First Names and Longevity

Laura Pinzur

Department of Economics

Pomona College

Gary Smith

Department of Economics

Pomona College

Correspondence:

Gary Smith

Fletcher Jones Professor

Pomona College

425 N. College Way

Claremont, CA 91711

phone: 909-607-3135; fax: 909-6231-8576

e-mail: gsmith@pomona.edu

First Names and Longevity

Summary.—A statistical relationship has been reported between people's initials and their life expectancy. Several studies have also reported that people with uncommon first names are perceived to be less intelligent, attractive, and likable than are people with more popular names. This leads to the possibility that such social stigmatization may affect the life expectancy of people with unpopular names. The California Department of Health Services mortality data base for the years 1960 through 2004 for 6.7 million white, nonHispanic decedents was used to compare the average age at death for decedents with the most popular and least popular first names. These data do not show a relationship between the popularity of a first name and life expectancy.

First Names and Longevity

Walton (1937) argued that a person's first name "may be a determining factor in his development of personality, acquisition of friends, and in all probability, in his success or failure in life." (p. 396) Several subsequent studies have examined the relationship between first names and how people are perceived by themselves and by others (Horne, 1986; Joubert 1993).

McDavid and Harari (1966) found a correlation between elementary school children's peer-group popularity and the desirability rating of their first names. They also found that uncommon names received low ratings and concluded that children with rare names may tend to be viewed negatively. In a later study, Harari and McDavid (1973) argued that rare names are generally considered to be less socially attractive and to invoke negative stereotypes. In a controlled experiment, they found that, for both boys and girls, teachers gave higher grades to the same essays if the purported author's first name was common than if it was rare.

Garwood (1976) found a positive relationship between school achievement and the desirability of names as judged by teachers, and Erwin and Calev (1984) found that school teachers gave lower grades to essays purportedly written by children with names judged to be unattractive by a survey of undergraduate students. Savage and Wells (1948) found that among 3,320 Harvard undergraduates, those with uncommon names were more likely to expelled for poor academic performance and more likely to be classified as "psychopathic personality" or psychoneurosis" by the Hygiene Department. Joubert (1983) found that

among 1,390 college undergraduates, those with uncommon names were less likely to graduate with honors.

Erwin (1993) examined the link between names and the perception of physical attractiveness by others. For this study, 36 male and 36 female undergraduate students rated the physical attractiveness of 3 male and 3 female photographs with randomly assigned first names that had previously been rated for attractiveness by 10 male and 10 female subjects. Photographs with attractive names were ranked significantly more physically attractive.

Busse and Seraydarian (1978) found a link between the frequency of first names and their desirability. Over 2,000 elementary and high school aged students rated 179 boy names and 246 girl names. There was a significant correlation between the name desirability rating and the frequency with which the name occurred in the sample for both boy names ($r = 0.67$, $p < 0.001$) and for girl names ($r = 0.62$, $p < 0.001$).

Levine and Willis (1994) compared a list of 260 children's names gathered from 100 participants to a list of 1,000 names chosen at random from the city directory. Any name not appearing on the directory list was classified as unusual. Twenty usual names were chosen at random from those names remaining. Two hundred judges were then asked to rate each name on a 5-point scale for success, morality, health, warmth, cheerfulness, and sex stereotype. The common names were given more desirable ratings in all six of these categories. Similarly, Karlin & Bell (1995) found that undergraduates were twice as likely to associate positive traits with common names than with uncommon names.

First-name effects have been found to be stronger among males than among females (Ellis & Beechley, 1954; Willis, Willis, & Gier, 1982; Anderson and Schmitt, 1990). Some

attribute this difference to the fact that girls are more likely to be given uncommon names (Anderson, 1985), so that uncommon names stand out more for boys than for girls.

Zweigenhaft (1977) presents evidence indicating that an uncommon name may not be a disadvantage. Similarly, Skinner (1984) looked at the grades of 597 students in an introductory psychology course and found that, while students with unique names did not do quite as well as other students, the observed differences did not have p-values less than 0.05 because of the small sample size. The general applicability of his results is further constrained, as he notes, by the fact that the sample is relatively homogeneous and academically successful.

The above research suggests that a name may be related to an individual's self-esteem and perception by others; it may also be related to one's health. Several studies have explored the relationship between youth self-esteem and adult well-being. Ray (2004) summarizes several studies indicating that mortality is affected by a person's social support system and that individuals with larger social networks have lower blood pressure compared to individuals with smaller social networks. Health can evidently be affected by how others perceive and relate to a person.

