names that have appeared in the top 100.

- Top200 dataset: same as above but including names that have appeared in the top 200.

- Top500 dataset: same as above but including names that have appeared in the top 500.

Table 1 shows the number of names in each dataset by sex and also provides the fraction of total births that these datasets account for. For example, the Top500 dataset contains a total of 1468 female names, which together account for 60.99% of all female births from 1940 to 2009.

Next, we examine the patterns in the name choice data. Table 4 shows the top ten female and male names for the years 1940, 1950, 1960, 1970, 1980, 1990, 2000, and 2009. It is clear that there is quite a bit of churn in popular names. For instance, of the 10 most popular female names in 1990, only 5 remained in the top 10 in 2000. To understand the patterns better, we plot the popularity of the top 6 female and males names from 1980 for the full span of our data, *i.e.*, from 1880 to 2009 (see Figures 2 and 3). The plots present clear visual evidence of cycles in the data.
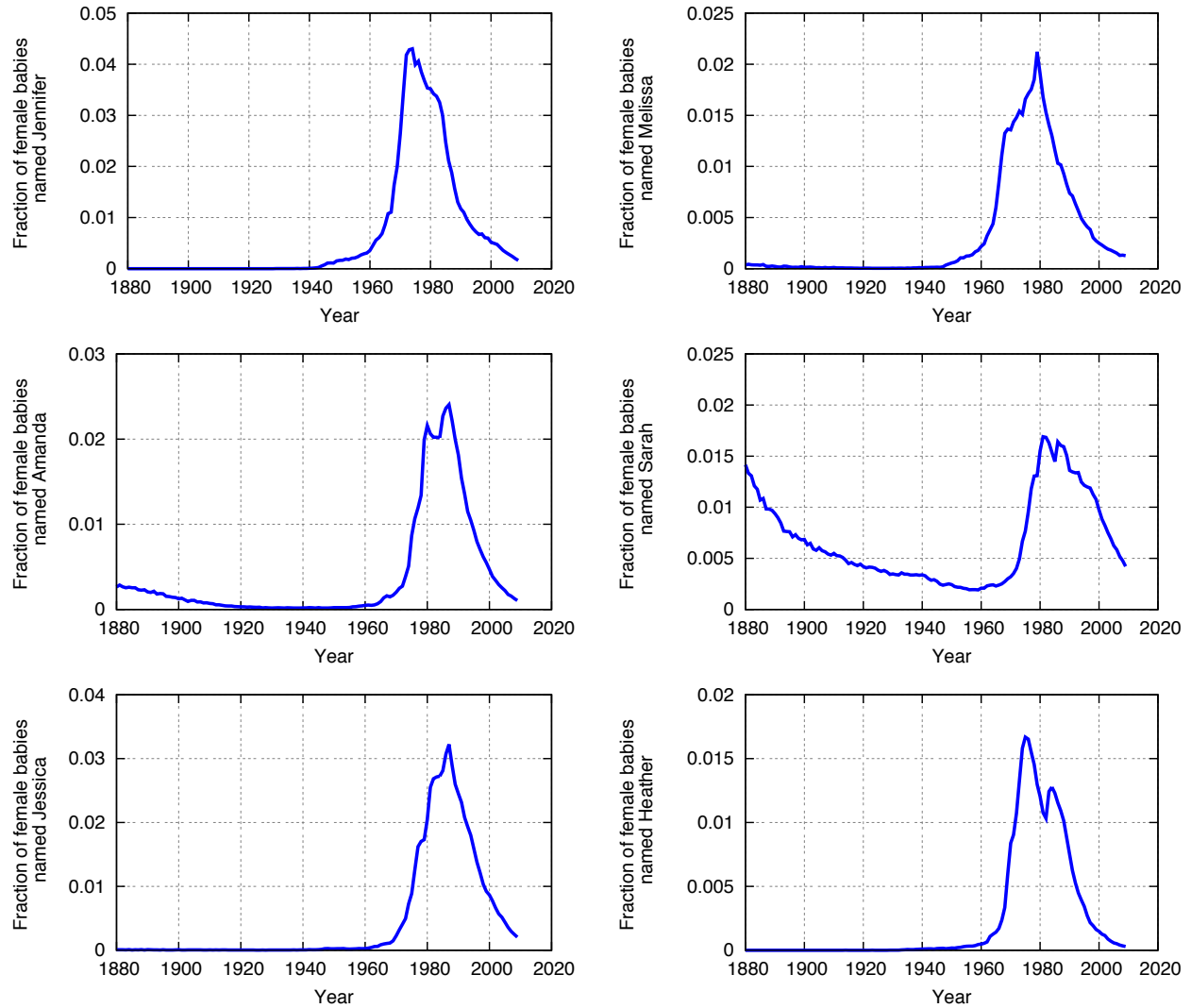
Figure 2: Popularity Curves of the Top 6 Female Baby Names in 1980. Top row: Jennifer, Melissa. Middle row: Amanda, Sarah. Bottom row: Jessica, Heather.

4.2 Data on Economic and Cultural Capital

To examine the two theories of fashion, we need data on the geographical (state-level) and longitudinal (yearly) variations in the economic and cultural capital of decision makers.

4.2.1 Economic Capital

We use a state's median household income at period t as a measure of the economic capital of the decision-makers from that state at t. Our income data comes from two sources – the decennial Census and the Social and Economic Supplements of the Current Population Survey (CPS). We retrieve data
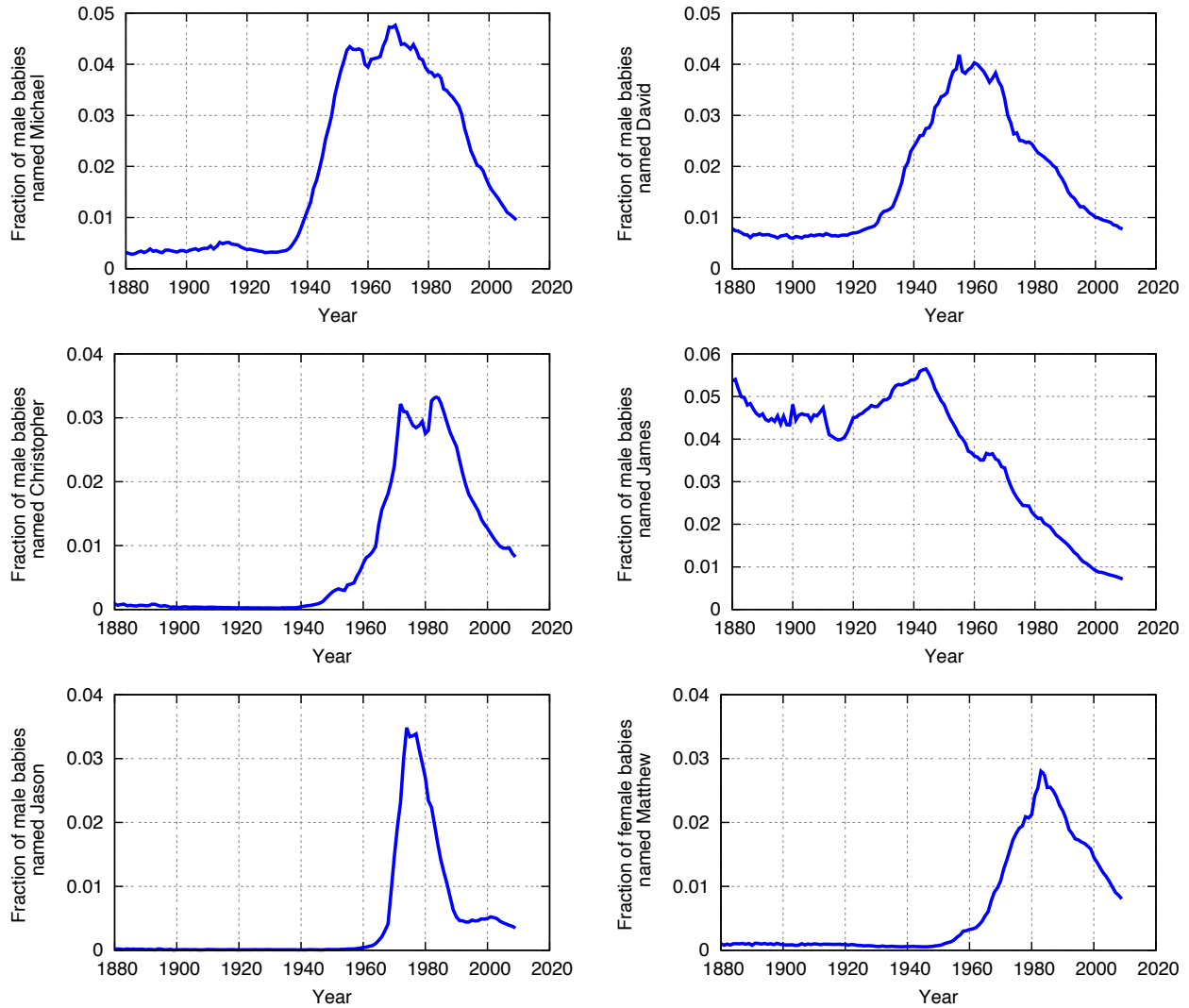
Figure 3: Popularity Curves of the Top 6 Male Baby Names in 1980. Top row: Michael, David. Middle row: Christopher, James. Bottom row: Jason, Matthew.

on the state-level median household income for 1970 and 1980 from the decennial census tables.[5] For 1984-2009, we obtain annual state level data on median household income from CPS. To calculate values for the intervening years we use linear interpolation. This is reasonable since a state's median income rarely exhibits wide year-to-year fluctuations.

The original data is in current dollars (*i.e.*, reported dollars). To get a normalized measure of wealth, we need to correct for both inflation over time and for geographic variations in cost of living.

---

[5]While the Census Bureau has asked income related questions from 1940, the wording used in the question formulation in 1940, 1950, and 1960 was different from that in use now (family vs. household income), making it difficult to combine the data from the former years with our current dataset.

We do this using the 'revised 2009 version of the Berry-Fording-Hanson (BFH) state cost of living index' (Berry et al., 2000). We denote this normalized metric as w_{jt}, the adjusted median household income of state j in period t. It is obtained as follows:

$$w_{jt} = \frac{\text{Median income of state } j \text{ in period } t}{\text{BFH cost of living index of state } j \text{ in period } t} \tag{1}$$

The BFH index is a measure of how costly a state is in comparison to a median state in 2007 (the index for the two middle states, New Mexico and Wyoming, is set to 100 in 2007). Table 2 lists the top and bottom five wealthiest states based on w_{jt} for 1970, 1980, 1990, and 2000.

4.2.2 Cultural Capital

Cultural capital is defined as an individual's knowledge of arts, literature, and culture (Dimaggio and Useem, 1974; Bourdieu, 1984). The most commonly used measures of cultural capital is education attainment, especially higher education (Robinson and Garnier, 1985; Cookson and Persell, 1987; Lamont and Lareau, 1988).

We use the percentage of adults in state j with a bachelor's degree or higher in period t, as a measure of the educational attainment of decision makers from that state in period t. This data comes from the U.S Census Bureau (for years 1970, 1980, 1990, and 2000, and interpolated for intervening years) and the CPS (annually for 2001-2006). As in the case of income, the absolute number of people with bachelor's degree is an imperfect metric of the relative cultural capital of decision makers in period t, especially since people have become more educated with time. Hence, for each state j in period t, we subtract the national average of the percentage of the adults with bachelors degree, and use this as the measure of the cultural capital c_{jt}. Table 3 lists the most and least educated states (based on c_{jt}) for 1970, 1980, 1990, and 2000.

5 Algorithmic Detection of Popularity Cycles

5.1 Definitions and Algorithm

Essentially, a cycle is an increase followed by a decrease (*i.e.*, an inverted V-shaped curve as that exhibited by Jennifer from 1940-2009 in Figure 2) or a decrease followed by an increase (*i.e.*, a V-shaped pattern, like the one exhibited by Sarah from 1880-1980 in Figure 2). However, not all

bumps and troughs are classified as cycles because they could simply stem from long-term trends in the data (decreasing or increasing) or cycles of popularity. While it is easy to visually identify popularity cycles in a small set of names (*e.g.*, Figures 2 and 3), visual identification is neither feasible nor consistent when analyzing a large set. Therefore, we now present a formal definition of a cycle, which we then use to detect and characterize cycles in the data. We start by providing some terminology. Consider a sequence of T real numbers, $x_1, x_2, ..., x_T$.

Definition 1. *We define operators $\prec$ and $\succ$ as follows:*

(a) $x_i \prec x_j$ if $x_i < x_j$ or if $x_i = x_j \wedge i < j$.

(b) $x_i \succ x_j$ if $x_i > x_j$ or if $x_i = x_j \wedge i > j$.

Definition 2. *We define a local minimum and a local maximum as follows:*

(a) x_i is a local minimum if $x_i \prec x_j$ for all $i - \tau \leq j \leq i + \tau$.

(b) x_i is a local maximum if $x_i \succ x_j$ for all $i - \tau \leq j \leq i + \tau$.

Using this notation, we now define a cycle as follows.

Definition 3. *A cycle $\mathcal{C}$ is a sequence of three values $\{x_i, x_j, x_k\}$ with $i < j < k$ that satisfies the following conditions:*

1. *x_i, x_k are local minimas and x_j is a local maximum or x_i, x_k are local maximas and x_j is a local minimum.*

2. *$Length(\mathcal{C}) \geq L$, where $Length(\mathcal{C}) = k - i$ is the distance between the first and last points of the cycle.*

3. *$Amplitude(\mathcal{C}) \geq M$, where $Amplitude(\mathcal{C}) = \min\{|x_i - x_j|, |x_j - x_k|\}$ is the amplitude of the cycle.*

To be classified as a cycle, a bump or trough has to be *significant* in both time and magnitude. We weed out insignificant deviations through two mechanisms: (1) A local maxima or minima has to dominate τ values to both its right and left (see Definition 2 and the first condition of Definition 3). Thus, a short-term increase in a curve that is on a decreasing trend is not classified as a local maxima

16

and vice-versa. This ensures that we only capture consistent increases and decreases, not shocks in time. Further, the total length of the cycle has to be at least L to ensure that we are capturing real patterns in the data and not shocks (see the second condition of Definition 3). (2) The amplitude of a cycle has to be greater than a baseline value M (see the third condition of Definition 3). For example, even if we find a name has followed an inverted V-shaped pattern, but if the magnitude is very small, then we do not classify it as a cycle.

5.2 Application of Algorithm to Name Choice Context

We now apply these definitions and algorithm to name choice context. Specifically, we set $\{M, \tau, L\} = \{0.00005, 4, 10\}$ and analyze the time series of f_{it} in the datasets of interest.[6] We perform our analysis on the Top50, Top100, Top200, and Top500 datasets and present the results in Table 5. Of the 361 Top50 names, more than 80% have seen at least one cycle of popularity. Moreover, a significant fraction (30%) has gone through two or more cycles of popularity. This suggests the presence of recurring fashion cycles. In datasets with less popular names, the fraction of names with fashion cycles is lower, but still quite significant. For example, more than 75% of the 1468 female names in Top500 have gone through at least one cycle. In Table 6, we present details on the main patterns of repeat cycles seen in the data. Different types of patterns are prevalent at varying frequencies. For instance, 13.6% of names in Top100 have gone through a $/\backslash\_/\backslash$ pattern, while 6.54% have gone through $\backslash\_/\backslash\_/$.

Interestingly, several names have gone through more than one cycle (see Figure 4). To better understand repeat cycles in names, we analyze the time it takes for cycles to repeat. We define 'dormancy length' as the period between two popularity cycles where the name is dormant or adoptions for the name are close to minimum. Formally:

Definition 4. *Given two adjacent cycles* $\mathcal{C}_1 = \{x_i, x_j, x_k\}$ *and* $\mathcal{C}_2 = \{x_l, x_m, x_n\}$ *such that* $|x_k - x_l| < d_t * Amplitude(\mathcal{C}_2)$, *where* $d_t < 1$ *is a dormancy threshold. The dormancy length is defined as* $l - k$.

