Alphabetism: The effects of surname initial and the risk of being otherwise undistinguished

Alexander S. Cauley
Jeffrey S. Zax

University of Colorado Boulder
Department of Economics
December 31, 2016

Participants in the Applied Micro Workshop at the University of Colorado Boulder have been generous with comments. This draft is preliminary and unworthy of citation. Please refrain. The authorial order represents relative contributions. Any resemblance to alphabetical ordering is purely coincidental.

1. Introduction

Individual experiences of life, and of economic success in particular, depend on many individual characteristics. Human capital is perhaps the most prominent among them, and its role has been the subject of extensive study. The role of identity, defined by ethnic or racial affiliation, national identity or gender has also received substantial attention. Other salient characteristics include non-cognitive skills, personality traits, height and appearance.

At the same time, names are central to identity. A small literature has exploited names to proxy for unobserved socio-demographic characteristics. A slightly larger literature has examined the role of alphabetization in academic publishing, and in a small number of other domains. However, the literature does not contain any comprehensive attempt to assess the more general effects, if any, of alphabetization.

This paper analyzes alphabetism, the relationship between the alphabetic rank of surname initial on experience in high school, investment in further accumulation of human capital through tertiary education, and labor market success in early and mid-adulthood. It demonstrates that those with surname initials ranked further from the beginning of the alphabet experience significantly, and in many cases substantively, worse outcomes through early labor market experiences. These effects disappear by mid-adulthood, presumably because they are superseded by observable characteristics that are more directly expressive of ability.

Section 2 reviews the literature describing the information content and economic effects associated with names and surname initials. Section 3 describes the econometric models and data employed here, including the twelve dependent variables that describe high school experiences, tertiary education accomplishment and labor market outcomes through mid-adulthood. Section 4 estimates general effects of surname initials on these outcomes. Section 5 distinguishes between

these effects on individuals who are and are not distinctive in terms of cognitive ability and physical attractiveness. Section 6 concludes.

2. The economic role of names

Names have important roles in economics. Their content can convey a substantial amount of economic information. Given names, surnames and their combination can identify ancestry and ethnicity. They can also identify socioeconomic status, and affect that status if changed.

Apart from name content, orderings by name can affect search outcomes. In turn, these can affect the distribution of resources and opportunities. Moreover, individuals whose rank in these orderings is disadvantaged may respond strategically by altering their position and participation in search-based activities. However, distinction in these activities can transcend the effects of alphabetic rank.

a. Name content and economic outcomes

Given names alone are informative. They are indicative of parental economic status (Aura and Hess, 2010; Olivetti and Paserman, 2015). Teachers may have lower expectations for students with given names that are associated with low socio-economic status (Figlio, 2005).

Either because given names are correlated with resources and attitudes or because they affect opportunities, they can influence economic outcomes. Given names that are distinctively African-American are associated with residence in poorer neighborhoods (Fryer and Levitt, 2004) and an array of poorer life outcomes (Aura and Hess, 2010). Immigrants to the United States enjoyed greater occupational success if they replaced their original given names with typical American alternatives (Biavaschi, et al., forthcoming). Men with first names associated with Croatian

nationalism were more likely to enlist in the Croatian Army and to be killed in the Croatian War of Independence (Jurajda and Kova, 2016).

Experiments suggest that combinations of given and surnames are also economically potent. Choices in common two-person experimental games depend on the full name of the partner (Fershtman and Gneezy, 2001). Its influence appears to be associated with behavioral expectations associated with the ethnicity indicated by that name.

Callback rates for synthetic on-line job applications are lower for those with complete names that suggest African-American origin in the United States (Bertrand and Mullainathan, 2004), Arabic origin in the United States (Widner and Chicoine, 2011), middle eastern origin in Sweden (Carlsson and Rooth, 2007) and Turkish origin in Belgium (Baert, et al., 2015). Callback rates for synthetic on-line apartment rental applications in the United States are lower for those with names that suggest African-American or Arabic identity (Carpusor and Loges, 2006).

Non-experimental evidence is partially supportive. Mutual funds in the United States experience significantly lower fund inflows if the fund manager has a complete name that is widely perceived as "foreign" (Kumar, et al., 2015). However, complete names have no important effects on academic outcomes in college, holding constant race (Foster, 2008).

Surnames alone convey important historical and demographic information. Clark (2014) and Clark and Cummins (2015) examine inter-generational mobility by characterizing the economic status of individuals from different generations bearing the same surname. Angelucci, et al. (2010) employ surnames to identify extended families. Arai and Skogman Thoursie (2009) demonstrate that, when foreign-born residents of Sweden change their surnames to names that are Swedish or ethnically neutral, their incomes rise significantly. Surname frequency is a source of inferences

regarding intergenerational mobility (Güell, et al., 2015), socieoeconomic status (Collado, et al., 2008) and longevity (Pena, 2013).[1]

b. Ordering and primacy through surname initial

In environments with multiple options that must be evaluated sequentially and whose characteristics are ex ante uncertain, search theory concludes that continued search must balance the cost of delaying choice against the potential benefit of identifying a superior option. The optimal stopping rule consists of identifying a minimum acceptable, or "reservation" quality for the choice, and concluding the search with the first option that meets or exceeds that quality (Kohn and Shavell, 1974; Albrecht, 2011). In simple contexts, the probability that an acceptable choice has been identified increases as search progresses. Consequently, the probability of being sampled declines with rank.[2]

This prediction is consistent with the results of several studies of academic publishing. Feenberg, et al. (forthcoming) demonstrate that papers that appear first in an on-line listing are significantly more likely to be viewed, downloaded and cited, even though the ordering is random. In a different on-line list with random order, Haque and Ginsparg (2009) again demonstrate that first-listed papers are downloaded and cited more often. In yet another random on-line list, Dietrich

[1] Rubinstein and Brenner (2014) invoke unobserved surnames and their ethnic identifications as the mechanism by which parental ethnicity affects economic outcomes.

[2] Carney and Banaji (2012) demonstrate that immediate responses to multiple, objectively equivalent, options are characterized by a significant preference for the first. This effect is not present when the same options are compared under circumstances that encourage reflection. The "primacy" effect that they identify is probably different from that associated with deliberative search. Instead, they speculate that it has origins in evolutionary success.

(2008) finds that papers listed first receive more citations.[3] Berger (2016) reports that articles listed first in randomly-ordered printed tables of contents receive more citations.

Ordering effects also appear in contexts other than academic publishing. Arbatskaya (2007) presents a model in which ordered, costly consumer search implies that, in equilibrium, prices for a homogeneous good vary systematically with search order. Ho and Imai (2008) demonstrate that first ballot position conveys an electoral advantage. Ginsburgh and van Ours (2003) demonstrate that outcomes in a performance competition depend on the randomly-assigned order of the competitors.[4]

Economic outcomes may depend on name initial as well as name content. Groups are often ordered and searched by alphabetic rank of initial. In these contexts, the probability that an acceptable choice has been identified increases as the search proceeds to letters with higher "rank". Consequently, economic opportunities diminish and, therefore, outcomes suffer with the rank of surname initial.

Multiple examples arise in business contexts. Jacobs and Hillert (2016) demonstrate that stocks with names that are ranked earlier in the alphabet experience higher trading volumes and lower liquidity than stocks in otherwise equivalent companies. Itzkowitz, et al. (2016) also demonstrate that stocks with names that are ranked earlier in the alphabet experience higher trading volumes. They present suggestive evidence that these stocks also have higher valuations. Hartzmark

[3] In the contexts of Dietrich (2008) and Haque and Ginsparg (2009), papers are listed in order of submission. Both consider the possibility that papers submitted earlier in the submission window are of higher quality from those submitted later. However, neither finds any evidence to suggest that this concern is substantive.

[4] However, in Ginsburgh and van Ours (2003), the first position is at the greatest disadvantage. One possible explanation is that judges are not fully-informed about the competitive task when the competition begins. They learn as they observe successive competitors.

(2015) demonstrates that, within a portfolio, the stock with the initial closest to the front of the alphabet is most likely to be sold.[5]

Alphabetic ordering effects are exploitable. Ang, Chua and Jiang (2010) demonstrate that the difference in value between company shares that are equivalent with the exception of differential voting rights is less when the shares with inferior rights are designated as "A" shares and those with superior rights are designated as "B" shares than when the designations are reversed. McDevitt (2014) predicts and verifies that businesses that adopt artificial names for the purpose of appearing at the front of alphabetical listings target infrequent customers, charge higher prices and provide inferior service.