Longitudinal studies suggest that these effects may persist. Trzesniewsky et al. (2006) measured self-esteem in 978 children at ages 11, 13, and 15. When their subjects were 26 years old, they collected data on the subjects' well-being. They found that adolescents with low-self esteem at ages 11, 13, and 15 tended to have more mental health problems, more physical health problems, more criminal convictions, and diminished economic prospects at

age 26. If names are associated with low self-esteem, names may also be associated with conditions that can affect life expectancy.

McGee and Williams (2000) examined the correlation between self-esteem and health-compromising behavior in 959 adolescents. The participants completed forms at ages 9, 11, and 13 assessing their academic self-esteem and overall self-esteem. At age 15, they were assessed for a number of health-compromising behaviors. Controlling for family background, lower self-esteem was related to eating disorders, early sexual activity, and suicidal thoughts.

Antonucci, Peggs, and Marquez (1989) also explored the relationship between self-esteem and physical health. Sixty-eight medical patients at a university-affiliated family practice center completed a questionnaire that measured self-esteem. Those patients with higher self-esteem had a better health status than those who reported lower levels of self-esteem. Since physical health is a predictor of life expectancy, we might expect that those with higher self-esteem also live longer.

One important measure of the relationship between first names and well-being is life expectancy: whether people with unpopular names have different mortality rates, on average, than do people with popular names. Christenfeld, Phillips, and Glynn (1999) reported that life expectancy was related to a person's 3-letter initials; specifically, that people with positive initials (such as ACE or WIN) live much longer than do people with negative initials (such as PIG or DIE). They report that in comparison to their controls, males with positive initials lived 4.5 more years, males with negative initials lived 2.8 fewer years, females with positive initials lived 3.4 more years, and that there was no difference for females with

negative initials. Abel and Kruger (2007) similarly reported that major league baseball players with positive initials lived 13 years longer than did persons with negative initials.

If initials, which are not used in everyday discourse, are related to longevity, it is possible that names, which are spoken and heard innumerable times in a person's lifetime, are related to life expectancy. We test this relationship by using California mortality data for the years 1960 through 2004 to compare the average age at death for white, nonHispanic males and females with the most popular and least popular names in each birth year.

Sample

The California Department of Health Services (1960-2004) maintains a mortality data base back to 1960 that identifies each decedent's name, gender, date of birth, date of death, and race or ethnicity. (They also have a mortality data base for 1905 to 1959 that identifies name, gender, date of birth, and date of death—but not race or ethnicity. In practice, these early years have few usable data because the recorded date of birth is usually "unknown.")

Because mortality varies by race and gender, we follow common practice (Christenfeld, *et al.*, 1999) by looking at white, nonHispanic decedents for 1960-2004, and separate the decedents by gender. Our primary analysis is of decedents who were at least 50 years old at the time of death, since we are interested in the long-term effects of one's name on mortality, not deaths that occur shortly after birth, during childhood, or during young adulthood (which would also involve the confounding effects of military service). For comparison, we also report results for all decedents, and for decedents who were at least 10 years old and at least 21 years old at the time of death.

We excluded decedents with no recorded first name or with only a single-letter first name (presumably the first initial of their name). There were a total of 3,447,476 male decedents with 55,927 distinct names and 3,209,199 female decedents with 68,354 distinct names.

Methods

Retrospective studies have many possible pitfalls. Here, one serious problem is that the popularity of names changes over time. Suppose, for example, that mortality rates are constant and do not depend on a person's name, but that there are more people named Hunter and Madison in later birth cohorts than in earlier cohorts. If so, people named Hunter and Madison, whether dead or alive, will tend to be younger than average. Suppose, using extreme assumptions to make the point, that the names Hunter and Madison have only been used in the past 20 years. All decedents named Hunter or Madison would consequently have died before the age of 20.

This problem can be circumvented by grouping decedents by birth year. Morrison and Smith (2005) show that grouping decedents by birth year provides a valid test of the null hypothesis that mortality rates are the same for different groups of decedents. Specifically, if two groups have the same mortality rates, then the expected value of the average age at death (AAD) over any horizon will be the same for both groups. For example, if we look at two groups with the same mortality rates who were born in 1900, the expected value of the AAD increases as we expand the horizon from, say, 25 to 50 to 100 years; however, the expected value of the AAD is the same for both groups whether we look at horizons of 25, 50, or 100 years.