---

[6]Note that if we set lower values of M, τ, and L we will find more cycles in the data. By setting relatively high values of these parameters, we are setting a higher bar for classifying a bump or trough as a cycle. Please see Table 13 in Online Appendix B.1 for a senstivity analysis to varying τ and M.
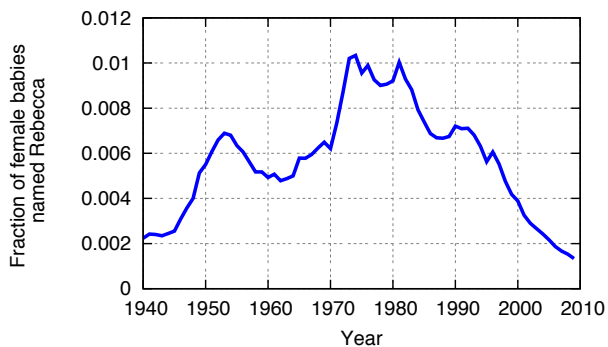
Figure 4: Popularity Curve of Rebecca

In Table 7, we provide the statistics for dormancy length when the dormancy threshold is defined as 10% (i.e., the change in values from the end of the first cycle to the beginning of the second cycle is less than 10% of the amplitude of the second cycle). For all four datasets, Top50, Top100, Top200, and Top500, the median dormancy period is between 3-8 years. However, a large number of names also remain dormant for significant periods before enjoying a resurgence. For example, the 75th quartile of dormancy length for Top100 male names is 29. Further, the dormancy periods are longer for female names compared to male names.
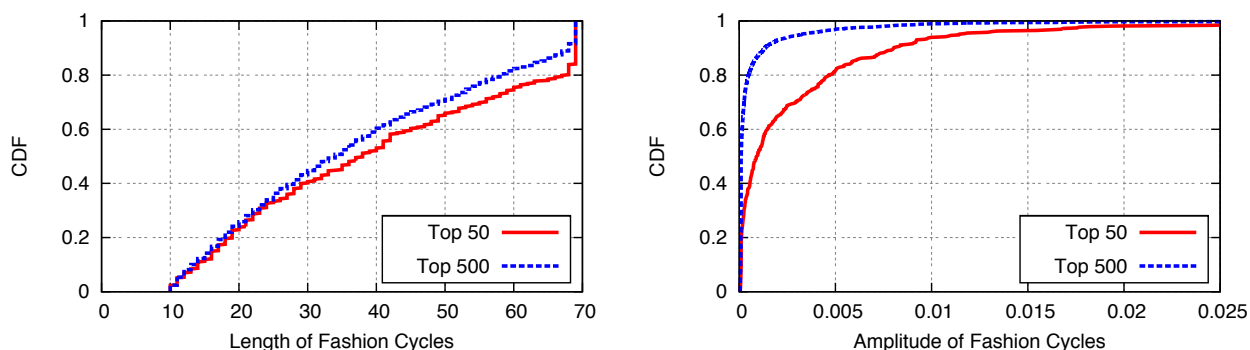
Figure 5: Distributions of the length and magnitude of cycles for Top50 and Top500 datasets.

The left panel of Figure 5 shows the distributions of cycle lengths and cycle magnitudes for the Top50 and Top500 datasets. Note that this is a CDF over cycles and not names; some names might have more than one cycle, in which case their data is represented multiple times, and some others might have no cycles in which case they are not represented in the graph. Please see Table 5 for data on the fractions of names with different numbers of cycles. Two patterns emerge from these two

18

figures. First, the cycle length distributions are similar for the two datasets. Second, while the cycle lengths span a large range, the median name enjoys a ≈ 35-year cycle. The right panel shows the magnitude distributions for the Top50 and Top500 datasets. Not surprisingly, the distribution for the Top50 dataset first-order stochastically dominates that of the Top500 dataset. Note that 20% of all names in Top50 have an amplitude of 0.005 or more, which implies that these names were chosen by more than 10000 parents per year at the peak of their popularity.

6   An Empirical Framework for Identifying Cycles

In the previous section, we saw that the data presents clear evidence of cycles. Our efforts to identify and classify these cycles were algorithmic. We provided a specific definition of a cycle and identified patterns in the data that satisfied our definition. In this section, we establish the presence of cycles using statistical analyses. There are two main reasons for developing a statistical framework over and above algorithmic methods. First, statistical methods are not influenced by user subjectivity unlike the algorithmic methods, which require the values of τ, L, and M as user input. Second, statistical methods can include other explanatory variables that drive these cyclical patterns.

Fashion cycles differ from standard product life cycles (Levitt, 1965; Day, 1981) in two important ways. First, they can potentially reappear. Theory models of signaling-based fashions predict such recurring fashions (Corneo and Jeanne, 1994), a prediction confirmed by casual observation (*e.g.*, skinny jeans). Recall that even in our setting, a significant fraction of names go through multiple cycles of popularity. Second, they have to be caused by social signaling. Both these properties have to be satisfied for a cycle to be defined as a fashion cycle. For example, repeat cycles can occur without social signaling simply driven by a firm's marketing activities. Similarly, social signaling can occur in non-conspicuous arenas unrelated to fashion. Hence, formally:

Definition 5. *An adoption curve is defined as a social-signaling based fashion cycle if:*

- *it satisfies statistical properties that can lead to repetitive cycles over sufficiently long periods.*

- *the cycles (if they exist) are caused by social signaling – either wealth signaling or cultural capital signaling.*

19

An empirical framework that seeks to identify the presence and cause of fashion cycles in data has to provide researchers tools to establish the two properties described above. In this section, we focus on the first aspect of the problem – identifying the presence of cycles in data using statistical tests. In §7, we outline the second part of our framework, wherein we present tools to test whether the cycles are indeed caused by social signaling.

6.1 Conditional Monotonicity Property

Observe that name cycles are inverted V-shaped rather than inverted U-shaped curves. In this respect, they resemble stock market and real estate bubbles rather than standard product life cycle curves. We know from finance literature that bubbles occur when consumers' utility and actions depend on their expectations and beliefs on others' valuation of the product rather than the inherent attributes of the product (Camerer, 1989). In such settings, small changes in consumers' beliefs and expectations can cause large shifts in behavior. Since consumers' behavior in fashion markets are also driven by their beliefs on what other consumers consider fashionable (Yoganarasimhan, 2012a), it is understandable that the popularity cycles of names follow similar patterns.

Note that unlike financial economists who are interested in bubbles, we are interested in cycles. A bubble is defined as an auto-regressive process that does not have a stable long-term mean. In contrast, cycle is an auto-regressive process that shows a clear cyclical behavior or generates an inverted V-shaped curve. Traditionally, the finance literature has used non-stationarity tests to identify bubbles in data (Diba and Grossman, 1988; Evans, 1991; Charemza and Deadman, 1995). However, we show that non-stationarity alone is not sufficient to generate cycles, and therefore provide a more precise framework for identifying cycles using the concept of 'Conditional Monotonicity'.

Let the popularity of a conspicuously consumed product i evolve as an AR(p) process, *i.e.*, an autoregressive process of order p, as follows:

$$y_{it} = \sum_{k=1}^{p} \phi_k y_{it-k} + \left[1 - \sum_{k=1}^{p} \phi_k\right] \eta_i + \epsilon_{it} \tag{2}$$

where $|\sum_{k=1}^{p} \phi_k| \leq 1$. y_{it} is a measure of product i's popularity in period t, $[1 - \sum_{k=1}^{p} \phi_k]\eta_i$ is an unobserved product fixed effect, and ϵ_{it} is a mean-zero shock. The multiplier $[1 - \sum_{k=1}^{p} \phi_k]$ in front of η_i ensures that the total effect of the unobservable in each period is always fixed at η_i.
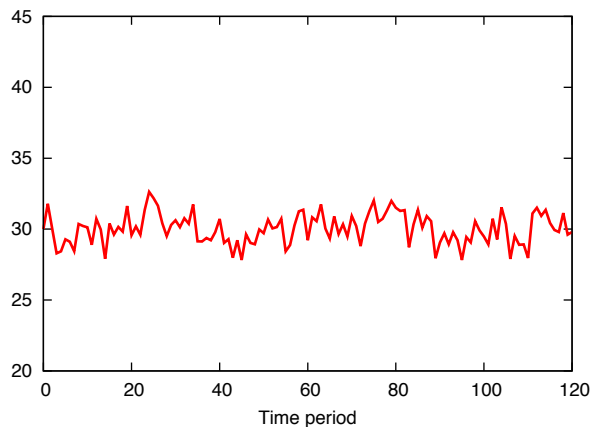
Figure 6: AR(1) stationary process with $\phi_1 = 0.5, \eta_i = 30$.
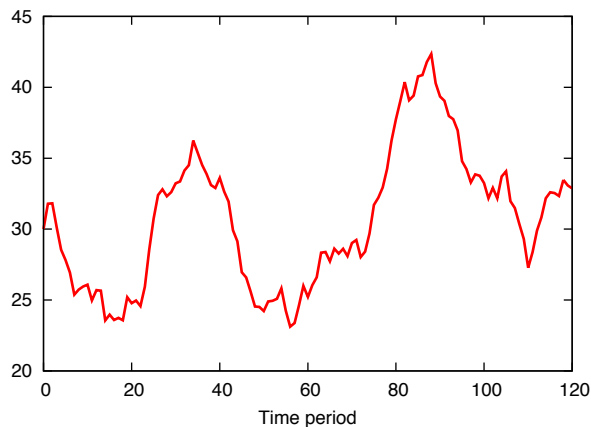
Figure 7: AR(2) conditionally monotonic process with $\phi_1 = 1.5, \phi_2 = -0.5, \eta_i = 30$.

This simple framework can be easily expanded to include other time-varying and time-invariant explanatory variables. Equation (2) can be rewritten as follows:

$$\Phi_p(L)y_{it} = \gamma_i + \epsilon_{it} \tag{3}$$

where $\Phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 ... - \phi_p L^p$, with L denoting the lag operator and $\gamma_i = [1 - \sum_{k=1}^{p} \phi_k]\eta_i$. Depending on the parameter values, this process is either stationary or non-stationary.

An AR(p) process is stationary if all the roots of the polynomial $\Phi_p(L)$ lie outside the unit circle. Under these conditions, a shock to the system dissipates geometrically with time and the resulting process is mean-reverting and stable. For example, if the popularity of name i follows a stationary process, then shocks to its popularity (*e.g.*, election of a president of a name i or the sudden fame of a celebrity with name i) will dissipate with time, and its popularity will soon return to its long-term average. See Fuller (1995) and Dekimpe and Hanssens (1995) for detailed discussions on stationary time-series models. In an AR(1) process, the stationarity condition boils down to $|\phi_1| < 1$, and it can be written as $y_{it} = \phi_1 y_{it-1} + (1 - \phi_1)\eta_i + \epsilon_{it}$. Figure 6 shows an AR(1) process with $\phi_1 = 0.5$ and $\eta_i = 30$. Note that this is a very stable process that oscillates around a constant mean of 30. The expectation of the tth realization of a stationary AR(p) series is a weighted mean of its last p realizations and the unobserved fixed effect η_i. Hence, every period, there is a constant pull towards the mean η_i, and this property makes a stationary process stable. Of course, an important implication of this stability is that a stationary process cannot give rise to popularity cycles significant in either

21

time or magnitude.

An AR(p) process is non-stationary if one or more of the roots of $\Phi_p(L)$ lie on the unit circle. When subjected to a shock, a non-stationary series does not revert to a constant mean, and its variance increases with time. If name choices are non-stationary, then shocks due to celebrities, politicians, *etc.*, can cause long-term shifts in name popularity. An AR(1) process with $\phi_1 = 1$ is non-stationary and is referred to as a random walk process. It can be written as $y_{it} = y_{it-1} + \epsilon_{it} \Rightarrow E(y_{it}) = y_{it-1}$. Therefore, at any point in time, the process evolves randomly in one direction or the other. While a random walk process is not mean reverting, it doesn't produce cycles of any significant magnitude either, because its specification does not imply consecutive increases or decreases. Hence, non-stationary is not sufficient to generate cycles. Below, we define the *Conditional Monotonicity Property* and describe its role in generating cycles. Let Δ be the first difference operator such that $\Delta y_{it} = y_{it} - y_{it-1}$.

Proposition 1. *A non-stationary AR(p) process with roots* $1, \frac{1}{c_1}, \frac{1}{c_2}, \ldots, \frac{1}{c_{p-1}}$*, where* $p \geq 2$ *and* $0 < c_1, c_2, \ldots, c_{p-1} \leq 1$*, is conditionally monotonic in the following sense:*

- *If* $\prod_{k=1}^{p-2}(1 - c_k L)\Delta y_{it-1} \geq 0$*, then* $E\left[\prod_{k=1}^{p-2}(1 - c_k L)\Delta y_{it}\right] = c_{p-1}\prod_{k=1}^{p-2}(1 - c_k L)\Delta y_{it-1} \geq 0.$

- *If* $\prod_{k=1}^{p-2}(1 - c_k L)\Delta y_{it-1} \leq 0$*, then* $E\left[\prod_{k=1}^{p-2}(1 - c_k L)\Delta y_{it}\right] = c_{p-1}\prod_{k=1}^{p-2}(1 - c_k L)\Delta y_{it-1} \leq 0.$

Proof: See Appendix A.1. □

According to Proposition 1, in a conditionally monotonic AR(p) process, there is a lower bound on $E(\Delta y_{it})$ if last $p-1$ periods' changes satisfy the constraint $\prod_{k=1}^{p-2}(1 - c_k L)\Delta y_{it-1} \geq 0$. So, conditional on past lags, the current y_{it} is expected to be at least $\prod_{k=1}^{p-2}(1 - c_k L)\Delta y_{it-1}$, irrespective of η_i. Similarly, there is an upper bound on $E(\Delta y_{it})$ if $\prod_{k=1}^{p-2}(1 - c_k L)\Delta y_{it-1} \leq 0$. Note that these bounds are not dependent on η_i. In certain lower order AR(p) processes, conditional monotonicity manifests itself as cycles.

We now demonstrate the implications of conditional monotonicity for an AR(2) process here and refer readers to Appendix A.1 for a general proof for an AR(p) process. Consider a non-stationary AR(2) process of the form $y_{it} = (1 + c_1)y_{it-1} - c_1 y_{it-2} + \gamma_i + \epsilon_{it}$, where $0 < c_1 \leq 1$ and $\gamma_i = (1 - \sum_{k=1}^{2}\phi_k)\eta_i = (1 - [(1+c_1) - c_1])\eta_i = 0$. Note that this process satisfies the requirements for conditional monotonicity, because its two roots are 1 and $\frac{1}{c_1}$, where $0 < c_1 \leq 1$. This series

22

can be rewritten as $y_{it} = y_{it-1} + c_1\Delta y_{it-1} + \epsilon_{it}$. When this series is on an increasing trend, it has a tendency to keep increasing because $E(y_{it}) = y_{it-1} + c_1\Delta y_{it-1} > y_{it-1}$, when $\Delta y_{it-1} > 0$. That is, conditional on an increase in the last period ($\Delta y_{it-1} > 0$), the series continues to increase in expectation. Similarly, when this series is on a decreasing trend, it has a tendency to keep decreasing because $E(y_{it}) = y_{it-1} + c_1\Delta y_{it-1} < y_{it-1}$, when $\Delta y_{it-1} < 0$. That is, conditional on a past decrease ($\Delta y_{it-1} < 0$), the series continues to decrease in expectation. In data, this property manifests itself as periods of consecutive increase followed by periods of consecutive decrease – a pattern that can be interpreted as cycles. Hence, we can establish the presence of fashion cycles in an AR(2) process by showing that the underlying process is conditionally monotonic. As an illustration, see Figure 7, which shows the presence of cycles in the conditionally monotonic process defined by $y_{it} = 1.5y_{it-1} - 0.5y_{it-2} + \epsilon_{it}$ and $\eta_i = 30$.