The ordering effects in academic publishing discussed above are amplified by alphabetical rankings. Richardson (2008) finds that journals disproportionately request reviews from referees with surname initials towards the beginning of the alphabet. Arsenault and Larivière (2015) present evidence that authors whose surname initials are ranked towards the front of the alphabet are more likely to be cited. Huang (2015) provides similar evidence, and demonstrates that the tendency to disproportionately cite papers by these authors is greater in disciplines where reference lists tend to be longer. This suggests that citations are generated by search procedures that favor papers encountered earlier.

In economics, these effects are compounded by the tendency to list authors alphabetically in multi-authored papers. Frandsen and Nicolaisen (2010, 613) report that, for publications in economics between 1978 and 2007, authorship was alphabetized in "roughly three-fourths" of multi-authored papers. Waltman (2012) reports that this percentage was 72.3 % between 2007 and 2011. Therefore, economists with surname initials ranked towards the beginning of the alphabet are

[5] However, Hartzmark (2015) demonstrates that the stock with the initial closest to the end of the alphabet also has a significantly higher likelihood of being sold. This effect has not been assessed elsewhere.

disproportionately likely to be first authors, their papers are disproportionately likely to appear towards the front of reference lists, and, presumably, disproportionately likely to be cited.

Moreover, Maciejovsky, et al. (2009) present evidence that, under alphabetical author ordering, economists tend to assign slightly less credit to authors in later positions. As a consequence, Einav and Yariv (2006) demonstrate that, within the highest ranked economics departments, faculty with initials that occur earlier in the alphabet are more likely to be tenured.[6]

These effects do not, however, extend to the highest professional recognitions. Hamermesh and Pfann (2012) find no significant relationship between alphabetical rank of surname initial and membership in the Econometric Society, receipt of honors from the American Economic Association, receipt of the John Bates Clark Award or the Nobel Memorial Prize in Economic Sciences.

As in the business context, alphabetical ordering effects in academic publishing invoke strategic responses. Ackerman and Brânzei (forthcoming) identify two conflicting incentives associated with alphabetical author lists. Authors with surname initials towards the front of the alphabet may shirk, because their effort will not affect their rank in the listing. However, if authors with surname initials towards the end of the alphabet are to have coauthors with earlier surname initials, they will choose those who do not shirk.[7]

[6] This effect appears to be driven by the convention of listing authors of multi-authored papers alphabetically, rather than differences in faculty behavior across surname initials. It also suggests that the highest ranked departments may, incorrectly, interpret authorship order as correlated with contribution rather than alphabetic rank.

[7] Similarly, Ackerman and Brânzei (forthcoming) identify two conflicting incentives with ordering authors by contribution. This ordering creates an obvious incentive to increase individual contributions. However, increased contributions that do not exceed the contribution of the author ranked just ahead do not alter the order. Therefore, this ordering also creates an incentive to contribute only slightly more than the author ranked just behind. In their model, these incentives can lead to greater shirking than under alphabetic ordering.

There is clear empirical evidence, cited above, that authors with surname initials towards the beginning of the alphabet receive amplified attention and credit. However, Efthyvoulou (2008) suggests that alphabetical authorship subjects them to a disadvantage as well. Were they to have made the largest contribution, their rank in an authorship list by contribution would be identical to their rank in an authorship list by alphabet. Therefore, their true contribution cannot be revealed in lists by alphabet and are difficult to reveal in lists by contribution.

Perhaps as a consequence, alphabetical authorship listings across multiple scholarly fields have become less common. Waltman (2012) demonstrates that the frequency of alphabetical authorship listings in multi-authored papers diminished across all areas of scholarship from 32.2% in 1981 to 15.9% in 2011.

Waltman (2012) also attributes this, in part, to increasing numbers of co-authors.[8] Frandsen and Nicolaisen (2010) confirm that alphabetical orderings become less common as the number of co-authors increases. This reduction occurs because alphabetical orderings are less likely to coincide with other ordering principles as the number of authors increases. In addition, authors with surname initials far from the beginning of the alphabet prefer collaborations where authorship lists follow some non-alphabetic ordering (van Praag and van Praag, 2008).

However, as shown above, alphabetization remains common among economists. Consequently, economists with surname initials ranked further from the beginning of the alphabet invoke at least three responses (van Praag and van Praag, 2008): First, they are more likely to choose non-alphabetical authorial orderings.

[8]As examples, Frandsen and Nicolaisen (2010) demonstrate that the shares of papers in both economics and information science with multiple authors increased from approximately one-third in 1978 to approximately two-thirds in 2007. The share of co-authored papers in high energy physics increased from approximately 72% to approximately 83%.

Second, they strategically choose their co-authorship strategies. Einav and Yariv (2006) and Kadel and Walter (2015) demonstrate that economists with surname initials ranked further from the beginning of the alphabet avoid participating in papers with more than three authors.

Moreover, Ong, et al. (2016) assert that, among authors with surname initials that are farther from the beginning of the alphabet, those with greater skill have an incentive to author singly. This reduces the risk that they will share credit with a less talented author who would precede in alphabetical order. This is consistent with citation counts, which, among single-authored papers, are greater for authors whose surname initial is further from the beginning of the alphabet.

In contrast, Ong, et al. (2016) assert that, among authors with surname initials that are closer to the beginning of the alphabet, those with greater skill are more likely to co-author. Because of their greater skill, they are more likely to find a skilled coauthor who is nevertheless content to receive second listing. Correspondingly, double-authored papers receive more citations as the first author's surname initial moves closer to the beginning of the alphabet.

Third, economists with surname initials further from the beginning of the alphabet tend to be less productive (van Praag and van Praag). This reflects an endogenous response to systematic disadvantage. Those who are consistently in later ordering ranks will have fewer incentives to invest in the skills necessary to take advantage of opportunities, should they arise.

Evidence of alphabetic bias arises outside of academic publishing, where choices occur between alternative individuals who are of roughly equivalent merit. The positive correlation between surname initial and secondary school test scores in Czechoslovakia is consistent with the hypothesis that these schools admit marginal students in alphabetical order (Jurajda and Münich, 2010). Potential donors with surname initials further from the beginning of the alphabet are less likely to make charitable donations, presumably because they are less likely to receive personal solicitations (Rosen and Meer, 2011). Law school faculty with surname initials closer to the

beginning of the alphabet are more likely to receive invitations to visit at other institutions (Merritt, 1999).

c. Name effects as secondary

In all of the examples of alphabetic ordering above, the ordering itself is never substantive. The primary characteristic of interest would be, in the context of businesses, economic returns. In the case of academic publishing, it would be the quality and relevance of an article. Ordering should be influential only when primary characteristics are difficult to assess or appear to be similar across options.

As examples, the ballot effects in Ho and Imai (2008) are most important in races and for candidates that attract little attention. The trading volume and liquidity effects in Jacobs and Hillert (2016) are most important for stocks in companies that are of lesser prominence. The effects associated with paternal and father-in-law origin in Rubinstein and Brenner (2014) are markedly stronger for individuals whose imputed skin tone is less indicative regarding this origin.[9]

Conversely, in contexts where the primary characteristics of interest can be readily assessed, ordering should be unimportant. Alphabetic effects may be absent in Hamermesh and Plann (2012) because, in the comparisons among outstanding economists, records of accomplishment are substantial and the numbers of such records are relatively few. In contexts such as these, where searches are over fewer options, each characterized by extensive relevant information, alphabetic orderings may be irrelevant.

[9] Similarly, Derous, et al. (forthcoming) suggest that skin tone is a more important indicator of identity than is name.

3. Models and Data

The intent here is to expand the investigation of alphabetic ordering effects beyond the limited domains described in the previous section. The next sections estimate these effects with respect to twelve individual outcomes representing experiences in high school, college and the labor force.

Section 4 estimates common effects of surname initial on all members of the sample described here. The regression equation employed for this purpose is model 1, where y_j represents each of the twelve dependent variables and *j* indexes sample members:

$$y_j = \beta_0 + \beta_{IQ} IQ_j + \beta_\alpha \alpha_j + \beta_{APPEARANCE} APPEARANCE_j + \boldsymbol{X}_j' \gamma + \varepsilon_j \quad (1)$$

The explanatory variable of interest is α_j, the index for surname initial, defined below. The coefficient of interest is β_α, the effect of surname initial.

The control variables IQ_j and $APPEARANCE_j$ measure, respectively, cognitive ability and physical appearance. They are distinguished from the rest of the control variables, X_j, because of the possibility that they may interact with surname initial. As discussed at the end of the previous section, alphabetic rank may be most important for individuals who are not distinguished in other relevant dimensions. Cognitive ability and physical appearance are, arguably, dimensions that are relevant to human capital accumulation and labor market success. Section 5 explores the hypothesis that alphabetic ordering is most important for those who are not distinctive in either of these dimensions.