Earlier studies have characterized names either by their frequency or by their desirability rating based on peer surveys. These are not independent characterizations since, as noted above, Busse and Seraydarian (1978) and Levine and Willis (1994) both found that frequencies are related to desirability ratings. Because we cannot conduct peer surveys of generations now deceased, we will characterize names by frequency.

In each birth year, we identified the ten most popular names and all of the unique names of decedents born in that year; i.e., we considered a name to be unpopular in that birth year if there was only one decedent with that name. Joubert (1983) and Anderson (1985) similarly define uncommon names as those that appear only once in a sample. For example, for the 1900 birth year, the ten most popular male names are shown in Table 1—a total of 17,608 decedents. For that same birth year, there were 3,124 male decedents with unique names. (The large number of people with unique names—the popularity of unpopular names—reminds us of Yogi's Berra's remark that, "Nobody goes there anymore; it's too crowded.)

For each birth year t, the average age at death was calculated for the decedents with the 10 most popular names (AAD_t^+) and for the decedents with unpopular names (AAD_t^-). If there were at least 10 decedents with unpopular names (in practice, there were at least 43 unpopular decedents in every birth year from 1861 through 1984), then the paired difference in the average age at death was calculated for that birth year

$$D_t = AAD_t^+ - AAD_t^-$$

For example, for male decedents born in 1900, the average age at death was 76.39 years for those with popular names and 76.80 years for those with unpopular names, a difference of -0.41 years.

The null hypothesis is that mortality rates are not related to names, so that the expected value of each paired difference is zero: $E[D_t] = 0$. Two-sided p values were calculated using a matched-pair t test.

Results

Over the entire period, there were a total of 1,049,737 males with popular names, 181,099 males with unpopular names, 649,536 females with popular names, and 199,915 females with unpopular names. Tables 1 and 2 show the ten most popular male and female names in our data base for the 1900 birth year, 1950 birth year, and all birth years. Some names were perennially popular: John and Mary were top-10 names in every single year. However, there have also been some changes in the top 10 over time. Table 3 shows several unusual names, including Nevermore, Strange, and WoollooMoolo.

Figures 1 and 2 show the male and female AADs for all birth years, which are over 100 for people born in 1860 (since the mortality data begin in 1960) and decline as the birth year approaches 1954, the last death year in our data with 50-year-old decedents.

The ideal data set would have complete mortality data for everyone with popular and unpopular names in every birth cohort. Thus, our data set is less than ideal. There are some birth cohorts with reasonably complete mortality data. However, there are also some early birth years (like 1860) for which many people died too young to be included in our study, and some later birth years (like 1950) for which many people died (or will die) too old to be included. Thus some birth years have decedents of all ages, while other birth years have mostly older decedents and other birth years have mostly younger decedents. Nonetheless, as explained earlier, our statistical test is valid with a well-defined p value.

Figures 1 and 2 show that, for every birth year, there is very little difference in the AAD for those with popular and unpopular names. This is true for birth years with decedents of all ages, for birth years with mostly older decedents, and for birth years with mostly younger decedents. The difference in AAD is less than one year in every birth year and is less than six months in 83 percent of the birth years.

Table 4 shows that the mean and median differences in the AADs over the entire period are slightly negative for both males and females (those with unpopular names lived slightly longer, on average, than did those with popular names). To test the robustness of our results, we also looked at all decedents, regardless of their age at time of death, and at decedents who were at least 10 years old and at least 21 years old at the time of death. The largest observed difference in the mean or median AAD is -0.26 years ($t = -9.59$, $p < 0.01$). This case (males who were at least 50 years old when they died) is the only instance of a p value less than 0.05 and, here, those with unpopular names lived slightly longer, on average, than those with popular names. The one statistically significant result may be due to the number of tests conducted.

Decedents Grouped by Death Year

We also grouped decedents by death year in order to confirm the point made by Morrison and Smith (2005) that the grouping of decedents by death year can be misleading if there are changes over time in the frequency of the characteristics we are interested in—here, names. For example, Table 2 suggests that the popularity of the name Mary has declined over time. (Of all decedents born in 1900, 4.2 percent were named Mary; of all decedents born in 1950,

2.8 percent were named Mary.) Morrison and Smith show that this sort of decline in later birth cohorts tends to increase the AAD of decedents grouped by death year.