Note that non-stationarity is necessary, but not sufficient for conditional monotonicity. Non-stationarity is necessary because in stationary processes, the conditional expectation $E(y_{it})$ remains dependent on η_i, precluding us from making any general statements on the relationship between $E(y_{it})$ and its past $(p-1)$ lags. For example, consider the stationary AR(2) process $y_{it} = \phi_1 y_{it-1} + \phi_2 y_{it-2} + \gamma_i + \epsilon_{it}$, where $\gamma_i = (1 - \sum_{k=1}^{2}\phi_k)\eta_i \neq 0$. In this case, $E(y_{it}) = \phi_1 y_{it-1} + \phi_2 y_{it-2} + \gamma_i$. Even when this process is on an increasing trend ($\Delta y_{it-1} > 0$), we cannot make the general claim that $E(y_{it}) > y_{it-1}$ because $E(y_{it})$ depends on η_i. Hence, non-stationarity is a necessary pre-requisite for conditional monotonicity. However, non-stationarity is not sufficient to induce consecutive periods of increase or decrease. For example, consider the non-stationary AR(2) process $y_{it} = 0.5y_{it-1} + 0.5y_{it-2} + \epsilon_{it}$ $\Rightarrow y_{it}(1-L)(1+0.5L) = \epsilon_{it}$. This process is not conditionally monotonic because one of its roots is -2, $i.e.$, $c_1 = -0.5 < 0$. Note that this doesn't give rise to consecutive increases or decreases because $E(y_{it}) = 0.5(y_{it-1} + y_{it-2}) < y_{it-1}$ when $\Delta y_{it-1} > 0$. Hence, we need conditional monotonicity, over and above non-stationarity, to establish the presence of cycles in the data.

Finally, a conditionally monotonic process needs to be observed for sufficiently long periods of time to generate cycles. While the property is defined over the change in the y_{it} from the last period, such changes need to be observed for a long enough period to observe a full cycle or multiple cycles

(as in Figure 7).

6.2 Application: Identifying Cycles in the Choice of Given Names

6.2.1 Model

We now expand Equation (2) to suit our specific context as follows:

$$y_{it} = const. + \sum_{k=1}^{p} \phi_k y_{it-k} + \alpha x_{it} + \beta z_i + \gamma_i + \epsilon_{it} \tag{4}$$

y_{it} denotes n_{it}, number of babies given name i in period t. We model this as a function of:

1. the last p lags of i's popularity. This captures the past trends in name i's popularity that can affect adoption by current parents.

2. x_{it}, time-varying factors that affect i's popularity. Here x_{it} consists of the number of babies of sex s_i born in year t and time dummies.

3. z_i, the time-invariant attributes of name i that affect its popularity – length, sex, and number of Biblical mentions.

4. a name fixed effect $\gamma_i = [1 - \sum_{k=1}^{p} \phi_k] \eta_i$, which comprises of time-invariant unobservables that affect name i's popularity such as its historical relevance, symbolism, and meaning.

5. a mean-zero shock ϵ_{it} that captures shocks to a name's popularity. This can stem from a variety of factors, including but not limited to, the rise or fall of celebrities, entertainers, and book characters.

   Further, we make the following assumptions about the model:

- Assumption 1: $E(\epsilon_{it}) = E(\gamma_i \cdot \epsilon_{it}) = E(\epsilon_{it} \cdot \epsilon_{ik}) = 0 \ \forall \ i,t,k \neq t$

  We follow the familiar error components structure, *i.e.*, ϵ_{it} is mean-zero and uncorrelated to γ_i for all i,t. It is allowed to be heteroskedastic, but assumed to be serially uncorrelated. Since this is an important and strong assumption, we test its validity after estimating the model using the Arellano-Bond (2) test.

- Assumption 2: $E(x_{ik} \cdot \epsilon_{it}) = E(x_{ik} \cdot \gamma_i) = E(z_i \cdot \epsilon_{it}) = E(z_i \cdot \gamma_i) = 0 \ \forall \ i,k,t$

  The time invariant attributes of a name are assumed to be uncorrelated to both ϵ_{it} and γ_i. x_{it} is assumed to be uncorrelated to γ_i because the long-term mean of any name i is unlikely to be correlated with the total births of either sex in year t. Moreover, since the decision to have a child

is unlikely to be influenced by shocks in the popularity of a specific name i, there is no reason to expect ϵ_{it} to be correlated to past, current, or future births.

Assumption 2 is specific to this context. In a different context where explanatory variables are pre-determined or potentially endogenous, it can be easily relaxed (see §7).

While both x_{it} and z_i are uncorrelated to the shock and the fixed effect, the same cannot be said of the lagged dependent variables – the dynamics of the model imply an inherent correlation between the lagged dependent variables $(y_{it-1},...y_{it-k})$ and the unobserved heterogeneity γ_i if $\gamma_i \neq 0$. Moreover, ys and ϵs are correlated by definition. Since the current error term affects both current and future popularity $\Rightarrow E(y_{ik} \cdot \epsilon_{it}) \neq 0$ if $k \geq t$. However, past popularity remains unaffected by future shocks $\Rightarrow E(y_{ik} \cdot \epsilon_{it}) = 0$ if $k < t$.

6.2.2 Estimation

We rewrite Equation (4) as follows and use this formulation in our subsequent analyses.

$$y_{it} = const. + \mu y_{it-1} + \sum_{k=1}^{p-1} \theta_k \Delta y_{it-k} + \alpha x_{it} + \beta z_i + \gamma_i + \epsilon_{it} \tag{5}$$

where $\mu = \sum_{k=1}^{p} \phi_k$ and $\theta_k = -\sum_{j=k+1}^{p} \phi_k$.

If all the panels in the dataset follow a non-stationary process, then $[1 - \sum_{k=1}^{p} \phi_k]\eta_i = 0 \Rightarrow \gamma_i = 0 \,\forall$ i. In such cases, the endogeneity bias due to the correlation between the lagged dependent variables $(y_{it-1},...,y_{it-k})$ and the name fixed effect η_i is not an issue, and in theory, Equation (4) can be consistently estimated using a pooled OLS (Bond et al., 2002). However, if the non-stationarity assumption is violated even for a few panels, pooled OLS estimates are inconsistent. Moreover, in §7, we consider models with endogenous time varying variables (x_{it}) and such models cannot be estimated using pooled OLS. Therefore, we avoid pooled OLS and look for estimators that can accommodate endogenous variables and are robust to deviations from non-stationarity.

The two commonly used methods of estimating panel data models, Random-effects estimation and Fixed-effects estimation cannot be used in a dynamic setting. The former requires explanatory variables to be strictly exogenous to the fixed effect γ_i, an untenable assumption if some panels are indeed stationary. The latter allows for correlation between γ_i and explanatory variables, but since it uses a within-transformation, it requires all time-varying variables to be strictly exogenous to ϵ_{it}.

This is impossible in a dynamic setting with finite T (Nickell, 1981). While theoretically, we can solve this problem by finding external instruments, it is difficult to find variables that affect lagged name popularity, but do not affect current name popularity. We therefore turn to the GMM style estimators of dynamic panel data models that exploit the lags and lagged differences of explanatory variables as instruments (Blundell and Bond, 1998). This methodology has been successfully applied by researchers in a wide variety of fields in marketing and economics (Acemoglu and Robinson, 2001; Durlauf et al., 2005; Clark et al., 2009; Yoganarasimhan, 2012b; Shriver, 2015). We briefly outline the method here and refer readers to the aforementioned papers for details.

System GMM Estimator: First, consider the first-difference of Equation (5).

$$\Delta y_{it} = \mu \Delta y_{it-1} + \sum_{k=1}^{p-1} \theta_k \Delta^2 y_{it-k} + \alpha \Delta x_{it} + \Delta \epsilon_{it} \tag{6}$$

Notice that first differencing has eliminated the fixed effect γ_i, thereby eliminating the potential correlation between the lagged dependent variables and γ_i. However, by first differencing we have introduced another kind of bias. Now the error term $\Delta \epsilon_{it}$ is correlated with the explanatory variable Δy_{it-1} through the error term ϵ_{it-1}. However, it is easy to show that lags and lagged differences of y_{it} from time period $t-2$ and earlier are uncorrelated to $\Delta \epsilon_{it}$, and can therefore function as instruments for Δy_{it-1} and $\Delta^2 y_{ik}$s. Also, since Δx_{it} is uncorrelated with $\Delta \epsilon_{it}$, it can instrument for itself. So we specify the following sets of moment conditions for Equation (6):

$$E(y_{ip} \cdot \Delta \epsilon_{it}) = 0 \; \forall \, p \leq t-2 \tag{7}$$

$$E(\Delta y_{ip} \cdot \Delta \epsilon_{it}) = 0 \; \forall \, p \leq t-2 \tag{8}$$

$$E(\Delta x_{it} \cdot \Delta \epsilon_{it}) = 0 \; \forall \, t \tag{9}$$

In theory, these moments are sufficient to identify ϕ_ks and α as long as the process is not first-order non-stationary. However, *a priori*, it is not clear whether or not these moment conditions are sufficient for identification in our context. So following Blundell and Bond (1998), we also consider moment

conditions for the level Equation (5).

$$E(\Delta y_{iq} \cdot (\gamma_i + \epsilon_{it})) = 0 \ \forall \, q \leq t-1 \tag{10}$$

$$E\left(\Delta^2 y_{iq} \cdot (\gamma_i + \epsilon_{it})\right) = 0 \ \forall \, q \leq t-1 \tag{11}$$

$$E(x_{it} \cdot (\gamma_i + \epsilon_{it})) = 0 \tag{12}$$

$$E(z_i \cdot (\gamma_i + \epsilon_{it})) = 0 \tag{13}$$

x_{it} and z_i can instrument for themselves since they are uncorrelated to both γ_i and ϵ_{it}. Lagged differences of y_{it} from period $t-1$ and earlier can be used as instruments for y_{it-1} and Δy_{it-k}s. The moment conditions (10) and (11) hold irrespective of the stationarity properties of the process. They only require the initial deviations of the dependent variable to be independent of its long-term average, which is a reasonable assumption in most settings, including ours.

Stacking the moments gives us a system GMM estimator that provides consistent estimates irrespective of the stationarity properties of the process. We employ a two-step version of the estimator because it is robust to panel-specific heteroskedasticity and increases efficiency. However, the standard errors of the two-step GMM estimator are known to be biased. Windmeijer (2005) proposed a correction for this bias, and we follow his method to obtain robust standard errors.

Serial Correlation and Lagged Dependent Variables: A key assumption in the method outlined above is that the error terms are not serially correlated. Serial correlation is problematic for two reasons. First, in the presence of serial correlation, the restrictions that we apply break down. For example, consider a scenario where errors follow a MA(1) process such that $\epsilon_{it} = \rho \epsilon_{it-1} + u_{it}$, where $E(u_{it}) = 0$ and $E(u_{it} \cdot u_{ik}) = 0 \ \forall \, k \neq t$. Then, for $q = t-1$, (10) can be expanded as: $E(\Delta y_{it-1} \cdot (\gamma_i + \rho \epsilon_{it-1})) = 0$. However, this moment condition is invalid because Δy_{it-1} is correlated with ϵ_{it-1}. Similarly, we can show that moment conditions (7), (8), and (11) also fail to hold in the presence of serial correlation. Second, the absence of serial correlation confirms the absence of omitted variable biases. See Section 7.1.3 for a detailed discussion on this issue. Therefore, for all the models that we estimate, we test the validity of the instruments and the absence of omitted variable bias using the Arellano-Bond test for serial correlation (Arellano and Bond, 1991).

6.2.3 Results and Discussion

We estimate the model on the four datasets of interest, Top50, Top100, Top200, and Top500, and present the results in Table 9. The instruments for each of the level and differenced equations are shown in the last four lines of the table itself. GMM refers to the instruments generated from the lagged dependent variables, and standard refers to exogenous variables that instrument for themselves.

In all the models, we find that the coefficient of Δy_{it-2} is insignificant, implying that the process is AR(2). So Equation (5) can be written as:

$$y_{it} = const. + \mu y_{it-1} + \theta_1 \Delta y_{it-1} + \alpha x_{it} + \beta z_i + \gamma_i + \epsilon_{it} \tag{14}$$

This process satisfies the conditional monotonicity property if and only if $\mu = 1$ and $0 < \theta_1 \leq 1$. Under these conditions, the two roots of the process are 1 and $\frac{1}{\theta_1}$, where $0 < \theta_1 \leq 1$. Hence, for all the four models, we test the following two hypothesis:

- H1: $\mu = 1$

- H2: $\theta_1 = \bar{\theta}_1$, where $\bar{\theta}_1$ is a positive constant such that $0 < \bar{\theta}_1 \leq 1$

The results from the hypothesis tests are shown in Table 9. First, for all the four models, we cannot reject the null of hypothesis H1 that $\mu = 1$. This suggests that the data generating process is non-stationary and contains a unit root. Second, in all the models, we cannot reject the null of hypothesis H2 that $\theta_1 = 0.47$. Together, these results present clear evidence for the existence of cycles in the data since they demonstrate the conditional monotonicity of the underlying process.

The Arellano-Bond test confirms that our model is not mis-specified; we cannot reject the null hypothesis of no second-order serial correlation in first-differenced error terms, *i.e.*, the tests present no evidence of serial correlation. This establishes the validity of our moment conditions and confirms the absence of omitted variable biases. Nevertheless, in all our models, we include time period dummies. They control for unobserved time-varying variables such as education, income, urbanization, and religious preferences, which may affect name choice.

In all the models, the coefficient of z_i is insignificant. This is understandable because in a truly non-stationary model, the impact of time-invariant observed attributes should also be zero, just like

the impact of the time-invariant unobserved attributes is zero. Recall that $\gamma_i = (1 - \sum_{k=1}^{p} \phi_k) \eta_i = 0$ because $(1 - \sum_{k=1}^{p} \phi_k) = 0$. Similarly, β can be expressed as $\beta = \hat{\beta}(1 - \sum_{k=1}^{p} \phi_k) = 0$.

Our results are robust to variations in model specification and data used. When we estimate the model with f_{it} (the fraction of babies given name i in period t) as the dependent variable instead of n_{it} (number of babies given name i in period t), the qualitative results remain unchanged. Similarly, the results are robust to the following changes in the data used – (1) inclusion of all the names that have been in top 1000 at least once (this dataset can be referred to as Top1000), (2) inclusion of a set of randomly picked names to our existing datasets, and (3) inclusion of observations prior to 1940 (*i.e.*, analyzing all the data from 1880-2009 instead of focusing on the data from 1940-2009).

7   Empirical Framework for Analyzing the Drivers of Fashion

In the last two sections, we saw that there is both algorithmic and statistical evidence for the existence of popularity cycles of large magnitudes in the data. In this section, using state-level variation in economic and cultural capital, we examine whether these cycles are consistent with one of the two signaling theories of fashion – 1) fashion as a signal of wealth, and (2) fashion as a signal of cultural capital. We also consider and rule out a series of alternative explanations.
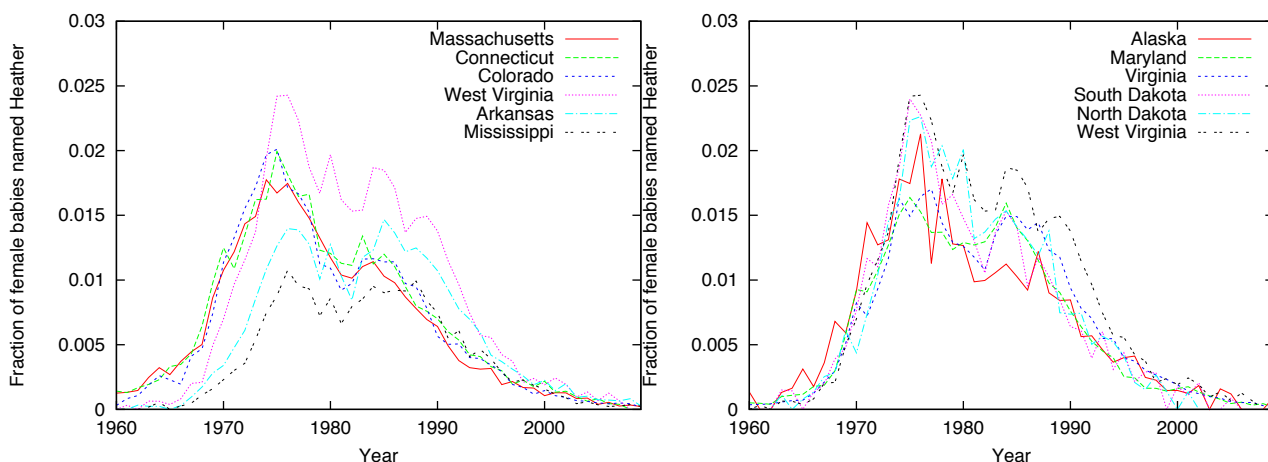
Figure 8: Popularity of Heather in three of the most and least educated states (left panel), and in three of the most and least wealthy states (right panel).