For this purpose, section 5 employs model 2:

$$y_j = \sum_{i\in I}\sum_{r\in R} \beta_{ir}(i_j r_j) + \sum_{i\in I}\sum_{r\in R}\sum_{k=1}^{2} \beta_{ir,k}(i_j r_j) x_{k,j} + \beta_{M_I M_R,\alpha}(M_I M_R)\alpha_j + \sum_{\substack{i\in I \\ (I\times R)\setminus\{M_I M_R\}}}\sum_{r\in R} \beta_{ir,\alpha}\,(i_j r_j)\alpha_j + \boldsymbol{X}_j'\gamma + \varepsilon_j \quad (2)$$

Sets $I = \{L_I, M_I, H_I\}$ and $R = \{L_R, M_R, H_R\}$ each consist of three hierarchical divisions of IQ and attractiveness, respectively. These subsets partition the full sample into low (L), intermediate (M), and high (H) levels for both. The cross between these two partitions yields nine subsamples, representing all combinations of low, intermediate and high IQ with low, intermediate and high appearance rating.

In this specification, $\beta_{ir}(i_j r_j)$ represent subsample-specific fixed effects. $\beta_{ir,k}(i_j r_j)$ represent subsample-specific coefficients for $x_{k,j}$, where $x_{1,j}$ and $x_{2,j}$ are IQ and attractiveness measures of individual *j*, respectively. $\beta_{ir,\alpha}(i_j r_j)$ are the coefficients for surname initial rank in the eight subsamples apart from that representing those with intermediate IQ and attractiveness scores. The coefficient for this last subsample, $(M_I M_R)$, is $\beta_{M_I M_R,\alpha}$, the effect of interest.

The explanatory variables $\mathbf{X}_j$ follow those in Zax and Rees (2002). As there, the analytical posture consists of observing each individual as they graduate from high school and predicting subsequent outcomes. Consequently, these variables describe individuals at that graduation.[10] While post-graduation choices may affect outcomes of interest that occurred further into their adult lives, the analysis here captures the effects of these choices in the characteristics at high school graduation upon which they were based.

The Wisconsin Longitudinal Study, or WLS (Herd, et al., 2014; http://www.ssc.wisc.edu/wlsresearch) provides the data employed here. The WLS population consists of 10,137 individuals, representing a random sample comprising one-third of all seniors graduating from high school in 1957 in Wisconsin. These individuals have been surveyed

[10] This construction holds constant completed education. All sample members are high school graduates at the time of observation for explanatory variables. None have had the opportunity yet to enroll in tertiary training.

intermittently from 1957 through 2011. The sample here consists of 3,281 males with complete data for all individual and family explanatory variables employed below.

Table 1 presents summary statistics for the explanatory variables that measure characteristics of the individual. Two of the variables, IQ score and high school rank, are direct measures of human capital. A third, measuring friends' intentions to attend college, serves as a proxy for the individual's ambitions regarding the acquisition of additional human capital.

IQ represents the individual's score on the Henmon-Nelson Test of Mental Ability, administered in the eleventh grade. It is normed so as to achieve an average score of 100, with a standard deviation of 15 (Gottfredson, 2009). Table 1 demonstrates that the sample here closely approximated these norms. The range of IQ scores was from 61 to 145, including individuals with limited and exceptional cognitive abilities. These abilities were, presumably, relatively fixed.

High school rank measures human capital accumulation during high school. It is the individual's percentile rank in his high school class upon graduation. The average of 45.5 indicates that this sample is skewed slightly towards those whose high school performance was weaker, presumably because it omits women. However, the range for this variable encompasses the entire range of possible values, from zero to 99. Zax and Rees (2002) argue that, in comparing students with the same IQ score and from the same high school class, differences in rank are most plausibly interpreted as reflecting differences in chosen effort.

"Friends' plans to attend college" is a binary recode of the WLS respondent's response to "What are most of your friends doing after high school?". This variable assigns the value of one to any response indicating intentions to continue schooling. The individual's own plans regarding college were presumably correlated positively with those of his friends.

The remaining three variables, "attractiveness rating", "relative body mass" and alphabetic rank of surname initial, measure personal characteristics that are not, themselves, important

Table 1. Summary statistics for individual explanatory variables

Variable	Mean	SD	Min	Max
Individual characteristics				
Alphabetical rank of surname initial	11.839	6.799	1	26
IQ	101.863	15.082	61	145
Facial attractiveness rating	0.0793	1.303	-4.011	4.149
Relative body mass - proxy for BMI	0.0161	0.829	-2.969	3.619
High school				
High school rank	45.5	28.115	0	99
Post-secondary education				
Friends' plan to attend college	0.404	0.491	0	1

The sample consists of 3,281 men.

components of human capital. They are, instead, characteristics that may affect participation in social interactions. These effects could arise because these "presentational characteristics" affect the individual's sense of social efficacy or elicit distinctive responses from others. The role of these characteristics in these interactions may therefore affect employment or returns to human capital.

The "attractiveness rating" and "relative body mass" variables both derive from visual examinations of high school year book photographs for the WLS subjects. The attractiveness rating is the WLS variable "meanrat_fcoder". It is the demeaned average of attractiveness ratings on an 11-point scale assigned by six female raters from approximately the same age cohort as the WLS respondents. "Relative body mass" is the WLS variable "srbmi". It is the average of body mass assessments assigned by three young female and three young male raters on an 11-point scale and then transformed into rater-specific Z-scores.

Lastly, "alphabetical rank of surname initial" is the explanatory variable of interest.[11] It represents a simple numerical correspondence between the letters of the alphabet, ordered

[11] The WLS provided surname initials to this study under strict confidentiality restrictions.

conventionally as "A" through "Z", and the ordered integers from one to 26. The average value of this variable, 11.8, indicates that "typical" surnames began with the letters "K" or "L".[12]

The assumption of linearity embodied in this transformation may appear restrictive. However, the intuitions that motivate this investigation are too general to imply any specific transformation. A fully non-parametric specification, consisting of letter fixed effects, is too cumbersome to be empirically useful. Other specifications, such as fixed effects for groups of adjacent letters, may relax the linearity assumption across groups but at the cost of an equality assumption within groups. The transformation here is, at least to some degree, validated by its performance in the regressions below.

Table 2 presents summary statistics for the explanatory variables that measure characteristics of the individual's family. With the exception of number of siblings, all variables are categorical. Of them, those measuring parental attitudes towards college attendance were presumably correlated positively with their sons' college ambitions.[13] The remainder, which describe household structure, parental educations and occupations, household income and father's ethnic background, describe basic characteristics of the household.

Fewer than 10% of households contained only one parent. Fewer than 10% of both fathers and mothers had college degrees. A large majority, 61.4% of individuals, reported that their parents encouraged them to attend college. The omitted category for household income consists of those with missing values for this variable, comprising 12.4% of the sample.

[12] Einav and Yariv (2006) and Ong, et al. (2016) employ the same assignment. Efthyvoulou (2008) employs the logarithm of this assignment. van Praag and van Praag (2008) employ both. Jurajda and Münich (2010) employ the numerical assignment, but also use the percentile of the last name by the alphabetical ranking. Huang (2015) employs the numerical assignment as well as fixed effects for groups of initials and for individual initials. Hamermesh and Pfann (2012) "hold constant for alphabetical location" without further explanation. Similarly, Merritt (1999) holds constant "alphabetic placement".

[13] The omitted category consists of parents who did not express opinions regarding college attendance.

Table 2. Summary statistics for household explanatory variables

Variable	Mean	SD	Min	Max
Household structure				
Both parents present	0.912	0.283	0	1
Number of siblings	3.086	2.472	0	26
Birth order	2.416	1.895	1	21
Father's education				
College	0.0954	0.294	0	1
High school	0.328	0.469	0	1
Missing	0.0698	0.255	0	1
Mother's education				
College	0.0933	0.291	0	1
High school	0.407	0.491	0	1
Missing	0.0749	0.263	0	1
Parental occupation				
Father has a white collar job	0.268	0.443	0	1
Mother has a white collar job	0.145	0.352	0	1
Household income				
Bottom 25%	0.207	0.405	0	1
Middle 50%	0.443	0.497	0	1
Top 25%	0.226	0.418	0	1
Below neighbors'	0.071	0.257	0	1
Above neighbors'	0.242	0.428	0	1
Parental attitude				
Parents encouraged college	0.614	0.487	0	1
Parents discouraged college	0.0305	0.172	0	1
Father's national/ethnic background				
British	0.109	0.311	0	1
Eastern European	0.0491	0.216	0	1
French	0.0463	0.21	0	1
German	0.489	0.499	0	1
Irish	0.0658	0.248	0	1
Mediterranean	0.0155	0.124	0	1
Polish	0.0637	0.244	0	1
Scandinavian	0.143	0.35	0	1
Minority	0.00518	0.0718	0	1
Missing	0.0131	0.114	0	1

All monetary variables are in 1992 dollars. The sample consists of 3,281 men.