In practice, a diverse mixture of names have been more or less popular over time and moved in and out of the top 10. It is consequently unclear how, on balance, a grouping by death year will affect the difference between the AADs of popular and unpopular names. All we know is that there can be effects either way.

Figures 3 and 4 show the AADs for male and female decedents with popular and unpopular names, grouped by year of death. These do, in fact, turn out to be quite misleading. A grouping by year of death suggests that a person's name can have a large effect on one's life expectancy—an effect that disappears when the decedents are grouped by year of birth.

Interestingly, the bias is one way for males and the other way for females. When grouped by year of death, male decedents with popular names generally had a lower AAD, while the opposite was true for female decedents. The respective mean and median differences in the AAD of those with popular and unpopular names are -0.92 and -1.36 years for males and 1.26 and 1.12 years for females.

Discussion

It has been reported that initials are related to life expectancy (Christenfeld, Phillips, and Glynn, 1999; Abel and Kruger, 2007). It has also been reported (for example, McDavid and Harari, 1966; Harari and McDavid, 1973; Erwin, 1993; Levine and Willis, 1994) that people with unpopular first names are perceived by themselves and by others as inferior to people with popular names—perceptions that might affect one's health. Nonetheless, we cannot infer causality from statistical correlations between first names and longevity because the parents'

choice of an unusual name may well reflect their personalities and other confounding factors that are related to life expectancy.

The most natural test of whether there is a statistical relationship between first names and longevity is to group the decedents by birth year and to compare the AAD of those with popular and unpopular names. If two groups have the same mortality rates, the expected value of the AAD over any horizon will be identical for the two groups. California mortality data for nonhispanic whites for the years 1960-2004 do not show a statistical relationship between longevity and name popularity. There is no substantial or statistically persuasive evidence that people with unpopular names die younger than do people with popular names. If there are any effects, they are probably small and likely related to ethnic, religious, cultural, and socioeconomic factors that are correlated to name choices.

References

Abel, E. L., & Kruger, M. L. (2007) Symbolic significance of initials on longevity. *Perceptual and Motor Skills*, 104, 179-182.

Anderson, T. (1985) Unique and common first names of males and females, *Psychological Reports*, 57, 204-206.

Anderson, T., and Schmitt, P. R. (1990) Unique first names in male and female psychiatric inpatients, *Journal of Social Psychology*, 130, 835-837.

Antonucci, T., Peggs, J., & Marquez, J. (1989) The relationship between self-esteem and physical health in a family practice population. *Family Practice Research Journal*, 9, 65-72.

Busse, T., & Seraydarian, L. (1978) Frequency and desirability of first names. *The Journal of Social Psychology*, 104, 143-144.

California Department of Health Services. (1960-2004) Death statistical master files. Sacramento, CA: Author.

Christenfeld, N., Phillips, D., & Glynn, L. (1999) What's in a name: mortality and the power of symbols. *Journal of Psychosomatic Research*, 47, 241-254.

Ellis, A., & Beechley, R. M. (1954) Emotional disturbance in children with peculiar names, *Journal of Genetic Psychology*, 3, 337-339.

Erwin, P. (1993) First names and perceptions of physical attractiveness. *The Journal of Psychology*, 127, 625-631.

Erwin, P., & Calev, A. (1984) The influence of Christian name stereotypes on the marking of children's essays. *British Journal of Educational Psychology*, 54, 223-227.

Garwood, S. (1976) First-name stereotypes as a factor in self-concept and school achievement. *Journal of Educational Psychology*, 68, 482-487.

Harari, H., & McDavid, J. (1973) Name stereotypes and teachers' expectations. *Journal of Educational Psychology*, 65, 222-225.

Horne, M. D. (1986) Potential significance or first names: a review, *Psychological Reports*, 59-839-845.

Joubert, C. E. (1983) Unusual names and academic achievement. *Psychological Reports*, 53, 266.

Joubert, C. E. (1993) Personal names as a psychological variable. *Psychological Reports*, 73, 1123-1145.

Karlin, N. J., & Bell, P. A. (1995) Assessing commonality and favorability of first names, *Psychological Reports*, 77, 97-98.