We start with a visual example using the popularity curve of Heather (Figure 8). The left panel shows Heather's popularity in the three most and three least educated states. It is clear that Heather
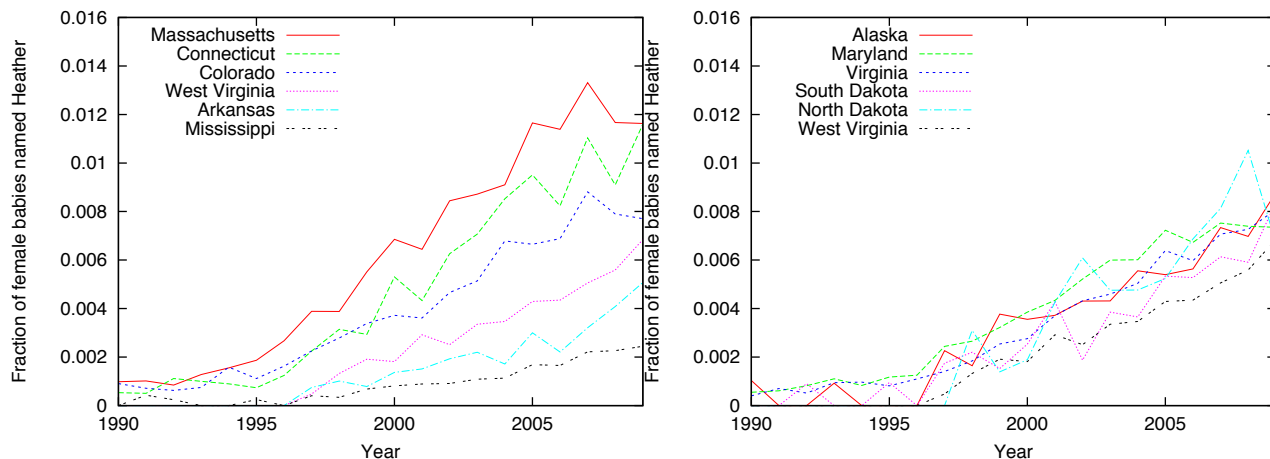
Figure 9: Popularity of Sophia in three of the most and least educated states (left panel), and in three of the most and least wealthy states (right panel).

became popular in the more educated states (Massachusetts, Connecticut, and Colorado) before it took off in the less educated ones (West Virginia, Arkansas, and Mississippi). Similarly, note that it starts dropping in popularity in the highly educated states first. However, we find no such patterns in the right panel, which shows Heather's popularity cycles in the three most and three least wealthy states. This pattern repeats in more recently popular names like Sophia too (Figure 9). Taken together, these patterns are suggestive evidence in support of the cultural capital theory. However, visual evidence from a few names is not conclusive, so we examine the data further for model-free patterns.

Preliminary examination of the data indicates that there are significant differences in when a name takes off and peaks across states. Table 8 compares the relative order of peaking in Colorado (high cultural capital state) and West Virginia (low cultural capital state). In the Top500 names, we find that 71.33% of the names peak in Colorado before West Virginia, *i.e.*, names tend to take off and peak in high cultural capital states before low cultural capital states. Nevertheless, even this is not conclusive because it does not control for other factors that affect name choice. So henceforth, we focus on empirical analysis.

To confirm that the cycles in the data are consistent with social signaling, our empirical tests should confirm the following two statements:

- The high types are the first to adopt a name followed by low types. Similarly, high types are the

30

first to abandon the name, followed by low types, *i.e.*, the rate of adoption is higher among the high types at the beginning of the cycle, while the opposite is true at the end of cycle.

- Adoption by high types has a positive impact on the adoption of the general population, while adoption by low types has a negative impact on the adoption of the general population.

where high types = wealthy and low types = poor if the wealth signaling theory is true, and high types = cultured and low types = uncultured if the cultural capital signaling theory is true. Below, we present two models that test the validity of each of these statements.

A potential issue with using state-level data to make inferences on individual behavior is aggregation-bias (Stoker, 1993; Blundell and Stoker, 2005). So in Appendix B.2, we explain how an individual-level model aggregates to the state-level models employed in this section.[7]

7.1 Interacting Wealth and Cultural Capital with Past Adoptions

7.1.1 Model

We expand the model of name popularity to the state level as follows:

$$y_{ijt} = const. + \sum_{k=1}^{p} \phi_k y_{ijt-k} + \lambda_w w_{jt} + \lambda_c c_{jt} + \delta_w w_{jt} y_{it-1} + \delta_c c_{jt} y_{it-1} + \alpha_1 x^1_{ijt} + \alpha_2 x^2_{ijt} + \beta z_i + \gamma_{ij} + \epsilon_{ijt} \quad (15)$$

- $y_{ijt} = n_{ijt}$, which is the popularity of name i in state j at time t.

- w_{jt} and c_{jt} are metrics of wealth and cultural capital of state j in period t.

- $w_{jt} y_{it-1}$ and $c_{jt} y_{it-1}$ captures the interaction between the lag of the total country level adoption of name i and wealth, cultural capital of state j.

- x^1_{ijt} are endogenous time-varying factors that affect name popularity such as past lags of the number of babies given name i at the national level (denoted as $y_{it-1}, y_{it-2}, etc.$).

- x^2_{ijt} are exogenous time-varying factors that affect name popularity such as the total number of babies born in state j in period t and time dummies.

- z_i are time-invariant attributes of the name already discussed in §6.2.1.

---

[7]The lowest level of geography in our data is a state. Hence, all our models in this section are specified at the state level. In a study on installed base effects in hybrid adoptions, Narayanan and Nair (2012) find that social effects tend to be stronger at lower geographical aggregations. So our use of a relatively high level of aggregation should, if anything, reduce the likelihood of our finding evidence in favor of social signaling.

- $\gamma_{ij} = [1 - \sum_{k=1}^{p} \phi_k] \eta_{ij}$ is an unobserved name-state fixed effect. This controls for the mean unobserved state-level preference for name i.

- ϵ_{ijt} is a i.i.d mean-zero time and state varying shock that affects the popularity of name i in state j. It captures differential exposure and other random effects, *e.g.*, the TV show with a lead character named i may randomly get aired in one market before another, or a local news item may mention name i in state j at time t.

The above model can be augmented to include state-time dummies. However, we found no significant effects for such dummies; so we skip them.

y_{it-1k}s are endogenous because y_{it-k} is a function of y_{ijt-k}, which in turn is a function of $\gamma_{ij} \Rightarrow E(y_{it-k} \cdot \gamma_{ij}) \neq 0$. Since the interaction terms $w_{jt}y_{it-1}$ and $c_{jt}y_{it-1}$ are functions of y_{it-1}, we treat them as endogenous too. We modify our earlier moment conditions to accommodate these changes and ensure that these correlations are not violated in our moment conditions. To avoid repetition, we do not describe the estimation strategy again. However, for each model estimated, we list the set of instruments for the level and first-differenced equation in the results tables.

7.1.2 Results and Discussion

Model N1 in Table 10 presents the results from the estimating the model on the Top50 dataset. We discuss the estimates from this model here and refer readers to §7.1.4 for robustness checks.[8]

In Model N1, the mean effect of c_{jt} is positive, while its interaction with past country level adoption $c_{jt} \cdot y_{it-1}$ is negative. So the total effect of c_{jt} is $c_{jt}(7.749 \times 10^{-2} - 2.780 \times 10^{-5} \cdot y_{it-1})$. Recall that c_{jt} is positive for states with high education and negative for states with low education (Table 3). For low values of y_{it-1} ($\approx y_{it-1} < 2787$), the overall impact of c_{jt} is increasing with education. Hence, at low values of y_{it-1}, the impact of education is increasingly positive for states with education higher than the national average ($c_{jt} > 0$), and increasingly negative for states with education lower than the national average ($c_{jt} < 0$). This suggests that high education states are

---

[8]Recall that names data is left-truncated at 5. At the state level, in Top50 names, 22.01% of the data is zero; in Top100, 29.54% of the data is zero; in Top200, 41.54% of the data is zero. We cannot tell how many of these are truly zeroes and how many are values less than 5. If we include less popular names, a significant fraction of these zeroes are likely to come from truncation. Truncation can adversely affect the quality and significance of the estimates. Hence, to keep the estimates clean, we avoid less popular names, and confine our analysis to the Top50 and Top100 datasets.

more likely, and low education states are less likely, to adopt a name at the early stages of the cycle (when its countrywide adoption is low). On the other hand, for high values of y_{it-1} (approximately, $y_{it-1} > 2787$), the exact opposite is true. Here, the overall impact of c_{jt} is increasingly negative for high education states ($c_{jt} > 0$) and increasingly positive for states with low education ($c_{jt} < 0$). That is, high education states are more likely to abandon a name as it becomes very popular, and the rate of abandonment is increasing with education. In contrast, low education states are more likely to adopt a name as it becomes very popular, and this rate of adoption is increasing as the education levels decrease.

The effect of the wealth metric, w_{jt}, is the opposite of that of education – the mean effect of w_{jt} is negative, while its interaction with past country level adoption $w_{jt} \cdot y_{it-1}$ is positive. This suggests that, after controlling for cultural capital, name cycles are starting with the less wealthy states and then spreading to the more wealthy ones. Thus, our results do not support wealth signaling theory, but are consistent with the cultural capital theory.

7.1.3 Controlling for Other Factors that Affect Name Choice

We now explain how our model controls for other factors that affect name choice discussed in §3.

Name Attributes – We control for time-invariant name attributes using both observed variables such as length, number of Biblical mentions, sex, and an unobserved state-name fixed effect γ_{ij}. γ_{ij} captures state j's preference for the name, its origin, symbolism, ease of pronunciation, *etc*. The inherent unobserved attractiveness of a name in a state can change over time and cause state-level trends in its popularity. Such trends are captured through lagged dependent variables (y_{ijt-k}s).

Familial, religious reasons, and assimilation, differentiation incentives – All these reasons can be grouped under the heading of peer-effects because they capture the impact of previous adoptions by others of same ethnicity, familial background, or religion on own adoption (Nair et al., 2010; Shriver et al., 2013). They are captured using lagged dependent variables. If lags are insufficient controls, then the model would suffer from serial correlation, which is not so in our case.[9] Hence, in all the

---

[9] We can illustrate the reasoning behind this using a simple example. Consider a scenario where parents' choices are influenced (among other things) by their need to fit in with a certain ethnic group. To this end, they may want to pick names that are currently popular in this group. Formally, let r_{ijt-1} denote the number of babies from the ethnic group given name

models that we estimate, we take care to add enough lags of past adoption on the right hand side to control for such time-varying name specific effects. We verify the adequacy of these controls using the Arellano-Bond test that confirms the absence of serial correlation.

Celebrity names – As discussed in §3.4, a popular lay theory of name adoption is based on celebrity adoption. While this has been refuted in earlier research (Lieberson, 2000), we now explain how our model controls for celebrity adoptions.

First, the impact of newly popular celebrities on naming decisions is captured through time varying error terms (ϵ_{ijt}). Once a celebrity is well-known, the past unobserved effect of her(/his) name on parents' choice (either through awareness of the name or due to adoptions by other parents) is accounted for through the lagged dependent variables. Again, the lack of serial correlation in the error terms ensures that these unobserved effects are adequately controlled for. More importantly, a celebrity-based theory cannot account for the differential rate of adoption (or abandonment) among different subsets of parents.

7.1.4 Robustness Checks

We conduct many checks to validate the robustness of our results. Below, we outline the key ones. First, we re-estimate the model with the Top100 dataset and confirm that the qualitative results remain the same (Model N2, Table 10). Next, according to Berry et al. (2000), the BFH index does not sufficiently normalize the cost of living for Alaska, making it look wealthier than it really is. Hence, we rerun our analysis after dropping it and confirm that the results are similar to our earlier ones (Model N3, Table 10).

---

i in state j in period $t-1$ and suppose that this number affects i's popularity in state j in period t. In our model, r_{ijt-1} is indirectly controlled for through y_{ijt-1} (which is a function of r_{ijt-1}), and we can be confident that this is an adequate control. Because, if it were not, then r_{ijt-1} would be a true omitted variable that would appear in the error-term as follows: $\epsilon_{ijt} = \delta r_{ijt-1} + u_{ijt}$, where $E(u_{ijt}) = E(u_{ijt} \cdot u_{ikt}) = 0 \; \forall \, t, k \neq t$. Moreover, since the number of parents from the ethnic group choosing name i in period $t-1$ is likely to be highly correlated to the number of parents from this group choosing name i at period $t-2$, r_{ijt-1} can be expressed as $r_{ijt-1} = \zeta r_{ijt-2} + v_{ijt-1}$, where $E(v_{ijt}) = E(v_{ijt} \cdot v_{ijk}) = 0 \; \forall \, t, k \neq t$. In that case, $E(\epsilon_{jit} \cdot \epsilon_{ijt-1}) = E((\delta r_{ijt-1} + u_{ijt}) \cdot (\delta r_{ijt-2} + u_{jit-1})) = \delta^2 E(r_{ijt-1} \cdot r_{ijt-2}) = \delta^2 \zeta \neq 0$. So if the lags of y_{ijt-1} do not sufficiently control for name specific time-varying factors that affect name popularity, then the model would suffer from serial correlation.

7.2 Impact of Adoption by States with the Highest, Lowest Cultural and Economic Capital

So far, our analysis was restricted to analysing the adoption patterns of a name within a state and relating it to the education and income of its residents. However, a given state may also be influenced by the adoption (or abandonment) of a name in high or low culture (or income) states. This influence can help in the spread of names across states, and explain the rise and fall of fashion cycles (see §7.2.2). So we now specify and estimate a model where we examine the impact of the difference in the adoption levels in the most and least cultured (and wealthy) states on the rest of the states.

7.2.1 Model and Results

Let $\{j_{1t}^w, j_{2t}^w, j_{3t}^w\}$ and $\{j_{4t}^w, j_{5t}^w, j_{6t}^w\}$ denote the three most and least wealthy states based on adjusted median income in period t. Let $d_{it-1}^m = \frac{1}{3} \left(\sum_{k=j_{1t}^w}^{j_{3t}^w} f_{ikt-1} - \sum_{k=j_{4t}^w}^{j_{6t}^w} f_{ikt-1} \right)$ be the difference between the mean popularity of name i in the most and least wealthy states in $t-1$. Similarly, let $d_{it-1}^c = \frac{1}{3} \left(\sum_{k=j_{1t}^c}^{j_{3t}^c} f_{ikt-1} - \sum_{k=j_{4t}^c}^{j_{6t}^c} f_{ikt-1} \right)$ be the differences between the mean popularity of name i in the most and least cultured states, based on c_{jt}. We use fractions instead of absolute numbers to control for differences in state populations.