The variable for father's national heritage differs substantially from more typical measures of race or ethnicity. The WLS, because of its geographic and temporal sampling frame, contains very few individuals with African-American or Hispanic heritage. As reported in table 2, "minorities" in

Table 3. Most common surname initials by nationality

Father's national/ethnic background	First	Frequency	Second	Frequency	Third	Frequency	N
British	S	0.104	H	0.096	C	0.087	357
Eastern European	B	0.106	K	0.099	S	0.099	370
French	D	0.132	L	0.132	B	0.105	152
German	S	0.153	B	0.104	K	0.096	1,606
Irish	M	0.167	C	0.097	D	0.083	216
Mediterranean	R	0.137	S	0.137	B	0.117	51
Polish	S	0.196	K	0.139	B	0.1	209
Scandinavian	S	0.102	J	0.096	H	0.085	469
Minority	C	0.177	H	0.177	P	0.118	17
Missing	B	0.117	H	0.116	B	0.093	44

Relative frequency distribution of surname initials (Table 3) is derived from our WLS sample. N = 3,281

the conventional sense comprise less than one percent of the sample. The conventional concerns with differences in outcomes that may be attributable to substantive racial or ethnic discrimination are, therefore, not relevant here.

The important distinctions in national heritage are largely between those with different European origins. While these distinctions are not generally associated with different experiences of discrimination, they may be relevant here because they could be associated with systematic differences in names, naming conventions, and therefore surname initials.

Table 3 demonstrates that the most common surname initials vary substantially across categories of national origin. In order to purge estimated surname initial effects of any influence arising from other attributes associated with national origin, models 1 and 2 include fixed effects for all of these national origin categories.

Table 4 presents summary statistics for the twelve dependent variables examined in the analysis below. Two of these variables measure outcomes of the high school experience. "Outstanding student" is a binary variable that represents the "Teacher's evaluation of graduate" and assigns the value of one to the response "Outstanding". "Favorable opinion of high school

Table 4. Summary statistics for individual dependent variables

Variable	Mean	SD	Min	Max
High school				
Outstanding Student	0.113	0.316	0	1
Favorable Opinion on high school classes	0.564	0.496	0	1
Post-secondary education				
Applied to college	0.341	0.474	0	1
Withdrew from college	0.344	0.475	0	1
Received a post-high school degree	0.444	0.497	0	1
Labor market				
Military service	0.504	0.5	0	1
Income score for first job	270.756	236.823	0	877
Siegel prestige score for first job	396.858	165.714	144	812
1974 employment earnings ($10,000s)	4.117	2.584	0	28.458
Siegel prestige score for employment in 1974	462.359	135.313	156	812
1992 employment earnings ($10,000s)	6.242	28.435	0	999.999
Siegel prestige score for employment in 1992	465.374	139.568	154	812

All monetary variables are in 1992 dollars. The sample consists of 3,281 men.

studies" is a binary recode of the WLS subject's response to the question "What is your opinion of your high school studies" with the value of one representing "Interesting, want to learn more".

Three outcome variables measure the individual's experience with tertiary education. "Applied to college" is a binary variable indicating whether the individual had applied to college in 1957. "Withdrew from college" is a binary variable indicating that the individual attended post-secondary school but did not report receipt of a degree. "Received a post-high school degree" is a binary variable indicating whether the individual had earned any tertiary degree as of 1992.

The remaining seven variables measure labor market experiences. Three, military service, income score and Siegel prestige score for the first job characterize the individual's first experiences. Two variables, income and Siegel prestige score, characterize the individual's employment, if any, in 1974 and 1992, at approximately ages 35 and 53.

4. The effects of alphabetism

This section estimates model 1 in order to examine the effects of surname initial rank on individual experiences in high school, participation in tertiary education, labor market activity as a young adult and later, in mid-career. All regressions include fixed effects for high school. These effects are necessary to imbue high school rank with meaning, because different standards across high schools imply that ranks in different schools are not directly comparable (Zax and Rees, 2002). In addition, here these fixed effects control for any systematic differences across high schools in the photographic techniques employed for yearbook pictures, upon which the attractiveness and body mass variables are based.

Table 5 presents estimates of equation 1 for the two high school outcome variables, whether an individual was recognized as an "outstanding student" and whether a student evaluated his high school classes favorably. The first represents an external evaluation of the student's high school performance. The second represents a self-reported evaluation of the high school experience. Both dependent variables are categorical. Accordingly, both regressions in table 5 are linear probability models.[14]

The equation for "outstanding student" demonstrates that, as would be expected, individuals with higher IQs and with higher high school ranks were significantly more likely to be identified as outstanding students. The same was true for those with friends planning to attend college and those whose parents encouraged college attendance, presumably reflecting shared ambitions.

[14] Variations in sample sizes across regressions here and in the following tables are attributable, with one exception, to differing incidences of missing values for the dependent variables. The exception occurs in table 6.

Table 5. Alphabetism in high school

Explanatory variables	Outstanding student	Opinion on high school classes
Individual characteristics		
Alphabetical rank of surname initial	-0.00128*	-0.00219*
	(0.000680)	(0.00122)
IQ	0.00199***	-7.22e-05
	(0.000502)	(0.000731)
High school rank	0.00354***	0.00461***
	(0.000345)	(0.000390)
Attractiveness	0.000555	-0.00896
	(0.00323)	(0.00680)
Relative body mass - proxy for BMI	0.00591	0.0129
	(0.00651)	(0.0111)
Friends' plan to attend college	0.0209*	0.162***
	(0.0124)	(0.0204)
Household characteristics		
Household structure		
Both parents present	0.0154	-0.0116
	(0.0193)	(0.0295)
Number of siblings	-0.00179	0.0105**
	(0.00222)	(0.00457)
Birth order	-0.000142	-0.0137***
	(0.00328)	(0.00515)
Father's education		
College	0.00857	0.0496*
	(0.0250)	(0.0282)
High school	-0.00388	0.0153
	(0.0114)	(0.0207)
Missing	0.0124	-0.0194
	(0.0193)	(0.0382)
Mother's education		
College	0.00517	-0.0138
	(0.0222)	(0.0307)
High school	-0.00973	0.0189
	(0.0108)	(0.0208)
Missing	0.00299	-0.0232
	(0.0206)	(0.0389)
Parental occupation		
Father has a white collar job	-0.0153	-0.00443
	(0.0150)	(0.0229)
Mother has a white collar job	0.0220	0.0337
	(0.0150)	(0.0230)

Table 5. Continued

Explanatory variables	Outstanding student	Opinion on high school classes
Household income		
Bottom 25%	0.00765	-0.00365
	(0.0144)	(0.0236)
Top 25%	0.0220	0.0361*
	(0.0149)	(0.0186)
Missing	0.00798	0.0341
	(0.0171)	(0.0310)
Below neighbors'	0.0215	-0.0339
	(0.0207)	(0.0332)
Above neighbors'	0.00493	-0.00807
	(0.0125)	(0.0210)
Parental attitude		
Parents encouraged college	0.0228*	0.215***
	(0.0120)	(0.0246)
Parents discouraged college	-0.0177	0.0678
	(0.0250)	(0.0543)
Father's national/ethnic background		
British	0.0214	0.0651
	(0.0205)	(0.0420)
Eastern European	0.00645	0.0177
	(0.0275)	(0.0479)
French	0.0777***	0.0333
	(0.0274)	(0.0541)
German	0.0214	0.0267
	(0.0164)	(0.0350)
Irish	-0.0396**	0.0204
	(0.0196)	(0.0490)
Mediterranean	0.0215	0.0689
	(0.0399)	(0.0658)
Scandinavian	0.00638	0.0138
	(0.0174)	(0.0406)
Minority	-0.0317	0.309***
	(0.0733)	(0.0882)
Missing	0.0344	0.0307
	(0.0490)	(0.0905)
Constant	-0.288***	0.144*
	(0.0541)	(0.0851)
Observations	3,281	3,196
R^2	0.198	0.244
High school FE	Y	Y

Standard errors are clustered at high school. *** p<0.01, ** p<0.05, * p<0.1

The estimated effects of both appearance and body mass were statistically insignificant. It seems plausible that either or both may have affected student interactions with high school colleagues and staff. These results suggest that any such effects were relegated to the social domain.