Levine, M., & Willis, F. (1994) Public reactions to unusual names. *The Journal of Social Psychology*, 134, 561-568.

McDavid, J., & Harari, H. (1966) Stereotyping of names and popularity in grade-school children. *Child Development*, 37, 453-459.

McGee, R., & Williams, S. (2000) Does low self-esteem predict health compromising behaviours among adolescents? *Journal of Adolescence*, 23, 569-582.

Morrison, S., & Smith, G. (2005) Monogrammic determinism? *Psychosomatic Medicine*, 67, 820-824.

Ray, O. (2004) How the mind hurts and heals the body. *American Psychologist*, 59, 29-40.

Savage B.M., and Wells, F. L. (1948) A note of singularity in given names, *Journal of Social Psychology*, 27, 271-272.

Trzesniewsky, K., Donnellan, M., Moffitt, T., Robins, R., Poulton, R., & Caspi, A. (2006) Low self-esteem during adolescence predicts poor health, criminal behavior, and limited economic prospects during adulthood. *Developmental Psychology*, 42, 381-390.

Walton, W. E. (1937) The affective value of first names, *Journal of Applied Psychology*, 21, 396-409

Willis, F. N.,Willis, L. A., and Gier, J. A. (1982) Given names, social class, and professional achievement, *Psychological Reports*, 51, 543-549.

Zweigenhaft, R. L. (1977) The other side of unusual first names, *The Journal of Social Psychology*, 103, 291-302.

Table 1

The Most Popular Male Names Among 1960-2004 Decedents

1900 Birth Year		1950 Birth Year		All Birth Years	
Name	Number	Name	Number	Name	Number
John	3,200	Robert	787	John	178,539
William	2,901	Michael	765	William	156,615
George	1,910	John	729	Robert	125,144
James	1,747	James	656	James	116,757
Charles	1,711	David	574	Charles	93,286
Joseph	1,665	William	531	George	90,023
Frank	1,291	Richard	510	Joseph	76,381
Robert	1,179	Thomas	367	Frank	59,405
Harry	1,043	Steven	331	Richard	54,506
Edward	965	Gary	322	Edward	51,897

Table 2

The Most Popular Female Names Among 1960-2004 Decedents

1900 Birth Year		1950 Birth Year		All Birth Years	
Name	Number	Name	Number	Name	Number
Mary	3,141	Linda	360	Mary	145,579
Helen	1,810	Mary	199	Helen	67,331
Ruth	1,502	Patricia	169	Margaret	64,073
Margaret	1,411	Susan	168	Dorothy	56,319
Anna	1,247	Deborah	136	Ruth	50,881
Dorothy	1,039	Nancy	128	Anna	46,248
Florence	1,020	Karen	122	Alice	35,264
Marie	968	Kathleen	120	Florence	34,135
Gladys	895	Sandra	119	Frances	33,712
Rose	882	Barbara	112	Marie	33,093

Table 3

Some Uncommon Names

Males	Females
Anaus	Annus
Coma	Bland
Dump	Dorka
Elf	Earless
Germe	Goon
Gross	Junko
Gurperminder	Naste
Queenie	Nevermore
Sad	Woolloomoolo
Strange	Yicke

Table 4

Paired Difference in Average Age at Death (AAD) for Decedents with Popular Names Minus AAD for Decedents with Unpopular Names

	Males				Females			
Minimum Age	Mean	Median	t value	p value	Mean	Median	t value	p value
0	0.08	-0.03	1.32	0.18	0.06	0.05	1.29	0.20
10	0.01	-0.02	0.31	0.76	0.04	0.08	0.95	0.34
21	-0.03	-0.09	-0.92	0.36	0.02	0.04	0.68	0.50
50	-0.25	-0.26	-9.59	< 0.01	-0.06	-0.08	-1.65	0.10