Let y_{ijt} denote the popularity of name i in state j in period t, where:

$$y_{ijt} = const. + \sum_{k=1}^{p} \phi_k y_{ijt-k} + \kappa_w d_{it-1}^w + \kappa_c d_{it-1}^c + \alpha_1 x_{ijt}^1 + \alpha_2 x_{ijt}^2 + \beta z_i + \gamma_{ij} + \epsilon_{ijt} \tag{16}$$

The interpretations of y_{ijt}, γ_{ij}, w_{jt}, c_{jt}, x_{ijt}^1, x_{ijt}^2, and z_i remain the same as in §7.1.1. As before, x_{ijt}^1 is treated as potentially endogenous in our estimation. d_{it-1}^w and d_{it-1}^c are unlikely to be correlated with either γ_{ij}, ϵ_{ijt}, or ϵ_{ijt-1} because they are difference metrics. The common country level preference for name i (say γ_i) is differenced out, as is any common time-varying shock ϵ_{it-1}. This allows us to treat $\{d_{it-1}^w, d_{it-1}^c\}$ and $\{\Delta d_{it-1}^w, \Delta d_{it-1}^c\}$ as exogenous variables in our estimation. As before, we list instruments used for all the models estimated in results tables.

Table 11 presents the results. Model P1 is estimated on the Top50 dataset. We find that the effect of d_{it-1}^c is positive and significant, which implies that names popular in the highly educated states and unpopular in the least educated states are more likely to be adopted by the rest of the population. This is consistent with the theory that fashion is a signal of cultural capital because parents' incentive

to adopt a name is increasing (decreasing) in the number of adoptions by the cultured (uncultured) states. However, in both models, d_{it-1}^w is insignificant, *i.e.*, we do not find any evidence in support of the wealth signaling theory.

Note that these results go over and beyond those in §7.1 in ruling out some alternative explanations. For instance, some prior work on name choice suggests that certain parents prefer unique names (Lieberson and Bell, 1992; Twenge et al., 2010). If preferences for education and uniqueness are correlated, then we would find that educated adopt unique names, which can potentially explain the positive interaction effect between education and past popularity. This positive interaction effect can also be explained using novelty-based explanation; *e.g.*, educated people may prefer to be on the cutting edge (use novel names) and less educated people may prefer to not be on the cutting edge. However, neither of these alternatives can explain the finding that adoption by high education states has a positive impact on others' adoption, and adoption by low education types has a negative impact on others' adoption (after controlling for name popularity, *i.e.*, uniqueness or novelty). These two findings are instead consistent with a vertical signaling based explanation.

7.2.2 Deconstructing a Fashion Cycle

Finally, we combine Models N1 and P1 into Model P2 (see Table 11). The patterns from this model allow us deconstruct culture-based fashion cycles: at the beginning of the cycle, when the name has not been adopted by anyone, the overall impact of c_{jt} is positive, which implies that cultured parents are more likely to adopt the name. This effect in turn gives rise to a situation, where the number of cultured parents who have adopted the name is higher than the number of uncultured parents who have adopted it, *i.e.*, $d_{it-1}^c > 0$. This increases the probability of adoption among everyone, but has a larger impact on the cultured parents at the beginning of the cycle (because the overall impact of c_{jt} is positive for low values of n_{it-1}). However, in time, when enough people have adopted the name (n_{it-1} is high), the cumulative impact of c_{jt} becomes negative. That is, cultured parents start abandoning the name, while the uncultured ones continue to adopt it. This in turn gives rise to a situation where the fraction of cultured parents who have adopted the name is lower than the number of uncultured parents, *i.e.*, $d_{it-1}^c < 0$. This dampens the adoption of the name among the entire population, with the

36

dampening effect being higher for the more cultured parents. This in turn pushes the name into a downward spiral, thereby ending the cycle.

7.2.3 Robustness Checks

We now present some specification checks to confirm the robustness of these results. First, we find that the results are robust to changes in the data used. In Model P3, we re-estimate the model with the Top100 dataset and find that the results are qualitatively similar (see Table 12). Next, in Model P4, instead of d_{it-1}^w and d_{it-1}^c, we consider $d_{it-1,t-2}^w$ and $d_{it-1,t-2}^c$, the differences in the mean adoptions between the three most and least wealthy (educated) states in years $t-1$ and $t-2$. The results in this model are similar to those from earlier models. In sum, we find that the results are robust to changes in model specification and data used.

8 Managerial Implications

Our findings have implications for marketing managers in the fashion industry. First, they provide an empirical framework to identify the drivers of fashion cycles in conspicuously consumed product categories. Second, they suggest that fashion should be seeded with consumers at the forefront of fashion cycles. For example, if consumers are interested in signaling cultural capital, the firms in that market should seed the product among the culturally savvy first. Over the last few years, seeding information with influentials has become a popular strategy among firms selling conspicuous goods; *e.g.*, Ford hired 100 social media savvy video bloggers to popularize its Fiesta car (Barry, 2009; Greenberg, 2010). However, finding effective seeds is a time-consuming and costly activity. In contrast, our findings suggest that even simple geography-based heuristics (at the state-level) can be used by fashion firms to find seeds. Interestingly, our findings also have implications for constraining market expansion. For example, fashion firms may want to withhold the product from low cultural capital consumers to keep the fashion cycle from dying too quickly by avoiding certain geographies.

Our main empirical framework is fairly general. It can be easily adapted to datasets from commercial settings. The model, as specified in Equation (15), can be modified to accommodate such data as follows: (a) include endogenous location-specific firm-level variables such as own price,

advertisement, and promotions into x^1_{ijt}; (b) capture the effect of past performance of the product both locally and globally using lagged dependent variables (y_{ijt}s and y_{it}s); (c) control for the effect of past competitive response and the effect of competitors' current prices and promotional strategies by including them as explanatory variables. Competition effects can be modeled as either endogenous or exogenous depending on the industry dynamics. If these are not observed by the researcher, they can be modeled as unobservables too; (d) control for the effect of industry-level trends and location specific factors using exogenous time-varying factors by including them in x^2_{ijt}. The model can then be estimated using the GMM-panel estimator discussed in §6.2.2 that controls for endogenous explanatory variables – this would be especially useful in the case of commercial products since we would expect prices and advertising expenditures to be correlated to unobserved product quality. Finally, as shown in Appendix B.2, our estimation framework can also be deconstructed and made to work for individual-level data. This can be useful if the firm has detailed information on its consumers, especially over multiple years.

9 Conclusion

Fashions and conspicuous consumption play an important role in marketing. However, empirical work on fashions is close to non-existent and we have no formal frameworks to identify the presence of fashion cycles in data or examine their drivers. In this paper we bridge this gap in the literature. First, we present algorithmic and statistical methods to identify the presence of cycles. In this context, we introduce the Conditional Monotonicity Property, and explain its role in giving rise to cycles. We also show how system GMM estimators can help researchers overcome potential endogeneity concerns and derive consistent estimates to establish the presence of cycles in data. Second, we apply our framework to the name-choice context and establish the presence of cycles in data. Third, we examine the potential drivers of fashion cycles in this setting, especially the two signaling theories of fashion. By exploiting longitudinal and geographical variations in parents' cultural and economic capital, we show that naming patterns are consistent with the cultural capital theory.

In sum, our paper makes two key contributions to the literature on fashion and conspicuous

consumption. First, from a methodological perspective, we present an empirical framework to identify the presence and cause of fashion cycles in data. Our method is applicable to a broad range of settings, wherein managers and researchers need to detect the presence of fashion cycles and examine their drivers. Second, from a substantive perspective, we establish the presence of large amplitude fashion cycles in names choice decisions and show that the patterns of these cycles are consistent with Bourdieu's cultural capital signaling theory.

Our analysis suffers from limitations that serve as excellent avenues for future research. First, the context of our data may not be best to examine the theories of fashion, especially the wealth signaling theory since names are costless. Thus, the magnitude and directionality of Bourdieu and Veblen effects are specific to our setting. Recall that given names are unique – they are not influenced by commercial concerns (advertisements, promotions, and so on) and are free (zero-price) for all potential adopters. This makes it difficult to extrapolate our point estimates to other commercial settings. Second, because we only have state-level data, our analysis is silent on within-state effects. It is possible that other types of peer effects are at play within smaller geographic areas that we miss in our across-state analysis. Analyzing and documenting such effects would be a useful next step.

We conclude with the observation that while fashion is an important driver of consumption in the modern society, it remains an under-studied topic in marketing. We hope that the empirical methods and substantive findings presented in this paper will encourage other researchers to undertake empirical studies of fashions in the future.

A Appendix

A.1 Proof of Proposition 1

Consider a non-stationary AR(p) process with roots $1, \frac{1}{c_1}, ..., \frac{1}{c_{p-1}}$, where $p \geq 2$ and $0 < c_1, ..., c_{p-1} \leq 1$. We know that y_{it} can be expressed as:

$$(1-L)(1-c_1 L)...(1-c_{p-1}L)y_{it} = \epsilon_{it} \tag{A-1}$$

$$\Rightarrow E[(1-c_1 L)...(1-c_{p-2}L)(1-c_{p-1}L)\Delta y_{it}] = 0 \tag{A-2}$$

$$\Rightarrow E[(1-c_1 L)...(1-c_{p-2}L)\Delta y_{it-1}] = c_{p-1}L(1-c_1 L)...(1-c_{p-2}L)\Delta y_{it} \tag{A-3}$$

$$\Rightarrow E\left[\left(\prod_{k=1}^{p-2}(1-c_k L)\right)\Delta y_{it-1}\right] = c_{p-1}\left(\prod_{k=1}^{p-2}(1-c_k L)\right)\Delta y_{it-1} \tag{A-4}$$

If $\left(\prod_{k=1}^{p-2}(1-c_k L)\right)\Delta y_{it-1} \geq 0$, then we have:

$$E[(1-c_1 L)...(1-c_{p-2}L)\Delta y_{it-1}] \geq 0 \qquad (A\text{-}5)$$

since $c_{p-1} > 0$ and $\left(\prod_{k=1}^{p-2}(1-c_k L)\right)\Delta y_{it-1}$.
If, on the other hand, $\left(\prod_{k=1}^{p-2}(1-c_k L)\right)\Delta y_{it-1} \leq 0$, then we have:

$$E[(1-c_1 L)...(1-c_{p-2}L)\Delta y_{it-1}] \leq 0 \qquad (A\text{-}6)$$

since $c_{p-1} > 0$ and $\left(\prod_{k=1}^{p-2}(1-c_k L)\right)\Delta y_{it-1} \leq 0$. Q.E.D.

Tables

	Top50		Top100		Top200		Top500	
	Female	Male	Female	Male	Female	Male	Female	Male
No. of names	218	143	366	275	648	488	1468	1115
Perc. of births	40.63	44.12	48.25	51.94	54.71	57.28	60.99	61.67

Table 1: Number of male and female names and the percentage of births (1940-2009) corresponding to these names.

Year	Top 5 States		Bottom 5 States	
	State	w_{jt}	State	w_{jt}
1970	Maryland	647.8	South Dakota	383.5
	Alaska	627	Maine	403.5
	California	621.6	Arkansas	408
	Michigan	615.3	Vermont	408
	Texas	588.5	Mississippi	411.2
1980	Alaska	601.1	South Dakota	342.3
	Maryland	565	Maine	351
	Michigan	510.7	Vermont	364.1
	Virginia	507.5	Mississippi	369.4
	Texas	498.9	Arkansas	371
1990	Alaska	672	South Dakota	383.8
	Maryland	646.3	North Dakota	389.4
	Virginia	580.5	Mississippi	393.6
	Delaware	572.7	West Virginia	393.7
	New Jersey	551.9	Arkansas	407.3
2000	Maryland	677.6	West Virginia	407.3
	Alaska	645.3	Kentucky	413.5
	Minnesota	636.4	Louisiana	422.2
	Delaware	626.6	Maine	423.5
	Virginia	607	Montana	432.1

Table 2: Top and bottom five wealthiest states based on adjusted median household income.

Year	Top 5 States		Bottom 5 States	
	State	c_{jt}	State	c_{jt}
1970	Colorado	4.2	Arkansas	-4.0
	Alaska	3.4	West Virginia	-3.9
	Utah	3.3	Kentucky	-3.5
	Hawaii	3.3	Alabama	-2.9
	Maryland	3.2	Tennessee	-2.9
1980	Colorado	6.7	West Virginia	-5.8
	Alaska	4.9	Arkansas	-5.4
	Connecticut	4.5	Kentucky	-5.1
	Maryland	4.2	Alabama	-4.0
	Hawaii	4.2	Mississippi	-3.9
1990	Massachusetts	6.9	West Virginia	-8.0
	Connecticut	6.9	Arkansas	-7.0
	Colorado	6.7	Kentucky	-6.7
	Maryland	6.2	Mississippi	-5.6
	New Jersey	4.6	Nevada	-5.0
2000	Massachusetts	8.8	West Virginia	-9.6
	Colorado	8.3	Arkansas	-7.7
	Maryland	7.1	Mississippi	-7.5
	Connecticut	7.0	Kentucky	-7.3
	New Jersey	5.4	Nevada	-6.2

Table 3: Top and bottom five states based on education *i.e.*, percentage of adults with bachelors degree (c_{jt}).

Rank	1940		1950		1960		1970	
	Female	Male	Female	Male	Female	Male	Female	Male
1	Mary	James	Linda	James	Mary	David	Jennifer	Michael
2	Barbara	Robert	Mary	Robert	Susan	Michael	Lisa	James
3	Patricia	John	Patricia	John	Linda	James	Kimberly	David
4	Judith	William	Barbara	Michael	Karen	John	Michelle	John
5	Betty	Richard	Susan	David	Donna	Robert	Amy	Robert
6	Carol	Charles	Nancy	William	Lisa	Mark	Angela	Christopher
7	Nancy	David	Deborah	Richard	Patricia	William	Melissa	William
8	Linda	Thomas	Sandra	Thomas	Debra	Richard	Tammy	Brian
9	Shirley	Donald	Carol	Charles	Cynthia	Thomas	Mary	Mark
10	Sandra	Ronald	Kathleen	Gary	Deborah	Steven	Tracy	Richard

Rank	1980		1990		2000		2009	
	Female	Male	Female	Male	Female	Male	Female	Male
1	Jennifer	Michael	Jessica	Michael	Emily	Jacob	Isabella	Jacob
2	Amanda	Christopher	Ashley	Christopher	Hannah	Michael	Emma	Ethan
3	Jessica	Jason	Brittany	Matthew	Madison	Matthew	Olivia	Michael
4	Melissa	David	Amanda	Joshua	Ashley	Joshua	Sophia	Alexander
5	Sarah	James	Samantha	Daniel	Sarah	Christopher	Ava	William
6	Heather	Matthew	Sarah	David	Alexis	Nicholas	Emily	Joshua
7	Nicole	Joshua	Stephanie	Andrew	Samantha	Andrew	Madison	Daniel
8	Amy	John	Jennifer	James	Jessica	Joseph	Abigail	Jayden
9	Elizabeth	Robert	Elizabeth	Justin	Elizabeth	Daniel	Chloe	Noah
10	Michelle	Joseph	Lauren	Joseph	Taylor	Tyler	Mia	Anthony

Table 4: Top 10 female and male names in 1940, 1950, 1960, 1970, 1980, 1990, 2000, and 2009.

No. of cycles	Top50		Top100		Top200		Top500	
	Female	Male	Female	Male	Female	Male	Female	Male
0	16.9	24.5	18.6	25.8	20.2	28.1	24.8	36.6
1	53.2	45.4	54.1	40.3	52.9	38.9	52.9	40.4
2	21.6	23.8	21.0	26.2	21.4	25.4	18.3	19.2
3	6.9	6.3	5.5	7.3	4.8	7.2	3.5	3.6
4	0.9	0	0.5	0.4	0.5	0.4	0.4	0.2
5	0.5	0	0.3	0	0.2	0	0.1	0
Total Percentage	100	100	100	100	100	100	100	100
No. of names	218	143	366	275	648	488	1468	1115

Table 5: Percentage of names with 0, 1, 2, 3, 4, and 5 cycles.