Of the remaining household characteristics, only two nationality effects were statistically significant. The absence of these effects suggest that recognition as "outstanding" was based almost entirely, and appropriately, on ability, ambition and performance.

These reassuring inferences do not, however, extend to the variable of principal interest. The effect of alphabetical rank of surname initial on this designation is significant, negative and substantively large. Two otherwise identical students whose surname initials differed in rank by ten places, the difference between "A" and "K", "H" and "S", or "N and "Y", as examples, would have differed in their probability of designation as outstanding by 1.28 percentage points. As the average probability of designation, from table 4, is 11.3%, this effect reduced the probability of designation for the student with the surname initial furthest from the beginning of the alphabet by more than 10%.

The regression for "opinion on high school classes" shares important similarities with that for "outstanding student". Students with higher high school rank, friends who intended to attend college and parents who encouraged college attendance were more likely to have favorable opinions about their high school classes.[15] Opinions were not affected by either appearance or body mass.

However, IQ had no effect on student opinions regarding their classes. It seems plausible that, unconditionally, differences in cognitive ability would have been associated with differences in

[15] The coefficients for number of siblings and birth order in this regression are both significant, of similar magnitude and opposite sign. Together, they imply that the addition of an older sibling, which would have increased both the number of siblings and birth order, would have had no substantive effect. However, the addition of a younger sibling, which would have increased the number of siblings but would not have changed birth order, induced a more favorable opinion regarding high school studies.

opinions regarding classes. For example, cognitive ability and appreciation for challenging courses might have been positively correlated. However, this correlation might have been of limited relevance if students of different abilities took different courses. Its relevance may have been further limited by the regression specification, which compares the effects of differences in cognitive ability for those whose high school performance and college ambitions were the same.

Regardless, the effect of alphabetic rank of surname initial on student opinions regarding their courses was, once again, significant and negative. Substantively, though, it was less important. Two otherwise identical students whose surname initials differed in rank by ten places would have differed in their probability of expressing favorable opinions of their courses by 2.19 percentage points. As the average probability of favorable opinions, from table 4, was 56.4%, this effect reduced the probability of a favorable opinion by the student with the surname initial furthest from the beginning of the alphabet by less than 5%.

Table 5 demonstrates that the alphabetical rank of surname initial affected both teacher evaluations of high school students and students' evaluations of their high school experience. Both evaluations may have been subject to distortions related to student characteristics that were not directly relevant to academic performance. However, the regressions in table 5 hold constant the non-academic characteristics that were most likely to have been salient, facial attractiveness and body mass. Consequently, the estimated effects of surname initial, themselves, are likely to capture the actual effect of alphabetic rank.

In addition, the substantive differences in the surname initial effects of table 5 may be informative regarding the behavioral mechanisms, discussed in the section 2, by which those effects might arise. Teachers were responsible for designation as an "outstanding student". The large effect of surname initial on the probability of achieving this designation suggests that, for teachers, ordering effects were important. In contrast, students were responsible for evaluating their courses.

Table 6. Post-secondary educational attainment

Explanatory variables	Applied to college	Withdrew from college	Received post-high school degree
Alphabetical rank of surname initial	-0.00293***	0.00562***	-0.00267**
	(0.00106)	(0.00184)	(0.00107)
IQ	-0.000458	-0.00174	0.00256***
	(0.000634)	(0.00108)	(0.000712)
High school rank	0.00294***	-0.00493***	0.00374***
	(0.000360)	(0.000601)	(0.000395)
Attractiveness	0.00646	-0.000214	-0.00406
	(0.00532)	(0.00866)	(0.00666)
Relative body mass - proxy for BMI	0.0117	0.000242	-0.000610
	(0.00908)	(0.0150)	(0.00911)
Friends' plans to attend college	0.156***	-0.0654**	0.108***
	(0.0209)	(0.0253)	(0.0176)
Parents encouraged college	0.0730***	-0.0282	0.145***
	(0.0177)	(0.0420)	(0.0191)
Observations	3,281	1,610	3,280
R^2	0.161	0.182	0.226
Additional household controls	Y	Y	Y
High school FE	Y	Y	Y

Standard errors are clustered at high school. *** p<0.01, ** p<0.05, * p<0.1

The smaller effect of surname initial on the probability of a favorable opinion suggests that students with surname initials further from the beginning of the alphabet were able to cope, at least to some degree, with the associated disadvantages.

Table 6 explores the relationships between the explanatory variables of greatest interest and tertiary education outcomes. It presents estimates from linear probability models for the probabilities of applying to college prior to high school graduation, withdrawing from college if ever enrolled prior to 1992, and receiving a college degree by 1992.[16] These estimates reinforce the themes apparent in table 5.

[16] The authors can provide complete results. Among the explanatory variables not presented in table 6, an individual was significantly more likely to apply to college if his father had graduated from college, significantly less likely to withdraw from college if either father or mother had graduated from college and significantly more likely to earn a college degree if either father or mother had graduated from college. Other explanatory variables did not display consistent significant effects. The sample for the regression analyzing withdrawal from college is restricted to those who ever enrolled.

As in table 5, better performance in high school was associated with more favorable outcomes. Those with higher high school ranks subsequently accumulated more human capital: they were significantly more likely to apply to college, significantly less likely to withdraw and significantly more likely to receive a college degree. The same was true for those whose friends intended to attend college. Those whose parents encouraged college attendance were significantly more likely to apply and to graduate. Holding constant high school performance and proxies for college ambitions, higher IQs were significantly associated with only higher probabilities of receiving a college degree.

Facial attractiveness and relative body mass had no significant effects on college outcomes. Their absence reinforces the implication of table 5 regarding these variables. Any relevance they may have had to the experience of young adults does not appear to have affected their investments in human capital.

However, this again did not hold for surname initial. As in table 5, individuals with surname initials ranked later in the alphabet had consistently inferior outcomes. The coefficients for surname initial rank are significant for all three outcomes. These coefficients imply that a difference of ten ranks in surname initial was associated with a reduction of 2.93 percentage points in the probability of applying to college, an increase of 5.62 percentage points in the probability of withdrawing after enrolling, and a reduction of 2.67 percentage points in the probability of receiving a college degree. Compared to the average probabilities from table 4 of, respectively, 34.1%, 34.4% and 44.4%, each of these differences was substantively large.

Table 7 explores the relationships between the explanatory variables of greatest interest and early employment outcomes. It presents a linear probability model for the probability of serving in the military.[17] It also presents regressions which describe the natural logarithm of the income

[17] Military service is a binary variable with one indicating an affirmative response to the question "Respondent ever been on active duty in the U.S. military or spent at least two months on active

Table 7. Initial employment

Explanatory variables	Military service	First employment	
		Income score	Prestige score
Alphabetical rank of surname initial	0.00465***	-0.00424	-0.749*
	(0.00148)	(0.00281)	(0.390)
IQ	0.00194**	0.00760***	1.065***
	(0.000771)	(0.00175)	(0.243)
High school rank	-0.00197***	0.00810***	1.765***
	(0.000402)	(0.000821)	(0.127)
Attractiveness	0.000988	-0.00190	-0.163
	(0.00798)	(0.0152)	(2.191)
Relative body mass - proxy for BMI	-0.00415	0.0181	-2.170
	(0.0118)	(0.0239)	(3.312)
Friends' plan to attend college	-0.0706***	0.190***	40.47***
	(0.0231)	(0.0461)	(6.759)
Parents encouraged college	-0.0634***	0.299***	47.75***
	(0.0211)	(0.0462)	(6.375)
Observations	3,281	3,086	3,087
R^2	0.035	0.213	0.311
Additional household controls	Y	Y	Y
High school FE	Y	Y	Y

Income score is in natural log. Standard errors are clustered at high school. *** p<0.01, ** p<0.05, * p<0.1

score[18] and the Siegel Occupational Prestige Score for the first job. These estimates suggest that employment was a superior option to military service. They further support the themes apparent in table 5.

Students with higher high school ranks were significantly less likely to have military experience, and had significantly higher incomes and prestige scores for their first job. The same was true for students whose friends intended to attend college and whose parents encouraged them to attend college. These results indicate that the more accomplished and ambitious were more likely to avoid military service and obtain better entry-level employment.

duty for training in the Reserves or National Guard?" The question was asked in 1992-3, so it is possible that some affirmative answers refer to military service later in life.

[18] The WLS documentation does not offer a thorough description of this variable, "ocix1" (http://www.ssc.wisc.edu/wlsresearch/documentation/waves/?wave=wls75&module=cjobh). It appears to be the median income of workers in an individual's occupation.