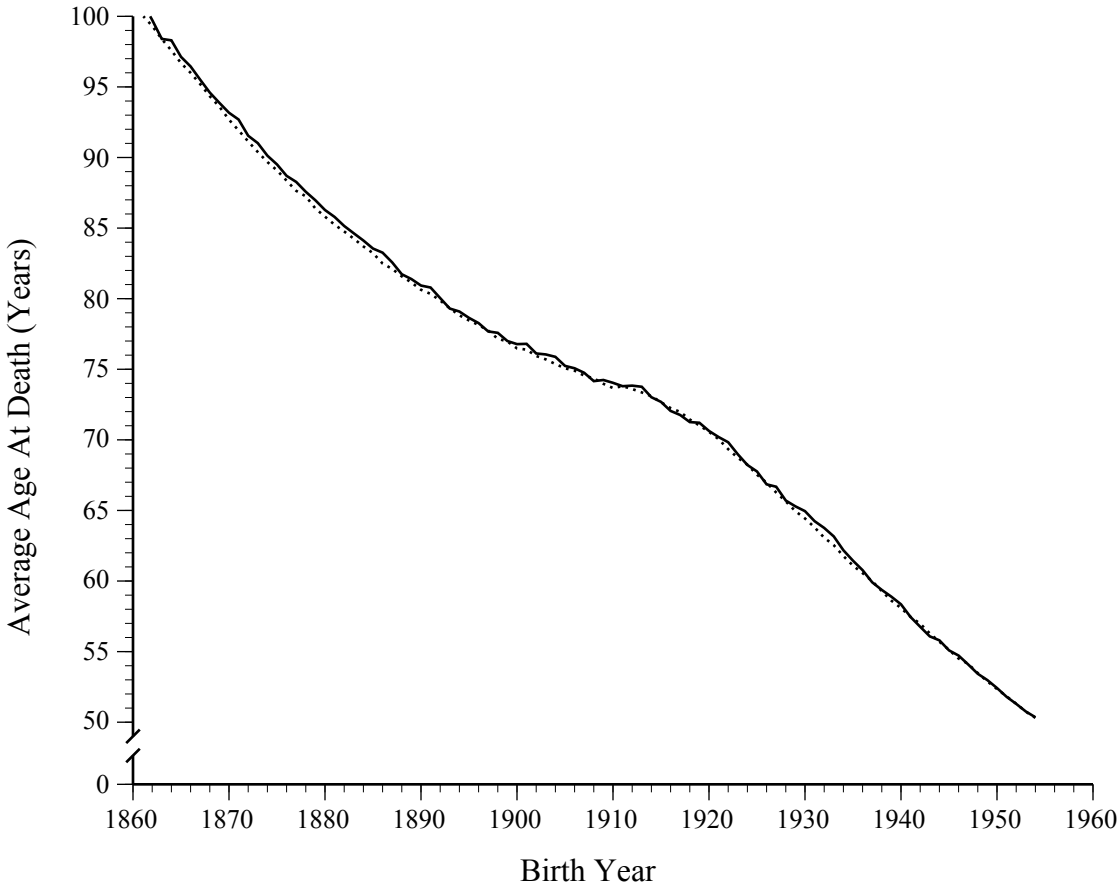

Figure 1 Average Age at Deathof Male Decedents with

Popular (dashed) and Unpopular (solid) Names, Grouped by Birth Year

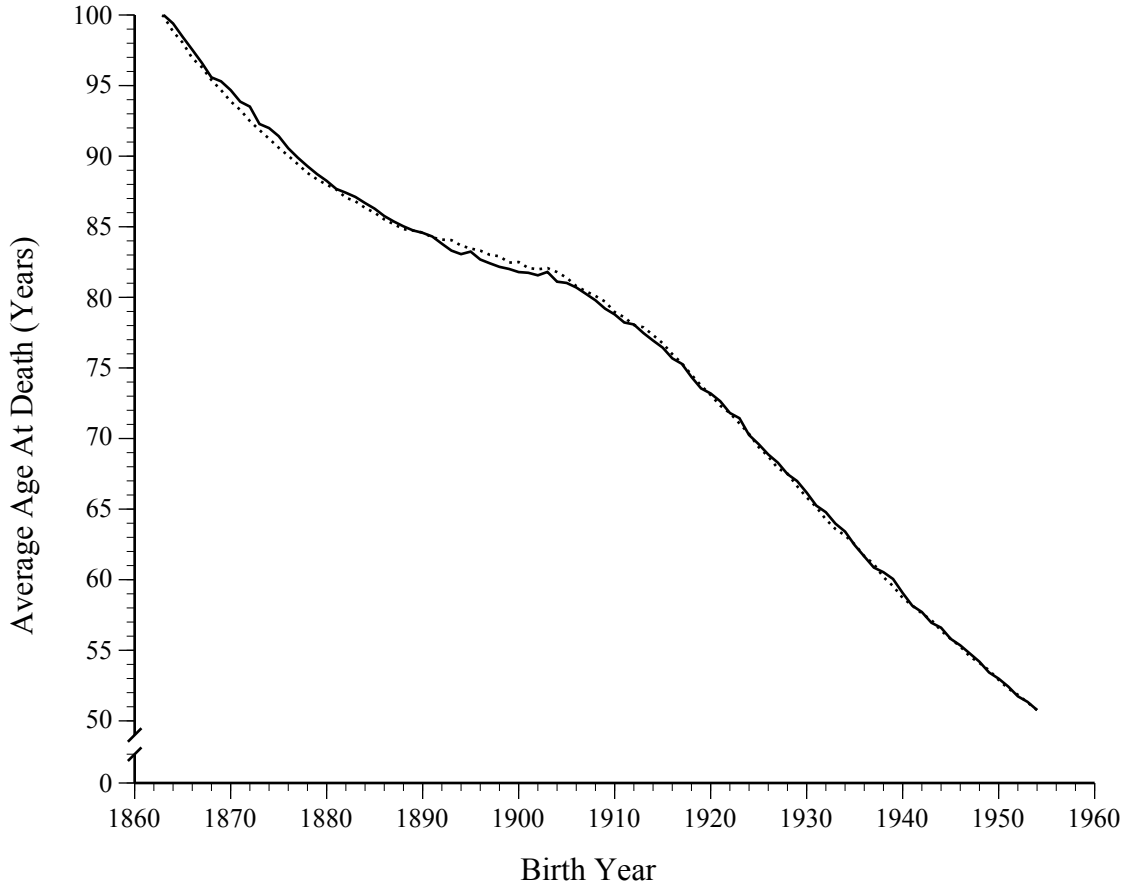