Pattern	Top50		Top100		Top200		Top500	
	Female	Male	Female	Male	Female	Male	Female	Male
/ or \	16.9	24.5	18.6	25.8	20.2	28.1	24.8	36.6
/\	26.1	22.4	24.6	16.4	18.2	13.3	12.1	8.5
\/	8.7	7.7	10.9	10.2	14.8	13.1	17.4	16.7
/\/	3.7	1.4	3.3	1.1	3.1	1.8	4.6	3.9
\/\	17.4	11.2	16.7	13.1	18.2	12.7	22.8	16.3
/\/\	13.3	17.5	11.5	16.4	10.0	14.8	8.0	10.0
\/\/	4.6	4.2	7.1	5.8	7.9	5.5	8.6	7.3
Other	9.3	11.1	7.3	11.2	7.6	10.7	1.7	0.7
Total Percentage	100	100	100	100	100	100	100	100
No. of names	218	143	366	275	648	488	1468	1115

Table 6: Cycle patterns in data.

Quartiles	Top50		Top100		Top200		Top500	
	Female	Male	Female	Male	Female	Male	Female	Male
25	2	2	2	2	2	2	2	2
50	8	4	6	3	5	3	4	3
75	32	16	29	11	16	8	11	8

Table 7: Distribution of dormancy lengths between cycles.

	Top50		Top100		Top200		Top500	
	Female	Male	Female	Male	Female	Male	Female	Male
CO peaks first	92	62	159	73	283	242	668	489
WV peaks first	54	0	87	0	151	0	340	125

Table 8: Relative ordering of when names peak in Colorado and West Virginia.

Table 9: Estimation results and Conditional Monotonicity tests. Dependent variable is n_{it}.

Variable		Model M1 Estimate	Std. Err.	Model M2 Estimate	Std. Err.	Model M3 Estimate	Std. Err.	Model M4 Estimate	Std. Err.
L.Number	n_{it-1}	0.9875***	0.0488	0.9893***	0.0197	0.9956***	0.0062	0.9958***	0.0042
L.ΔNumber	Δn_{it-1}	0.4785**	0.2004	0.4739***	0.1348	0.4708***	0.1154	0.4675***	0.0728
L2.ΔNumber	Δn_{it-2}	0.0630	0.1684	0.0580	0.1540	0.0479	0.0902	0.0451	0.0582
Length	l_i	10.849	18.661	3.4902	37.249	0.8112	3.4730	0.1787	0.3737
Bible	bib_i	0.3207	1.7048	0.1634	0.8375	0.0570	0.1353	0.0083	0.0418
Sex	s_i	-161.0	317.13	-13.31	2887.0	-3.2654	187.19	2.1934	48.589
Total babies	$\Gamma_{s_i t}$	-5.5e-4	0.0011	-4.5e-6	0.0133	-8.7e-7	0.0009	1.5e-5	0.0002
Const.	k	873.80	3283.2	31.656	25551	29.084	2346.8	-19.81	353.94
Time dummies		Yes		Yes		Yes		Yes	
Cond. Monotonicity Test									
H1: $\mu = 1$		Do not reject		Do not reject		Do not reject		Do not reject	
z-stat (p-value)		-0.256 (0.401)		-0.543 (0.295)		-0.710 (0.239)		-0.970 (0.166)	
H2: $\theta_1 = 0.47$		Do not reject		Do not reject		Do not reject		Do not reject	
z-stat (p-value)		0.042 (0.516)		0.029 (0.512)		0.007 (0.504)		-0.034 (0.488)	
AR-Bond (2) Test		Do not reject		Do not reject		Do not reject		Do not reject	
Test statistic (p-value)		-0.825 (0.409)		-0.816 (0.414)		-1.126 (0.260)		-1.524 (0.128)	
Corr. $(y, \hat{y})$		0.9962		0.9963		0.9964		0.9966	
Root MSE		887.9		691.3		533.2		358.5	
MAE		392.2		268.8		178.5		92.88	
Diff. Eqn. Instr.	GMM	L(2/3).{$n_{it},\Delta n_{it}$}		L(2/3).{$n_{it},\Delta n_{it}$}		L(2/3).{$n_{it},\Delta n_{it}$}		L(2/3).{$n_{it},\Delta n_{it}$}	
	Standard	$\Delta\Gamma_{s_i t}$		$\Delta\Gamma_{s_i t}$		$\Delta\Gamma_{s_i t}$		$\Delta\Gamma_{s_i t}$	
Level Eqn. Instr.	GMM	LΔ{$n_{it},\Delta n_{it}$}		LΔ{$n_{it},\Delta n_{it}$}		LΔ{$n_{it},\Delta n_{it}$}		LΔ{$n_{it},\Delta n_{it}$}	
	Standard	$s_i, l_i, u_i, \Gamma_{s_i t}$		$s_i, l_i, u_i, \Gamma_{s_i t}$		$s_i, l_i, u_i, \Gamma_{s_i t}$		$s_i, l_i, \Gamma_{s_i t}$	
No. of names, years		361, 67		641, 67		1136, 67		2583, 67	
Dataset used		Top50		Top100		Top200		Top500	

Note: *** $\Rightarrow p \leq 0.01$, ** $\Rightarrow p \leq 0.05$, * $\Rightarrow p \leq 0.1$

44

Variable		Model N1		Model N2		Model N3	
		Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Lagged Dep. Var.	n_{ijt-1}	0.9967***	0.2547×10^{-2}	0.9946***	0.2017×10^{-2}	0.9958***	0.2648×10^{-2}
	Δn_{ijt-1}	0.6498×10^{-1}***	0.1418×10^{-1}	0.3144×10^{-1}***	0.8188×10^{-2}	0.6441×10^{-1}***	0.1144×10^{-1}
	Δn_{ijt-2}	0.1484***	0.8726×10^{-2}	0.1398***	0.6794×10^{-2}	0.1475***	0.8648×10^{-2}
	Δn_{ijt-3}	0.1210***	0.5229×10^{-2}	0.1237***	0.4429×10^{-2}	0.1202***	0.5617×10^{-2}
	Δn_{ijt-4}	0.8465×10^{-1}***	0.6763×10^{-2}	0.8733×10^{-1}***	0.4787×10^{-2}	0.8458×10^{-1}***	0.6648×10^{-2}
Name Char.	l_i	0.1630***	0.2888×10^{-1}	0.7061×10^{-1}***	0.1506×10^{-1}	0.1604***	0.2810×10^{-1}
	bib_i	0.5866×10^{-2}***	0.2144×10^{-2}	0.2479×10^{-2}***	0.1237×10^{-2}	0.5869×10^{-2}***	0.2050×10^{-2}
	s_i	-0.5408***	0.7784×10^{-1}	-0.2858***	0.3995×10^{-1}	-0.5447***	0.7676×10^{-1}
Cultural Capital	c_{jt}	0.7749×10^{-1}***	0.2965×10^{-1}	0.3851×10^{-1}***	0.1620×10^{-1}	0.5425×10^{-1}**	0.3062×10^{-1}
	$c_{jt}.n_{it-1}$	-0.2780×10^{-4}***	0.7630×10^{-5}	-0.2410×10^{-4}***	0.6070×10^{-5}	-0.2290×10^{-4}***	0.7860×10^{-5}
Economic Capital	w_{jt}	-0.2757×10^{-1}***	0.1656×10^{-2}	-0.1674×10^{-1}***	0.9275×10^{-3}	-0.2679×10^{-1}***	0.1827×10^{-2}
	$w_{jt}.n_{it-1}$	0.9110×10^{-5}***	0.3780×10^{-6}	0.8630×10^{-5}***	0.3210×10^{-6}	0.9260×10^{-5}***	0.4140×10^{-6}
Other	n_{it-1}	0.9978×10^{-3}***	0.2708×10^{-3}	0.1766×10^{-2}***	0.1880×10^{-3}	0.1097×10^{-2}***	0.2297×10^{-3}
	n_{it-2}	-0.5499×10^{-2}***	0.2197×10^{-3}	-0.5974×10^{-2}***	0.1451×10^{-3}	-0.5616×10^{-2}***	0.1845×10^{-3}
	Γ_{s_ijt}	0.4880×10^{-5}	0.4830×10^{-5}	0.8720×10^{-5}***	0.2530×10^{-5}	0.5700×10^{-5}	0.4880×10^{-5}
	Const.	0.1335×10^{2}***	0.8790	0.8079×10^{1}***	0.4777	0.1286×10^{2}***	0.9775

AR-Bond (2) Test Test stat (p-value)	Do not reject -0.8918 (0.3725)	Do not reject -1.3415 (0.1798)	Do not reject -0.9758 (0.3292)
No. of names, states, years	361, 50, 34	641, 50, 34	361, 49, 34
Dataset used	Top50	Top100	Top50

Diff. Eqn. Instr.	GMM	L(2/4)$[n_{ijt}, \Delta n_{ijt}, n_{it}, c_{jt}.n_{it-1}, w_{jt}.n_{it-1}]$
	Standard	$\Delta[\Gamma_{s_ijt}, c_{jt}, w_{jt}]$
Level Eqn. Instr	GMM	L2 $\Delta[n_{ijt}, \Delta n_{ijt}, n_{it}, c_{jt}.n_{it-1}, w_{jt}.n_{it-1}]$
	Standard	$[s_i, l_i, bib_i, \Gamma_{s_ijt}, c_{jt}, w_{jt}]$

Note: *** $\Rightarrow p \leq 0.01$, ** $\Rightarrow p \leq 0.05$, * $\Rightarrow p \leq 0.1$

Table 10: Interacting wealth and cultural capital with past adoptions. Dependent variable is n_{ijt}.

	Variable	Model P1		Model P2	
		Estimate	Std. Err.	Estimate	Std. Err.
Lagged Dep. Var.	n_{ijt-1}	$0.1006\times10^{1***}$	0.4257×10^{-2}	0.9965^{***}	0.1970×10^{-2}
	Δn_{ijt-1}	$0.6911\times10^{-1***}$	0.2291×10^{-1}	$0.6383\times10^{-1***}$	0.8756×10^{-2}
	Δn_{ijt-2}	0.1511^{***}	0.2151×10^{-1}	0.1499^{***}	0.7894×10^{-2}
	Δn_{ijt-3}	0.1171^{***}	0.8039×10^{-2}	0.1216^{***}	0.4859×10^{-2}
	Δn_{ijt-4}	$0.8363\times10^{-1***}$	0.1602×10^{-1}	$0.8450\times10^{-1***}$	0.5837×10^{-2}
Name Char.	l_i	0.2386^{***}	0.4218×10^{-1}	0.11081^{***}	0.2965×10^{-1}
	bib_i	$0.1031\times10^{-1***}$	0.2202×10^{-2}	$0.46011\times10^{-2***}$	0.2211×10^{-2}
	s_i	-0.7938^{***}	0.1151	-0.4897^{***}	0.7873×10^{-1}
Cultural Capital	c_{jt}			$0.8071\times10^{-1***}$	0.2752×10^{-1}
	$c_{jt}.n_{it-1}$			$-0.2740\times10^{-4***}$	0.6950×10^{-5}
	d^c_{it-1}	$0.2076\times10^{3***}$	0.4423×10^{2}	$0.1465\times10^{3***}$	0.4699×10^{2}
Economic Capital	w_{jt}			$-0.2821\times10^{-1***}$	0.1621×10^{-2}
	$w_{jt}.n_{it-1}$			$0.9290\times10^{-5***}$	0.3660×10^{-6}
	d^w_{it-1}	0.5278×10^{2}	0.7038×10^{2}	$0.2448\times10^{3***}$	0.5875×10^{2}
Other	n_{it-1}	$0.4446\times10^{-2***}$	0.2577×10^{-3}	$0.9712\times10^{-3***}$	0.2059×10^{-3}
	n_{it-2}	$-0.4874\times10^{-2***}$	0.2627×10^{-3}	$-0.5523\times10^{-2***}$	0.1509×10^{-3}
	Γ_{s_ijt}	$-0.1660\times10^{-4***}$	0.8390×10^{-5}	0.5430×10^{-5}	0.4090×10^{-5}
	Const.	0.9494^{***}	0.3415	$0.1379\times10^{2***}$	0.8359

AR-Bond (2) Test		Do not reject		Do not reject	
Test stat (p-value)		-.5381 (0.5905)		-0.8741 (0.3821)	
Diff. Eqn. Instr.	GMM	L(2/4) $[n_{ijt},\Delta n_{ijt},n_{it}]$		L(2/4) $[n_{ijt},\Delta n_{ijt},n_{it},c_{jt}.n_{it-1},w_{jt}.n_{it-1}]$	
	Standard	$\Delta[\Gamma_{s_ijt},d^c_{it-1},d^w_{it-1}]$		$\Delta[\Gamma_{s_ijt},c_{jt},w_{jt},d^c_{it-1},d^w_{it-1}]$	
Level Eqn. Instr.	GMM	L2$\Delta[n_{ijt},\Delta n_{ijt},n_{it}]$		L2$\Delta[n_{ijt},\Delta n_{ijt},n_{it},c_{jt}.n_{it-1},w_{jt}.n_{it-1}]$	
	Standard	$[s_i,l_i,bib_i,\Gamma_{s_ijt},d^c_{it-1},d^w_{it-1}]$		$[s_i,l_i,bib_i,\Gamma_{s_ijt},c_{jt},w_{jt},d^c_{it-1},d^w_{it-1}]$	
No. of names, states, years		361, 50, 35		361, 50, 34	
Dataset used		Top50		Top50	

Note: $^{***} \Rightarrow p \leq 0.01$, $^{**} \Rightarrow p \leq 0.05$, $^{*} \Rightarrow p \leq 0.1$

Table 11: Impact of adoption by high and low types. Dependent variable is n_{ijt}.

Table 12 (rotated).