Holding constant high school performance and ambition, students with higher IQ scores were significantly more likely to serve in the military, and obtained first jobs with significantly higher income and prestige scores. All three of these effects seem plausible as demand-side responses: among those with equivalent high school performance and subsequent ambition, both the military and employers preferred individuals with greater cognitive ability.

Once again, of the three variables measuring presentational characteristics, neither attractiveness nor relative body mass had significant effects on any of the table 7 outcomes. However, surname initial continued to exert influence, and continued to do so by conferring advantages on those with initials closer to the front of the alphabet. The coefficient on surname initial is significantly positive in the linear probability model for military service and significantly negative in the regression for the prestige score of the first job. An increase of ten in alphabetic rank increased the probability of military service by 4.65 percentage points, or nearly one-tenth of the average probability of 50.4%. The same increase in alphabetic rank reduced the prestige score by 7.49 points, or approximately two percent of the average score, 396.9.

Table 8 explores the determinants of earnings and prestige scores for employment in 1974, at approximately age 35, and in 1992, at approximately age 53. Broadly, measures of innate human capital, high school effort and proxies for ambition at the end of high school were all associated with superior outcomes in both years. In contrast, presentational characteristics had become generally less influential.

As in all previous regressions, higher high school ranks were significantly associated with better outcomes at both ages. The same was true for ambition, as proxied by friends' plans for college and parents encouragement for college enrollment, with the exception of the insignificant coefficient for the latter in the regression for 1992 log earnings. In addition, higher IQ scores were significantly associated with higher earnings and prestige in both years.

Table 8. Employment in adulthood

Explanatory variables	Employment in 1974		Employment in 1992	
	Earnings	Prestige score	Earnings	Prestige score
Alphabetical rank of surname initial	-2.13e-05	0.214	0.000858	-0.203
	(0.00145)	(0.324)	(0.00238)	(0.353)
IQ	0.00294***	1.199***	0.00644***	1.432***
	(0.000920)	(0.206)	(0.00169)	(0.199)
High school rank	0.00144***	1.287***	0.00302***	0.895***
	(0.000423)	(0.107)	(0.000797)	(0.120)
Attractiveness	0.0131	3.359*	0.0119	0.478
	(0.00831)	(1.727)	(0.0126)	(2.080)
Relative body mass - proxy for BMI	-0.00289	-3.208	-0.00679	-4.887*
	(0.0111)	(2.614)	(0.0201)	(2.818)
Friends' plan to attend college	0.0689***	20.71***	0.0835**	22.43***
	(0.0243)	(5.445)	(0.0356)	(6.021)
Parents encouraged college	0.0686***	35.21***	0.0406	41.46***
	(0.0237)	(5.484)	(0.0381)	(6.362)
Observations	2,694	3,220	2,426	2,863
R^2	0.077	0.262	0.092	0.208
Additional household controls	Y	Y	Y	Y
High school FE	Y	Y	Y	Y

Earnings are in natural logs. Standard errors are clustered at high school. *** p<0.01, ** p<0.05, * p<0.1

As in all regressions for tables 5, 6 and 7, both attractiveness and relative body mass make no statistically significant contributions to three of the four regressions in table 7. The former is significant in the regression for 1974 prestige score, and the latter in the regression for the 1992 prestige score. These coefficients indicate, respectively, that more attractive and less massive individuals had jobs with greater prestige. Given the absence of similar effects in all previous regressions, these results probably do not indicate systematic effects.

In contrast to all regressions for tables 5, 6 and 7, alphabetical rank surname initial is insignificant in all table 8 regressions. This implies that the effects of surname initial rank dissipated as adulthood progressed. Presumably, the accumulation of productive skills and a record of accomplishment eventually superseded any ordering or conditioning effects associated with alphabetic rank.

5. The interaction between other forms of distinction and surname initial

Section 4 estimates effects for alphabetic rank of surname initial that are common to all sample members. However, section 2 suggests that a characteristic that is not of primary salience may be unimportant for those who are distinguished in more salient dimensions. The previous section offered implicit support for this suggestion, in that surname initial rank had no effects on labor market outcomes for older adults, who presumably had informative work histories.

This section explores this suggestion in greater detail. It assumes that cognitive ability and appearance were of primary importance. Accordingly, it distinguishes between the effects of surname initial rank on those who had distinctive cognitive abilities and appearances, and those who did not.

For the purpose of the analysis below, individuals with low IQ scores were those whose scores were more than one standard deviation below the sample average. Intermediate scores were those within one standard deviation of the sample average. High IQ scores were those that were more than one standard deviation above the sample average. Similarly, low, intermediate and high attractiveness scores were those that were, respectively, more than one standard deviation below, within one standard deviation of, and more than one standard deviation above the sample average. In the notation of section 3,

$$\text{Individual } j \in \begin{cases} L_I \text{ if } IQ_j \in [61,86] \\ M_I \text{ if } IQ_j \in [87,117] \\ H_I \text{ if } IQ_j \in [118,145] \end{cases} \qquad \text{Individual } j \in \begin{cases} L_R \text{ if Attractiveness}_j \in [-4.011,-1.24) \\ M_R \text{ if Attractiveness}_j \in [-1.24,1.37) \\ H_R \text{ if Attractiveness}_j \in [1.37,4.149] \end{cases}$$

Table 9 Sample size by IQ-attractiveness strata

	Low IQ group	Intermediate IQ group	High IQ group	Total
Low attractiveness group	91 (2.77%)	337 (10.27%)	88 (2.68%)	516 (15.73%)
Intermediate attractiveness group	334 (10.18%)	1,533 (46.72%)	341 (10.39%)	2,208 (67.29%)
High attractiveness group	73 (2.23%)	377 (11.49%)	107 (3.26%)	557 (16.98%)
Total	498 (15.18%)	2,247 (68.49%)	536 (16.34%)	3,281 (100%)

Table 9 reports the numbers of individuals within each of the nine subsamples. The four subsamples with either high or low scores for both attractiveness and IQ together comprise 10.9% of the sample. The four subsamples with intermediate scores for one and high or low scores for the other comprise 42.3% of the sample. The subsample with intermediate scores for both IQ and attractiveness includes 46.7% of the sample.[19]

Table 10 presents results for the regressions of table 5 with this expanded specification. It reports only the coefficients for alphabetic rank of surname initial within each stratum of the IQ-by-appearance categorization, $\beta_{M_I M_R,\alpha}$.[20] These coefficients confirm that alphabetic rank of surname initial was important only for those who were not distinguished in terms of cognitive ability or appearance. Significant effects occur only for those who were of intermediate IQ and average

[19] Stratifications that place more of the sample in the extreme categories yield less distinctive results. Empirically, it seems that "within one standard deviation of the sample average" is an accurate implementation of "undistinguished". However, it is possible that some of the differences in statistical significance apparent in the following tables are the consequences of different subsample sizes, rather than differences in behavioral responses.

[20] The authors can provide complete results for these and all subsequent regressions.

Table 10. Alphabetism in high school by IQ-attractiveness strata

Explanatory variables	Outstanding student	Opinion on high school classes
Alphabetical rank of surname initial		
Low IQ group		
Low attractiveness group	0.00336	0.00458
	(0.00292)	(0.00907)
Intermediate attractiveness group	-0.00130	-0.00312
	(0.00170)	(0.00376)
High attractiveness group	0.000769	-0.00535
	(0.00246)	(0.00808)
Intermediate IQ group		
Low attractiveness group	0.000335	-0.000378
	(0.00171)	(0.00357)
Intermediate attractiveness group	-0.00233**	-0.00457***
	(0.00103)	(0.00170)
High attractiveness group	0.00325	0.000905
	(0.00224)	(0.00329)
High IQ group		
Low attractiveness group	-0.00380	-0.00194
	(0.00497)	(0.00643)
Intermediate attractiveness group	-0.00376	0.000773
	(0.00292)	(0.00327)
High attractiveness group	-0.000992	0.00464
	(0.00456)	(0.00587)
Observations	3,281	3,196
R^2	0.224	0.252
Controls	Y	Y
High school FE	Y	Y

Standard errors are clustered at high school. *** p<0.01, ** p<0.05, * p<0.1

attractiveness, in partition ($M_I M_R$). In magnitude, these effects are approximately double those estimated for the entire sample in table 4.