Figure 2 Average Age at Death of Female Decedents with

Popular (dashed) and Unpopular (solid) Names, Grouped by Birth Year

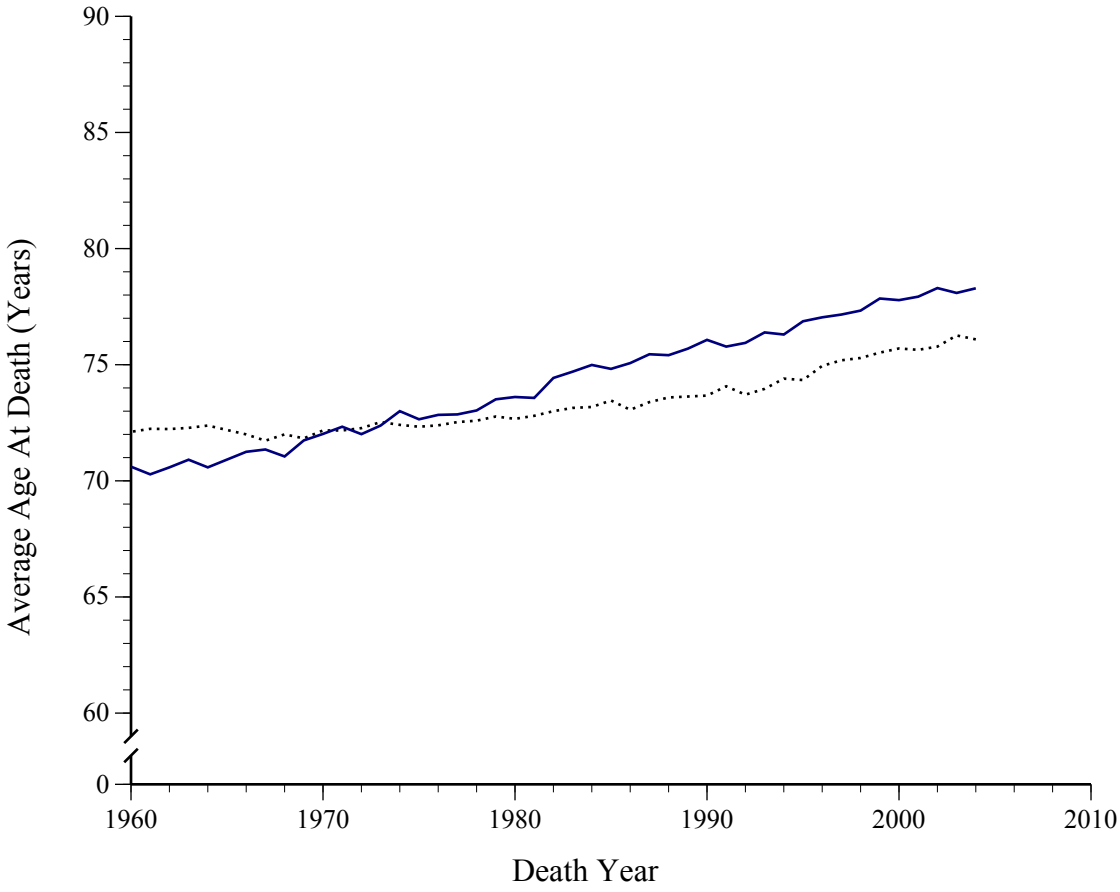

Figure 3 Male Decedents with Popular (dashed) and Unpopular (solid) Names;