Variable		Model P3 Estimate	Model P3 Std. Err.	Model P4 Estimate	Model P4 Std. Err.
Lagged Dep. Var.	n_{ijt-1}	0.9943***	0.2052×10^{-2}	0.9965***	0.2372×10^{-2}
	Δn_{ijt-1}	0.3056×10^{-1}***	0.8447×10^{-2}	0.6405×10^{-1}***	0.1241×10^{-1}
	Δn_{ijt-2}	0.1412***	0.6291×10^{-2}	0.1497***	0.8586×10^{-2}
	Δn_{ijt-3}	0.1242***	0.4170×10^{-2}	0.1216***	0.5043×10^{-2}
	Δn_{ijt-4}	0.8726×10^{-1}***	0.5232×10^{-2}	0.8503×10^{-1}***	0.6330×10^{-2}
Name Char.	l_i	0.4663×10^{-1}***	0.1550×10^{-1}	0.1105***	0.2983×10^{-1}
	bib_i	0.2126×10^{-2}*	0.1221×10^{-2}	0.4570×10^{-2}**	0.2198×10^{-2}
	s_i	-0.2762***	0.4047×10^{-1}	-0.4922***	0.7896×10^{-1}
Cultural Capital	c_{jt}	0.3711×10^{-1}***	0.1643×10^{-1}	0.7730×10^{-1}***	0.2902×10^{-1}
	$c_{jt}.n_{it-1}$	-0.2250×10^{-4}**	0.6160×10^{-5}	-0.2690×10^{-4}**	0.7410×10^{-5}
	d^c_{it-1}				
	$d^c_{it-1,t-2}$	0.1836×10^{3}***	0.4352×10^{2}	0.1216×10^{3}***	0.4749×10^{2}
Economic Capital	w_{jt}	-0.1693×10^{-1}***	0.9309×10^{-3}	-0.2832×10^{-1}***	0.16449×10^{-2}
	$w_{jt}.n_{it-1}$	0.8740×10^{-5}***	0.3220×10^{-6}	0.9320×10^{-5}***	0.3750×10^{-6}
	d^w_{it-1}				
	$d^w_{it-1,t-2}$	0.2223×10^{3}***	0.5183×10^{2}	0.2520×10^{3}***	0.6142×10^{2}
Other	n_{it-1}	0.1754×10^{-2}***	0.1873×10^{-3}	0.9326×10^{-3}***	0.2371×10^{-3}
	n_{it-2}	-0.5983×10^{-2}***	0.1455×10^{-3}	-0.5499×10^{-2}***	0.1884×10^{-3}
	$\Gamma_{s,jt}$	0.9170×10^{-5}	0.2540×10^{-5}	0.5430×10^{-5}	0.4620×10^{-5}
	Const.	0.8217×10^{1}***	0.4788	0.1385×10^{2}***	0.8605

AR-Bond (2) Test Test stat (p-value)		Do not reject $-1.5606\ (0.1186)$		Do not reject $-0.9662\ (0.3340)$	
Diff. Eqn. Instr.	GMM	L(2/4) $[n_{ijt},\Delta n_{ijt},n_{it},c_{jt}.n_{it-1},w_{jt}.n_{it-1}]$		L(2/4) $[n_{ijt},\Delta n_{ijt},n_{it},c_{jt}.n_{it-1},w_{jt}.n_{it-1}]$	
	Standard	$\Delta[s_{ijt},c_{jt},w_{jt},d^c_{it-1},d^w_{it-1}]$		$\Delta[\Gamma_{s,jt},c_{jt},w_{jt},d^c_{it-1,t-2}]$	
Level Eqn. Instr.	GMM	L2$\Delta[n_{ijt},c_{jt}.n_{it},c_{jt}.n_{it-1},w_{jt}.n_{it-1}]$		L2$\Delta[n_{ijt},\Delta n_{ijt},n_{it},c_{jt}.n_{it-1},w_{jt}.n_{it-1}]$	
	Standard	$[s_i,l_i,bib_i,\Gamma_{s,jt},c_{jt},w_{jt},d^c_{it-1},d^w_{it-1}]$		$[s_i,l_i,bib_i,\Gamma_{s,jt},c_{jt},w_{jt},d^c_{it-1,t-2},d^w_{it-1,t-2}]$	
No. of names, states, years		$641, 50, 34$		$361, 50, 34$	
Dataset used		Top100		Top50	

Note: *** $\Rightarrow p \leq 0.01$, ** $\Rightarrow p \leq 0.05$, * $\Rightarrow p \leq 0.1$

Table 12: Impact of adoption by high and low types (contd.); dependent variable is n_{ijt}.

References

D. Acemoglu and J. A. Robinson. A Theory of Political Transitions. *The American Economic Review*, 91(4): 938–963, 2001.

W. Amaldoss and S. Jain. Pricing of Conspicuous Goods: A Competitive Analysis of Social Effects. *Journal of Marketing Research*, 42(1):30–42, 2005.

M. Arellano and S. Bond. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(2):277–97, 1991.

L. W. Banner. American Beauty: A Social History Through Two Centuries of the American Idea, Ideal, and Image of the Beautiful Woman. *New York: Alfred A. Knopf*, 1983.

K. Barry. Ford Bets the Fiesta on Social Networking. *Wired*, 2009.

M. Baxter and R. G. King. Measuring Business Cycles: Approximate Band-pass Filters for Economic Time Series. *Review of Economics and Statistics*, 81(4):575–593, 1999.

J. Berger and G. Le Mens. How Adoption Speed Affects the Abandonment of Cultural Tastes. *Proceedings of the National Academy of Sciences*, 106(20):8146–8150, 2009.

W. D. Berry, R. C. Fording, and R. L. Hanson. An Annual Cost of Living Index for the American States, 1960-95. *Journal of Politics*, 60(2):550–567, 2000.

M. Bertrand and S. Mullainathan. Are Emily and Greg more Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *The American Economic Review*, 94(4), 2004.

P. Besnard. The Study of Social Taste Through First Names: Comment on Lieberson and Bell. *The American Journal of Sociology*, 100(5):1313–1317, 1995.

H. Blumer. Fashion: From Class Differentiation to Social Selection. *The Sociological Quarterly*, 10:275–291, 1969.

R. Blundell and S. Bond. Initial conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*, 87(1):115–143, 1998.

R. Blundell and T. Stoker. Heterogeneity and Aggregation. *Journal of Economic Literature*, 43(2):347–391, 2005.

S. Bond, C. Nauges, and F. Windmeijer. Unit Roots and Identification in Autoregressive Panel Data Models: A Comparison of Alternative Tests. In *International Conferences on Panel Data*, 2002.

P. Bourdieu. *Distinction: A Social Critique of the Judgement of Taste*. Harvard University Press, 1984.

A. F. Burns and W. C. Mitchell. Measuring Business Cycles. *NBER Books*, 1946.

C. Camerer. Bubbles and Fads in Asset Prices. *Journal of Economic Surveys*, 3(1):1–41, 1989.

W. W. Charemza and D. F. Deadman. Speculative Bubbles with Stochastic Explosive Roots: The Failure of Unit Root Testing. *Journal of Empirical Finance*, 2(2):153–163, 1995.

P. K. Chintagunta, S. Gopinath, and S. Venkataraman. The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets. *Marketing Science*, 29(5):944–957, 2010.

C. Clark, U. Doraszelski, and M. Draganska. The Effect of Advertising on Brand Awareness and Perceived Quality: An Empirical Investigation using Panel Data. *Quantitative Marketing and Economics*, 7(2): 207–236, 2009.

P. W. Cookson and C. H. Persell. *Preparing For Power: America's Elite Boarding Schools*. Basic Books, 1987.

G. Corneo and O. Jeanne. A Theory of Fashion Based on Segmented Communication. *Discussion Paper Serie A*, 462, 1994.

F. Davis. *Fashion, Culture, and Identity*. University of Chicago Press, 1994.

G. S. Day. The Product Life Cycle: Analysis and Applications Issues. *Journal of Marketing*, 45:60–67, 1981.

M. Dekimpe and D. Hanssens. The Persistence of Marketing Effects on Sales. *Marketing Science*, 14(1):1–21, 1995.

B. T. Diba and H. I. Grossman. Explosive Rational Bubbles in Stock Prices? *The American Economic Review*,

78(3):520–530, 1988.

P. Dimaggio and M. Useem. Social Class and Arts Consumption: The Origins and Consequences of Class Differences in Exposure to the Arts in America. *Theory and Society*, 5(2):141–161, 1974.

S. Durlauf, P. Johnson, and J. Temple. Growth Econometrics. In P. Aghion and S. Durlauf, editors, *Handbook of Economic Growth*, volume 1 of *Handbook of Economic Growth*, pages 555–677. Elsevier, 2005.

G. W. Evans. Pitfalls in Testing for Explosive Bubbles in Asset Prices. *American Economic Review*, 81(4): 922–30, 1991.

G. S. Frings. *Fashion: From Concept to Consumer*. Englewood Cliffs, NJ: Simon & Schuster, 1991.

R. G. Fryer and S. D. Levitt. The Causes and Consequences of Distinctively Black Names. *The Quarterly Journal of Economics*, 69(3), 2004.

W. A. Fuller. *Introduction to Statistical Time Series*. Wiley Interscience, 1995.

G. Gammage and S. L. Jones. Orgasm in Chrome: The Rise and Fall of the Automobile Tailfin. *Journal of Popular Culture*, 1974.

K. Greenberg. Ford Fiesta Movement Shifts into High Gear. *Marketing Daily*, 2010.

T. M. Gureckis and R. L. Goldstone. How You Named Your Child: Understanding the Relationship Between Individual Decision Making and Collective Outcomes. *Topics in Cognitive Science*, 1(4):651–674, 2009.

M. W. Hahn and R. A. Bentley. Drift as a Mechanism for Cultural Change: An Example from Baby Names. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270:S120–S123, 2003.

R. J. Hodrick and E. C. Prescott. Postwar US Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking, pages=1–16, year=1997, publisher=JSTOR*.

E. Karni and D. Schmeidler. Fixed Preferences and Changing Tastes. *The American Economic Review*, 80: 262–267, 1990.

T. Kisbye. Name-Borrowing Mechanisms. The Impact of English Masculine Personal Names on a Major Danish Town Community. In *Proceedings of the Thirteenth International Congress of Onomastic Sciences*, 1981.

M. Lamont and A. Lareau. Cultural Capital: Allusions, Gaps and Glissandos in Recent Theoretical Developments. *Sociological Theory*, 6(2):153–168, 1988.

S. Levitt and S. J. Dubner. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. William Morrow/HarperCollins, 2005.

T. Levitt. Exploit the Product Life Cycle. *Harvard Business Review*, 43, 1965.

S. Lieberson. *A Matter of Taste: How Names, Fashions, and Culture Change*. Yale University Press, 2000.

S. Lieberson and E. O. Bell. Children's First Names: An Empirical Study of Social Taste. *The American Journal of Sociology*, 98(3):511–554, 1992.

S. Lieberson and F. B. Lynn. Popularity as a Taste: An Application to the Naming Process. *Journal of the International Council of Onomastic Sciences*, 38:235–256, 2003.

P. V. Marsden. *Network diversity, Substructures and Opportunities for Contact*. Number 397-410. ed. C. M. Meyer, R. S. Scott, New York: Cambridge, 1990.

M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a Feather: Homophily in Social Networks. *Ann Rev Sociol 2001*, 27(415-444), 2001.

H. L. Mencken. *The American Language: An Inquiry into the Development of English Language in the United States*. Alfred A. Knopf, 1963.

C. M. Miller, S. H. McIntyre, and M. K. Mantrala. Toward Formalizing Fashion Theory. *Journal of Marketing Research*, pages 142–157, 1993.

H. Nair, P. Manchanda, and T. Bhatia. Asymmetric Social Interactions in Physician Prescription Behavior: The Role of Opinion Leaders. *Journal of Marketing Research*, 47(5):883–895, 2010.

S. Narayanan and H. Nair. Estimating Causal Installed Base Effects: A Bias Correction Approach. *Journal of Marketing Research*, 2012.

S. J. Nickell. Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6):1417–26, 1981.

W. Pesendorfer. Design Innovation and Fashion Cycles. *American Economic Review*, 85:771–792, 1995.

J. Richardson and A. L. Kroeber. *Three Centuries of Women's Dress Fashions: A Quantitative Analysis*, volume 5. University of California Press, 1940.

D. E. Robinson. Style Changes: Cyclical, Inexorable, and Foreseeable. *Harvard Business Review*, 53(6): 121–131, 1975.

R. Robinson and M. Garnier. Class Reproduction among Men and Women in France: Reproduction Theory on its Home Ground. *The American Journal of Sociology*, 91(2), 1985.

A. Rossi. Naming Children in Middle-class Families. *American Sociological Review*, 1965.

S. K. Shriver. Network Effects in Alternative Fuel Adoption: Empirical Analysis of the Market for Ethanol. *Marketing Science*, 2015.

S. K. Shriver, H. S. Nair, and R. Hofstetter. Social Ties and User-Generated Content: Evidence from an Online Social Network. *Management Science*, 59(6):1425–1443, 2013.

G. B. Sproles. Analyzing Fashion Life Cycles: Principles and Perspectives. *The Journal of Marketing*, pages 116–124, 1981.

T. Stoker. Empirical Approaches to the Problem of Aggregation over Individuals. *Journal of Economic Literature*, pages 1827–1874, 1993.

M. Sun, X. M. Zhang, and F. Zhu. To Belong or to Be Different? Evidence from a Large-Scale Field Experiments. Working Paper, 2014.

R. Taylor. John Doe, Jr.: A Study of his Distribution in Space, Time, and the Social Structure. *Social Forces*, 53 (1), 1974.

G. J. Tellis, R. Niraj, and E. Yin. Why and How Quality Wins over Network Effects and What it Means. *Journal of Marketing Research*, pages 135–149, 2009.

O. Toubia, J. Goldenberg, and R. Garcia. Improving Penetration Forecasts Using Social Interactions Data. *Management Science*, 60(12):3049–3066, 2014.

J. M. Twenge, E. M. Abebe, and W. K. Campbell. Fitting in or Standing Out: Trends in American Parents' Choices for Childrens Names, 1880–2007. *Social Psychological and Personality Science*, 1(1):19–25, 2010.

T. Veblen. *The Theory of the Leisure Class – An Economic Study of Institutions*. George Allen & Unwin Ltd, 1899.

F. Windmeijer. A Finite Sample Correction for the Variance of Linear Coeffcient Two-step GMM Estimators. *Journal of Econometrics*, 126(1):25–51, 2005.

E. G. Withycombe. *The Oxford Dictionary of English Christian Names*. Oxford University Press, 1977.

H. Yoganarasimhan. Cloak or Flaunt? The Fashion Dilemma. *Marketing Science*, 31(1):74–95, 2012a.

H. Yoganarasimhan. Impact of Social Network Structure on Content Propagation: A Study using YouTube Data. *Quantitative Marketing and Economics*, 10(1):111–150, 2012b.

B Online Appendix

B.1 Sensitivity Analysis to Varying τ and M

No. of cycles	$\tau=4, M=0.00005$		$\tau=5, M=0.00005$		$\tau=4, M=0.000075$		$\tau=4, M=0.0001$	
	Female	Male	Female	Male	Female	Male	Female	Male
0	18.6	25.8	18.0	28.4	21.6	28.4	37.9	37.1
1	54.1	40.3	57.7	43.6	54.1	40.4	41.5	35.6
2	21.0	26.2	20.5	22.9	18.3	24.7	15.9	22.9
3	5.5	7.3	3.5	5.1	5.2	6.2	4.2	4.4
4	0.5	0.4	0.3	0	0.8	0.3	0.5	0
5	0.3	0	0	0	0	0	0	0
Total Percentage	100	100	100	100	100	100	100	100
No. of names	366	275	366	275	366	275	366	275

Table 13: Percentage of cycles at different values of τ and M in the Top100 dataset.

B.2 Derivation of Aggregate Model from an Individual Level Model

B.2.1 Basic Model

Let individual q's probability of adopting name i at time t be:

$$y_{qit} = const. + \sum_{k=1}^{p} \phi_k y_{ijt-k} + \rho_1 w_{qt} + \rho_2 c_{qt} + \rho_3 w_{qt} y_{it-1} + \rho_4 c_{qt} y_{it-1} + \rho_5 x_{it}^1 + \rho_6 x_{it}^2 + \rho_7 z_i + \tau_{iqt} \quad \text{(B-1)}$$

The interpretation of the variables is similar to that in §7.1.1. The probability that agent q will adopt name i is a function of the past adoptions in her local neighborhood (state), her own wealth and cultural capital (w_{qt} and c_{qt}), interaction effects between her wealth, cultural capital and past adoptions by others, some endogenous factors that affect her affinity for name i (x_{ijt}^1 which includes total adoptions), some exogenous time-varying factors (x_{it}^2), time invariant name attributes (z_i), and an unobserved taste for name i (τ_{iqt}).

Summing Equation (B-1) over all potential adopters in state j at time t, and then dividing the resulting equation by the number of potential adopters, we have:

$$\bar{y}_{ijt} = const. + \sum_{k=1}^{p} \phi_k y_{ijt-k} + \rho_1 \bar{w}_{jt} + \rho_2 \bar{c}_{jt} + \rho_3 \bar{w}_{jt} y_{it-1} + \rho_4 \bar{c}_{qt} y_{it-1} + \rho_5 x_{it}^1 + \rho_6 x_{it}^2 + \rho_7 z_i + \bar{\tau}_{ijt} \quad \text{(B-2)}$$

Here, $\bar{w}_{jt}$ and $\bar{c}_{jt}$ are the mean wealth and cultural capital of state j at time t. $\bar{\tau}_{ijt}$ is the mean unobserved preference of potential adopters in state j, at time t, for name i. This can be rewritten as: $\bar{\tau}_{ijt} = \gamma_{ij} + e_{ijt}$, *i.e.*, we can extract out the mean preferences of residents of state j for name i and write the rest as a mean zero error term that varies with time. With these transformations, Equation (B-2) can be rewritten as:

$$\bar{y}_{ijt} = const. + \sum_{k=1}^{p} \phi_k y_{ijt-k} + \rho_1 \bar{w}_{jt} + \rho_2 \bar{c}_{jt} + \rho_3 \bar{w}_{jt} y_{it-1} + \rho_4 \bar{c}_{qt} y_{it-1} + \rho_5 x_{it}^1 + \rho_6 x_{it}^2 + \rho_7 z_i + \gamma_{ij} + e_{ijt} \quad \text{(B-3)}$$

This model is analogous to the aggregate model specified in §7.1.1, *i.e.*, all the parameter estimates from this aggregate model can be interpreted as individual level parameters with the right multipliers.