Table 11 presents the same elaboration on the regressions of table 6. Once again, the effects of surname initial rank in all three regressions are significant for those with intermediate IQ scores and average attractiveness. The signs are identical to those estimated for the entire sample in table 6. The coefficient magnitudes are larger, slightly so with regard to the probability of applying to

Table 11. Post-secondary educational attainment by IQ-attractiveness strata

Explanatory variables	Applied to college	Withdrew from college	Received post-high school degree
Alphabetical rank of surname initial			
Low IQ group			
Low attractiveness group	0.0144**	-0.0392**	0.00275
	(0.00680)	(0.0159)	(0.00836)
Intermediate attractiveness group	-0.000963	0.00918	0.00303
	(0.00357)	(0.0101)	(0.00369)
High attractiveness group	-7.69e-05	-0.0157*	0.00536
	(0.00877)	(0.00907)	(0.00815)
Intermediate IQ group			
Low attractiveness group	-0.00441	0.00240	0.00128
	(0.00342)	(0.00579)	(0.00334)
Intermediate attractiveness group	-0.00335*	0.00917***	-0.00591***
	(0.00174)	(0.00275)	(0.00169)
High attractiveness group	-0.00435	-0.000513	-0.00309
	(0.00327)	(0.00490)	(0.00319)
High IQ group			
Low attractiveness group	-0.0126*	-0.00355	-0.00607
	(0.00738)	(0.00593)	(0.00668)
Intermediate attractiveness group	-0.00267	0.00473	-0.00318
	(0.00361)	(0.00308)	(0.00301)
High attractiveness group	-0.00104	0.00108	0.00839
	(0.00526)	(0.00553)	(0.00609)
Observations	3,281	1,610	3,280
R^2	0.178	0.214	0.239
Controls	Y	Y	Y
High school FE	Y	Y	Y

Standard errors are clustered at high school. *** p<0.01, ** p<0.05, * p<0.1

college, markedly so with regard to the probabilities of withdrawing having enrolled, and of receiving a college degree.

These regressions also display incidental significance for three of the other 24 surname initial rank coefficients. These results may indicate that other behavioral mechanisms may also have been linked to surname initial rank. However, the inconsistency of these effects across the regressions of table 11, compounded by inconsistency across the other tables in this section, suggests that they are likely to have been inconsequential.

Table 12. Initial employment by IQ-attractiveness strata

Explanatory variables	Military service	First employment	
		Income score	Prestige score
Alphabetical rank of surname initial			
Low IQ group			
Low attractiveness group	-0.00729	-0.0109	-0.0387
	(0.00866)	(0.0205)	(2.365)
Intermediate attractiveness group	-0.000708	0.0119	1.225
	(0.00412)	(0.00842)	(1.055)
High attractiveness group	-0.00350	-0.000379	-1.613
	(0.00948)	(0.0187)	(2.136)
Intermediate IQ group			
Low attractiveness group	0.00324	-0.00783	-0.364
	(0.00455)	(0.00859)	(1.143)
Intermediate attractiveness group	0.00618***	-0.0114***	-1.386**
	(0.00211)	(0.00420)	(0.569)
High attractiveness group	0.00653	0.0110	0.739
	(0.00441)	(0.00865)	(1.254)
High IQ group			
Low attractiveness group	0.000646	0.00369	-2.043
	(0.00689)	(0.0166)	(2.747)
Intermediate attractiveness group	0.00753*	-0.00216	-1.475
	(0.00409)	(0.00733)	(1.159)
High attractiveness group	0.00195	0.00144	-0.427
	(0.00651)	(0.0158)	(2.149)
Observations	3,281	3,086	3,087
R^2	0.045	0.221	0.318
Controls	Y	Y	Y
High school FE	Y	Y	Y

Income score is in natural log. Standard errors are clustered at high school. *** p<0.01, ** p<0.05, * p<0.1

Table 12 expands the regressions of table 7 with the interactions between IQ score, attractiveness score and surname initial rank. Here, surname initial rank is associated with only one significant coefficient in strata other than that with intermediate IQ and attractiveness scores.

However, in this strata, the coefficients for surname initial rank in the regressions for military service and the prestige score for first employment are significant, of the same sign and larger in magnitude than the corresponding coefficients for the entire sample in table 7. Moreover, the coefficient for surname initial rank in the first employment income score regression is much larger than the insignificant coefficient for the entire sample in table 7, and statistically significant.

Table 13. Employment in adulthood by IQ-attractiveness strata

Explanatory variables	Employment in 1974		Employment in 1992	
	Earnings	Prestige score	Earnings	Prestige score
Alphabetical rank of surname initial				
Low IQ group				
Low attractiveness group	0.0102	-0.477	-0.00502	-0.318
	(0.00791)	(1.709)	(0.0152)	(2.671)
Intermediate attractiveness group	-0.00379	0.884	-0.00946	0.848
	(0.00362)	(0.826)	(0.00707)	(1.059)
High attractiveness group	0.00949	-0.313	0.0185	-1.632
	(0.0111)	(2.043)	(0.0132)	(2.152)
Intermediate IQ group				
Low attractiveness group	0.00162	0.944	-0.0125**	-0.772
	(0.00445)	(0.918)	(0.00576)	(1.104)
Intermediate attractiveness group	-0.00385**	-0.303	0.00101	-0.0834
	(0.00174)	(0.425)	(0.00345)	(0.502)
High attractiveness group	0.00586	0.767	-0.00437	0.375
	(0.00637)	(0.898)	(0.00752)	(1.251)
High IQ group				
Low attractiveness group	0.0169	2.837	0.0354	-0.987
	(0.0142)	(2.277)	(0.0254)	(1.938)
Intermediate attractiveness group	8.20e-05	0.113	0.00259	-0.544
	(0.00383)	(1.053)	(0.00648)	(0.958)
High attractiveness group	0.00647	-1.020	0.0290	-2.474
	(0.00723)	(1.668)	(0.0224)	(1.876)
Observations	2,694	3,220	2,426	2,863
R^2	0.09	0.272	0.114	0.213
Controls	Y	Y	Y	Y
High school FE	Y	Y	Y	Y

Earnings are in natural logs. Standard errors are clustered at high school. *** p<0.01, ** p<0.05, * p<0.1

Finally, table 13 presents the elaboration of table 8. Only two coefficients for surname initial rank are statistically significant in table 13, one for the subsample with intermediate IQ and intermediate attractiveness scores. The coefficient magnitudes for this variable in this strata in the other three regressions do not differ systematically from those for the entire sample in table 8. Therefore, it appears as though the effects of surname initial rank dissipate with age for this subsample, as well as for the entire sample.

6. Conclusion

The analyses presented here demonstrate that alphabetism has significant and substantial negative effects on men with surname initials ranked further from the beginning of the alphabet in high school, college and first labor market experiences. Those with higher-ranked initials are less likely to be recognized as an outstanding students in high school, less likely to have favorable opinions of their high school experience, less likely to apply to college while in high school, less likely to remain in college if admitted and less likely to earn a college degree. They are also more likely to have military experience and to have first jobs with lower occupational prestige scores.

These effects appear to be driven by the experience of alphabetic orderings, both in assigning opportunities and in conditioning individuals to be receptive to opportunities. Those whose surname initials are ranked further from the beginning of the alphabet are presumably offered fewer opportunities and are less prepared to take advantage of opportunities that are offered under alphabetic assignments.

These effects also appear to be driven by experiences of prominence in other domains. They are inconsequential for those who are distinctive, either through especially low or especially high scores, in cognitive ability or physical attractiveness. They are consistently present only for those who are of intermediate rank in both. For these men, who attract the least attention, the further disregard associated with later placement in the alphabet is especially harmful.

References

Ackerman, Margaret and Simina Brânzei (forthcoming) "The authorship dilemma: Alphabetical or contribution?", Journal of Autonomous Agents and Multi-Agent Systems.

Ang, James, Ansley Chua and Danling Jiang (2010) "Is A better than B? How affect influences the marketing and pricing of financial securities", Financial Analysts Journal, Vol. 66, No. 6, November/December, 40-54.

Angelucci, Manuela, Giacomo De Giorgi, Marcos A. Rangel, and Imran Rasul (2010) "Family networks and school enrolment: Evidence from a randomized social experiment", Journal of Public Economics, Vol. 94, Nos. 3-4, April, 197-221.

Arai, Mahmood and Peter Skogman Thoursie (2009) "Renouncing personal names: An empirical examination of surname change and earnings", Journal of Labor Economics, Vol. 27, No. 1, 127-147.

Arbatskaya, Maria (2007) "Ordered search", Rand Journal of Economics, Vol. 38, No. 1, Spring, 119-126.

Arsenault, Clément and Vincent Larivière (2015) "Is paper uncitedness a function of the alphabet?", ISSI (International Society for Scientometrics and Infometrics), 286-287.

Aura, Saku and Gregory D. Hess (2010) "What's in a name?", Economic Inquiry, Vol. 48, No. 1, January, 214–227.