Grouping by Death Year Can be Misleading

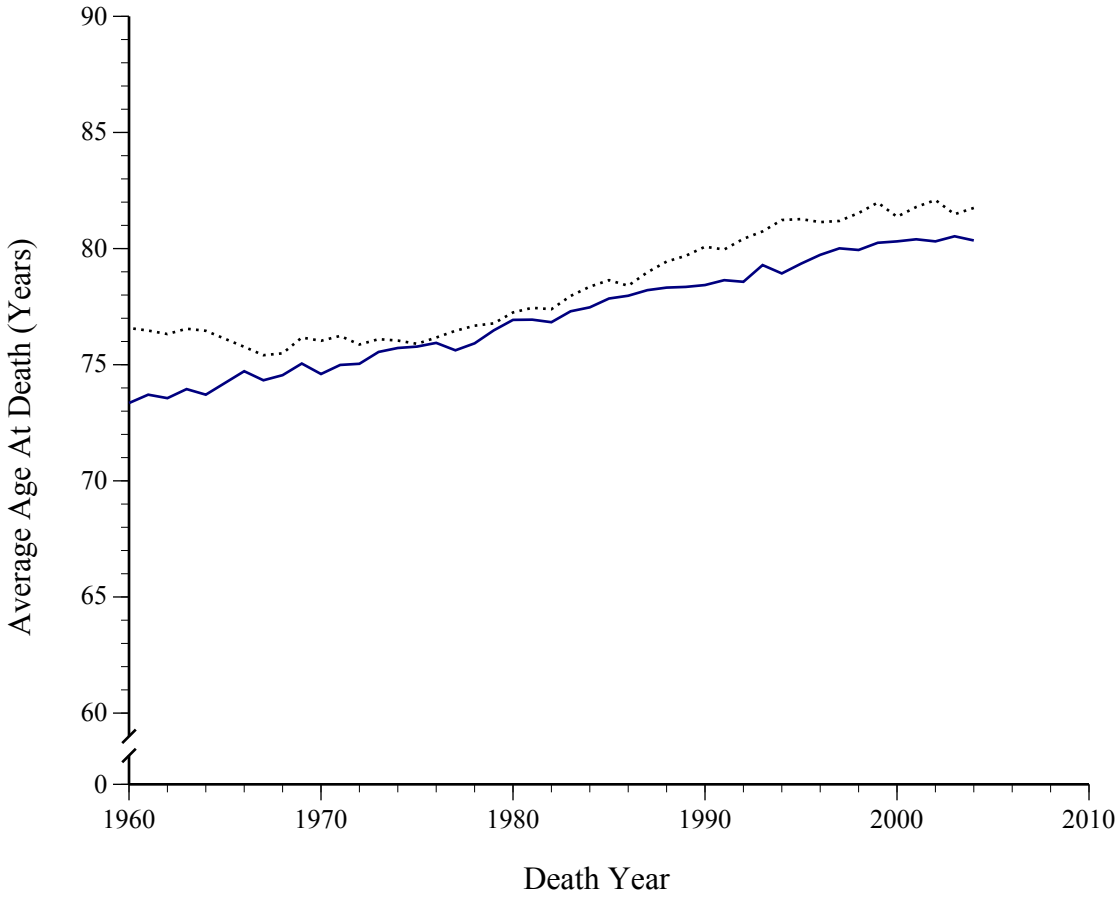

Figure 4 Female Decedents with Popular (dashed) and Unpopular (solid) Names;

Grouping by Death Year Can be Misleading