B.2.2 Expanded Model with within State Effects

We now expand the above model with within state effects. Let:

- y_{jit-1}^{hw} = the number of high wealth parents who have adopted name i in state j at time $t-1$.
- y_{jit-1}^{lw} = the number of low wealth parents who have adopted name i in state j at time $t-1$.
- y_{jit-1}^{hc} = the number of high culture parents who have adopted name i in state j at time $t-1$.
- y_{jit-1}^{lc} = the number of low culture parents who have adopted name i in state j at time $t-1$.

We now expand Equation (B-4) so that individual q's probability of adopting name i at time t is also affected by the number of high and low types that have adopted the name within state j:

$$y_{qit} = const. + \sum_{k=1}^{p} \phi_k y_{ijt-k} + \rho_1 w_{qt} + \rho_2 c_{qt} + \rho_3 w_{qt} y_{it-1} + \rho_4 c_{qt} y_{it-1} + \rho_5 x_{it}^1 + \rho_6 x_{it}^2 + \rho_7 z_i$$
$$+ \rho_8 w_{qt} y_{jit-1}^{hw} + \rho_9 w_{qt} y_{jit-1}^{lw} + \rho_{10} c_{qt} y_{jit-1}^{hc} + \rho_{11} c_{qt} y_{jit-1}^{lc} + \tau_{iqt} \tag{B-4}$$

Aggregating this over all potential adopters in state j, we have:

$$\bar{y}_{ijt} = const. + \sum_{k=1}^{p} \phi_k y_{ijt-k} + \rho_1 \bar{w}_{jt} + \rho_2 \bar{c}_{jt} + \rho_3 \bar{w}_{jt} y_{it-1} + \rho_4 \bar{c}_{qt} y_{it-1} + \rho_5 x_{it}^1 + \rho_6 x_{it}^2 + \rho_7 z_i$$
$$+ \rho_8 \bar{w}_{jt} y_{ijt-1}^{hw} + \rho_9 \bar{w}_{jt} y_{ijt-1}^{lw} + \rho_{10} \bar{c}_{jt} y_{ijt-1}^{hc} + \rho_{11} \bar{c}_{jt} y_{ijt-1}^{lc} + \gamma_{ij} + e_{ijt} \tag{B-5}$$

where $\bar{\tau}_{ijt} = \gamma_{ij} + e_{ijt}$, as before. Thus, the new error-term of the aggregate model can be written as:

$$\epsilon_{ijt} = \rho_8 \bar{w}_{jt} y_{ijt-1}^{hw} + \rho_9 \bar{w}_{jt} y_{ijt-1}^{lw} + \rho_{10} \bar{c}_{jt} y_{ijt-1}^{hc} + \rho_{11} \bar{c}_{jt} y_{ijt-1}^{lc} + e_{ijt} \tag{B-6}$$

First, note that much of the variation in these terms can be extracted out using lagged dependent variables (y_{ijt-1}s) and interaction effects of mean wealth/culture with state-level adoptions ($\bar{w}_{jt} y_{ijt-1}$s). For instance, we can rewrite the above equation as:

$$\epsilon_{ijt} = \rho_8 \bar{w}_{jt} y_{ijt-1} + \bar{w}_{jt}(\rho_9 y_{ijt-1}^{lw} - \rho_8 y_{ijt-1}^{hw}) + \rho_{10} \bar{c}_{jt} y_{ijt-1} + \bar{c}_{jt}(\rho_{10} y_{ijt-1}^{lc} - \rho_{11} y_{ijt-1}^{hc}) + e_{ijt} \tag{B-7}$$

The terms $\rho_8 \bar{w}_{jt} y_{ijt-1}$ and $\rho_{10} \bar{c}_{jt} y_{ijt-1}$ can, of course, be pulled out and used directly in the estimation since they are observables. Thus, the residual error-term is:

$$\epsilon_{ijt}' = \bar{w}_{jt}(\rho_9 y_{ijt-1}^{lw} - \rho_8 y_{ijt-1}^{hw}) + \bar{c}_{jt}(\rho_{10} y_{ijt-1}^{lc} - \rho_{11} y_{ijt-1}^{hc}) + e_{ijt} \tag{B-8}$$

Second, some terms can again be decomposed and written as functions of past state-level adoptions, aggregate wealth and cultural capitals, and their interactions. For example, $y_{ijt-1}^{lw}, y_{ijt-1}^{hw}$ can be written as:

$$y_{ijt-1}^{lw} = \mathcal{F}_{ht}(y_{ijt-1},...y_{ijt-p-1}, y_{it-2}, x_{it-1}, \bar{w}_{jt-1}, \bar{c}_{jt-1}, z_i, \gamma_{ij}) \tag{B-9}$$
$$y_{ijt-1}^{hw} = \mathcal{F}_{lt}(y_{ijt-1},...y_{ijt-p-1}, y_{it-2}, x_{it-1}, \bar{w}_{jt-1}, \bar{c}_{jt-1}, z_i, \gamma_{ij}) \tag{B-10}$$

Thus, much of the remaining variation in $y_{ijt-1}^{lw}, y_{ijt-1}^{hw}$ is captured through these lag variables and name-state fixed effects. Third, since many of the instruments in the estimation are for the first-differenced equation, the error-terms used in estimation are $\epsilon_{ijt}' - \epsilon_{ijt-1}'$. It is well-known that

first-differencing significantly assuages aggregation issues in models like this by differencing out much of the variation in the error terms, making the first-differenced error terms to be independent of endogenous explanatory variables. Please see Stoker (1993) for details.

Nevertheless, some remnant variation may still remain significant. If so, it will lead to serial correlation in estimated errors (through correlations in adoptions among high and low types across consecutive years). The main advantage of our estimator is that it allows us to test this empirically.

After estimating the model, and obtaining the parameters and error terms, we test for serial correlation in error terms using the Arellano-Bond (2) test. If the test rejects the hypothesis of no serial correlation, then it implies that the presence of within-state effects has invalidated our aggregated social effects. That is, if we find that $E(\widetilde{\epsilon}_{jit} \cdot \widetilde{\epsilon}_{ijt-1}) \neq 0$, where:

$$
\begin{aligned}
E(\widetilde{\epsilon}_{jit} \cdot \widetilde{\epsilon}_{ijt-1}) &= E[\bar{w}_{jt}(\rho_9 y^{lw}_{ijt-1} - \rho_8 y^{hw}_{ijt-1}) + \bar{c}_{jt}(\rho_{10} y^{lc}_{ijt-1} - \rho_{11} y^{hc}_{ijt-1}) + e_{ijt}) \\
&\quad \cdot (\bar{w}_{jt-1}(\rho_9 y^{lw}_{ijt-1} - \rho_8 y^{hw}_{ijt-1}) + \bar{c}_{jt-1}(\rho_{10} y^{lc}_{ijt-1} - \rho_{11} y^{hc}_{ijt-1}) + e_{ijt-1})] \quad \text{(B-11)}
\end{aligned}
$$

then the estimates from the aggregated model are inconsistent. If instead $E(\widetilde{\epsilon}_{jit} \cdot \widetilde{\epsilon}_{ijt-1}) = 0$, then the estimates of state-level social effects are consistent even if we do not have information on within state or more local neighborhood-level effect.

Thus, the presence of local/within-state social effects does not invalidate aggregate-level social effects if the Arellano-Bond (2) test is satisfied.

B.3 Popularity Cycles of the Top Three Female and Male Names from 2000 and 2009
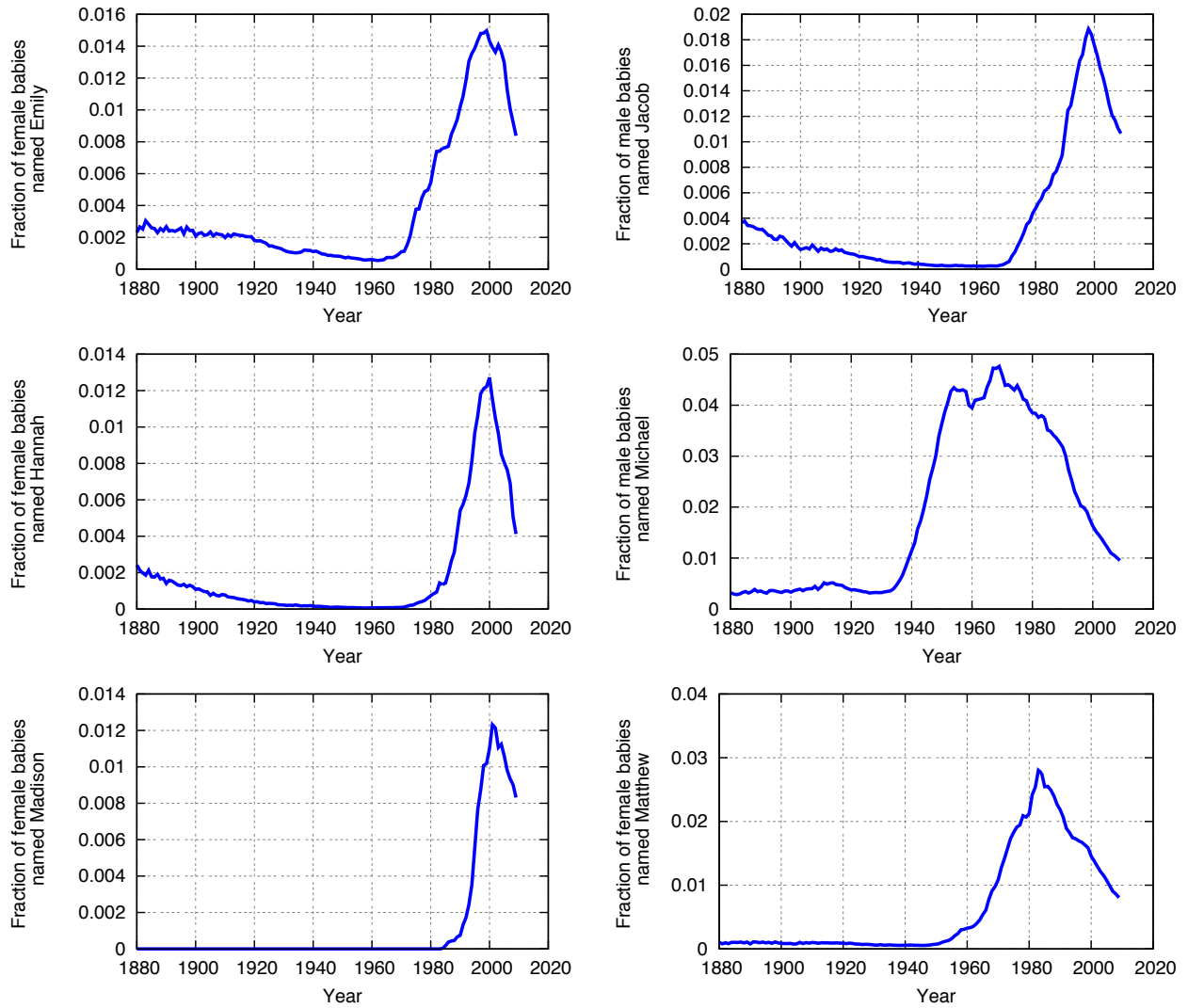

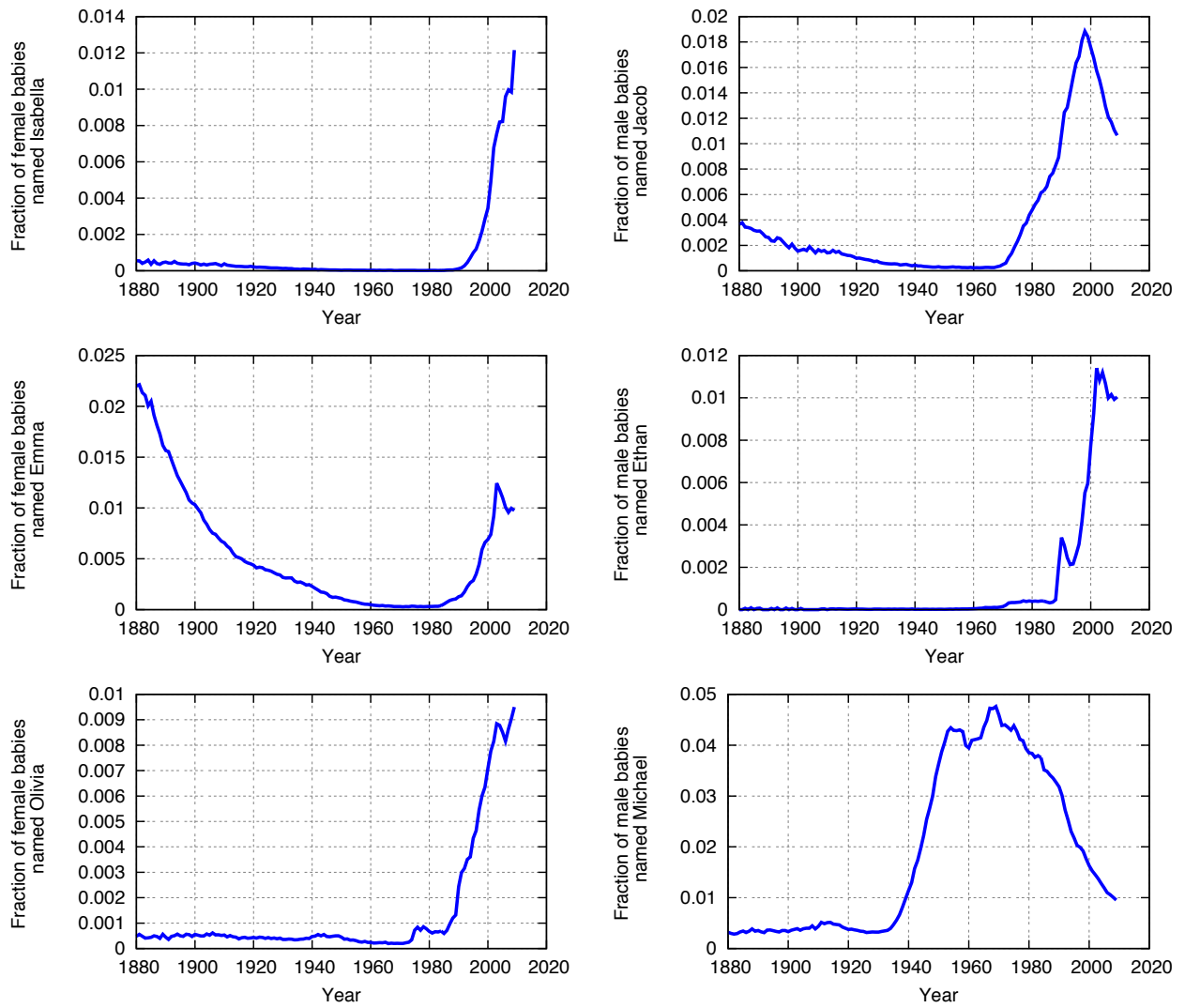Figure 1: Popularity Curves of the Top Three Female and Male Baby Names in 2000.

Figure 2: Popularity Curves of the Top Three Female and Male Baby Names in 2000.