Baert S, Cockx B, Gheyle N, et al. (2015) "Is there less discrimination in occupations where recruitment is difficult?", Industrial and Labor Relations Review, Vol. 68, No. 3, May, 467–500.

Berger, Jonah (2016) "Does presentation order impact choice after delay?", Topics in Cognitive Science, Vol. 8, 670-684.

Bertrand, Marianne and Sendhil Mullainathan (2004) "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination", The American Economic Review, Vol. 94, No. 4, September, 991-1013.

Biavaschi, Costanza, Corrado Giulietti and Zahra Siddique (forthcoming) "The economic payoff of name Americanization", Journal of Labor Economics.

Carlsson, Magnus and Dan-Olof Rooth (2007) "Evidence of ethnic discrimination in the Swedish labor market using experimental data", Labour Economics, Vol. 14, 716-729.

Carney, Dana R. and Mahzarin R. Banaji (2012) "First is best", PLoS ONE, Vol. 7, No. 6, June, e35088. doi:10.1371/journal.pone.0035088

Carpusor, Adrian G. and William E. Loges (2006) "Rental discrimination and ethnicity in names", Journal of Applied Social Psychology, Vol. 36, No. 4, 934-952.

Clark, Gregory, and Neil Cummins (2015) "Intergenerational Wealth Mobility in England, 1858–2012: Surnames and Social Mobility", The Economic Journal, Vol. 125, No. 582, February, 61–85.

Collado, M. Dolores, Inacio Ortuño Ortín and Andrés Romeu (2008) "Surnames and social status in Spain", Investigaciones Económicas, Vol. 32, No. 3, 259-287.

Derous, Eva, Roland Pepermans and Ann Marie Ryan (forthcoming) "Ethnic discrimination during resume screening: Interactive effects of applicants' ethnic salience with job context", Human Relations, 1-23.

Dietrich, J. P. (2008) "The importance of being first: Position dependent citation rates on arXiv:astro-ph", Publications of the Astronomical Society of the Pacific, Vol. 120, 224-228.

Efthyvoulou, Georgios (2008) "Alphabet economics: The link between names and reputation", The Journal of Socio-Economics, Vol. 37, 1266-1285.

Einav, Liran and Leeat Yariv (2006) "What's in a surname? The effects of surname initial on academic success", Journal of Economic Perspectives, Vol. 20, No. 1, Winter, 175-188.

Feenberg, Daniel, Ina Ganguli, Patrick Gaulé and Jonathan Gruber (forthcoming) "It's good to be first: Order bias in reading and citing NBER working papers", Review of Economics and Statistics.

Fershtman, Chaim, and Uri Gneezy (2001) "Discrimination in a segmented society: An experimental approach", The Quarterly Journal of Economics, Vol. 116, No. 1, February, 351–77.

Figlio, David N. (2005) Names, Expectations and the Black-White Test Score Gap, National Bureau of Economic Research Working Paper 11195.

Foster, Gigi (2008) "Names will never hurt me: Racially distinct names and identity in the undergraduate classroom", Social Science Research, Vol. 37, 934-952.

Frandsen, Tove Faber and Jeppe Nicolaisen (2010) "What is in a name? Credit assignment practices in different disciplines", Journal of Informetrics, Vol. 4, 508-517.

Fryer, Roland G. Jr. and Steven D. Levitt (2004) "The causes and consequences of distinctively black names", The Quarterly Journal of Economics, Vol. 119, No. 3, August, 767-805.

Ginsburgh, Victor A. and Jan C. van Ours (2003) "Expert opinion and compensation: Evidence from a musical competition", The American Economic Review, Vol. 93, No. 1, March, 289-296.

Gottfredson, Linda S. (2006) "Logical fallacies used to dismiss the evidence on intelligence testing", Chapter 1 in Phelps, Richard F. ed., Correcting Fallacies About Educational and Psychological Testing, American Psychological Association, Washington, D.C.

Güell, Maia, José V. Rodríguez Mora, and Chris I. Telmer (2015) "The Informational Content of Surnames, the Evolution of Intergenerational Mobility, and Assortative Mating", Review of Economic Studies, Vol. 82, No. 2, April, 693–735.

Hamermesh, Daniel S. and Gerard A. Pfann (2012) "Reputation and earnings: The roles of quality and quantity in academe", Economic Inquiry, Vol. 50, No. 1, January, 1-16.

Haque, Asif-ul and Paul Ginsparg (2009) "Positional effects on citation and readership in arXiv", Journal of the American Society for Information Science and Technology, Vol. 60, No. 11, 2203–2218.

Hartzmark, Samuel (2015) "The worst, the best, ignoring all the rest: The rank effect and trading behavior", Review of Financial Studies, Vol. 28, No. 4, 1024-1059.

Herd, Pamela, Deborah Carr and Carol Roan (2014) "Cohort profile: Wisconsin longitudinal study (WLS)", International Journal of Epidemiology, Vol. 43, 34-41.

Ho, Daniel E. and Kosuke Imai (2008) "Estimating causal effects of ballot order from a randomized natural experiment: The California alphabet lottery, 1978-2002", Public Opinion Quarterly, Vol. 72, No. 2, Summer, 216-240.

Itzkowitz, Jennifer, Jesse Itzkowitz and Scott Rothbort (2016) "ABCs of trading: Behavioral biases affect stock turnover and value", Review of Finance, 663-692.

Jacobs, Heiko and Alexander Hillert (2016) "Alphabetic bias, investor recognition, and trading behavior", Review of Finance, Vol. 20, No. 2, 693-723.

Jurajda, Štěpán and Dejan Kovač (2016) What's a Name in a War, IZA Discussion Paper No. 10331.

Jurajda, Štěpán and Daniel Münich (2010) "Admission to selective schools, alphabetically", Economics of Education Review, Vol. 29, 1100-1109.

Kadel, Annke and Andreas Walter (2015) "Do scholars in economics and finance react to alphabetical discrimination?", Finance Research Letters, Vol. 14, 64-68.

Kohn, Meir G. and Steven Shavell (1974) "The theory of search", Journal of Economic Theory, Vol. 9, 93-123.

Kumar, Alok, Alexandra Niessen-Ruenzi and Oliver G. Spalt (2015) "What is in a Name? Mutual Fund Flows When Managers Have Foreign-Sounding Names", Review of Financial Studies, Vol. 28, No. 8, August, 2281-2321.

Maciejovsky, Boris, David V. Budescu, and Dan Ariely (2009) "The researcher as a consumer of scientific publications: How do name-ordering conventions affect inferences about contribution credits?", Marketing Science, Vol. 28, No. 3, May-June, 589-598.

McDevitt, Ryan C. (2014) "'A' business by any other name: Firm name choice as a signal of firm quality", Journal of Political Economy, Vol. 122, No. 4, 909-944.

Merritt, Deborah Jones (1999) "Calling professor AAA: How to visit at the school of your choice", Journal of Legal Education, Vol. 49, No. 4, December, 557–563.

Olivetti, Claudia and M. Daniele Paserman (2015) "In the name of the son (and the daughter): Intergenerational mobility in the United States, 1850-1940", The American Economic Review, Vol. 105, no. 8, August, 2695-2724.

Ong, David, Ho Fai Chan, Benno Torgler and Yu (Alan) Yang (2016) Endogenous Selection Into Single and Coauthorships by Surname Initials in Economics and Management, working paper.

Pena, Pablo A. (2013) Surname Frequency and Lifespan, working paper.

van Praag, C. Mirjam and Bernard M.S. van Praag (2008) "The benefits of being professor A (rather than Z)", Economica, Vol. 75, 782-796.

Richardson, Michael L. (2008) "Alphabetic bias in the selection of reviewers for the American Journal of Roentgenology", American Journal of Roentgenology, Vol. 191, No. 6 Supplement, December, 213–216.

Rosen, Harvey S. and Jonathan Meer (2011) "The ABCs of charitable solicitation", Journal of Public Economics, Vol. 95, 363-371.

Rubinstein, Yona and Dror Brenner (2014) "Pride and prejudice: Using ethnic-sounding names and inter-ethnic marriages to identify labor market discrimination", The Review of Economic Studies, Vol. 81, No. 1, January, 389-425.

Waltman, Ludo (2012) "An empirical analysis of the use of alphabetical authorship in scientific publishing", Journal of Informetrics, Vol. 6, 700-711.

Widner, Daniel and Stephen Chicoine (2011) "It's all in the name: Employment discrimination against Arab Americans", Sociological Forum, Vol. 26, No. 4, 806–823.

Zax, Jeffrey S. and Daniel I. Rees (2002) "IQ, academic performance, environment and earnings", The Review of Economics and Statistics, Vol. 84, No. 4, November, 600-616.