When Does Beauty Pay? A Large Scale Image Based Appearance Analysis on Career Transitions

Nikhil Malik, Param Vir Singh, Dokyun Lee, Kannan Srinivasan

Abstract

We investigate the dynamic effects of beauty over an individual's career. Using a fine grained longitudinal sample on career milestones and educational background of 7436 individuals selected from an online professional social network and employing computer vision methods for rating attractiveness of individuals, we find that men enjoy a beauty premium early in their career which disappears later in the career. In contrast, women do not receive a beauty premium early in their career. Rather they receive a beauty premium later in their career. We show these effects through a survival analysis where attractive men are found to progress faster in their career early on and women are found to progress faster in their later career in comparison to their unattractive counterparts respectively. We find that the overall beauty premium is greater for women. These results are robust to a number of control variables and individual unobserved heterogeneity. We provide theoretical reasoning that rationalizes these findings.

1. Introduction

Discrimination against individuals based on beauty has generated immense interest both among researchers and practitioners (Mobius and Rosenblat 2005, Hamermesh 2011, Cohen 2017, Ruffle and Shtudiner 2014). Any discrimination in the labor market based on attractiveness would make the hiring and promotion processes inefficient. As a result, organizations have become cognizant of unconscious discrimination for not only beauty but also against females (Rivers and Barnett 2016, Chicago tribune) and minorities (Vega 2015, money.cnn). In fact, all employees at Google go through an unconscious bias training that is aimed at minimizing the subconscious biases (positive or negative) that affect behavior (Feloni 2016, Business Insider).

Existing literature has often studied beauty bias in an experimental setup (Mobius and Rosenblat 2005, Andreoni and Petrie 2006). Individuals are asked to play hypothetical roles such as jury member (kulka and kessler 1978) or manager (Watkins and Jones 2015) choosing between likely convicts or able workers respectively. These studies highlight bias when an evaluator has limited information to differentiate candidates on true quality. Other studies (Hamermesh and Biddle 1994, Zebrowitz and Donald 1991) that look at archival data, show a beauty premium over a significant career length without revealing whether bias consistently effects an individual over his or her entire career or not. At what point beauty bias plays a role is extremely important to understand the mechanism at play. These long-term studies typically find positive premium for both men and women overall. Once again ignoring the time dynamics likely masks different mechanisms at play for the two genders.

In this study, we investigate the dynamic effects of beauty bias in the labor market to identify different factors that result in beauty premium for men and women. Using detailed panel data on career milestones and education background of 7436 individuals selected from a professional social network, we show that the beauty bias assists attractive men (2 s.d. above mean) move up 22.6% faster early in the career when employers have little information about the true quality of the candidates. However early in the career attractive women may not be able to extract the same advantage, they may be perceived to not be serious participants in the job market or face jealous reactions from their competitors. But if analyzed over mid-career and later career jobs they would appear to outperform their plain looking counterparts. We find attractive women (2 s.d. above mean) are 33.6% more likely to move up in any given period in this phase. Studies have shown mixed results on whether beauty is penalizing (Ruffle and Shtudiner 2014) or rewarding (Hamermesh and Biddle 1994) for women. We assert that the confusion likely arises from focus on different career phases in these studies.

The size of our study (7436 primary and 92540 auxiliary profiles) provides a unique external validity on impact of looks on professional careers. However, breadth of our study presented two key challenges: we need to rate attractiveness of large number of subjects and establish a measure of success which is comparable across industries and career stages. Using computer vision techniques, we build a supervised machine learning model that is trained to mimic attractiveness scoring done by human raters. This model is invariant to temporary characteristics such as: clothing, hair style and expressions. Secondly, we extend upon traditional labor economics methods (Baker et al 1994 and Gayle et al 2012) that establish preference order between jobs utilizing observed pairwise job switches. This approach works well when the hierarchy is well defined, allowing quantification of success. Our unstructured data is highly sparse, including over 100,000 unique employers and 100,000 unique titles. Therefore we improve upon the basic idea by building a dense representation of jobs and running a variant of Page Rank algorithm to calculate a measure of job's rank. The resulting job ranks are highly intuitive and can be used by other studies which investigate career transitions from large scale unstructured data.

Our paper makes three main contributions. First, we show that the while the beauty bias exists, it makes an impact for men and women at different career stages. Men are affected early in their career whereas women are affected later in their career. Our findings help uncover the mechanism behind beauty bias. Second, our findings are based on one of the largest samples of longitudinal archival data employed to study beauty bias which provides external validity to the findings of prior experimental small sample studies. Third, our computer vision based machine learning method for rating attractiveness and our method for ranking jobs has wide applicability in any labor economics study on appearances.

2. Literature Review and Theoretical Development

Existing literature establishes that attractive looks are associated with positive perception of social skills, mental health and personality (Dion et al 1972). These qualities align well with requirements of labor market. Studies on politicians (Olivola and Todorov 2010), lawyers (Biddle and Hamermesh

1998), doctors and retail managers (Watkins and Jones 2015) therefore find an impact of beauty on success. However, several of these studies have been conducted in experimental settings where individuals making the decision have had little interaction with the candidate, highlighting situations where an individual has limited information to differentiate subjects. In such a scenario, people making the decision construe a candidate's true quality from superficial cues from looks. Studies that use archival data (Hamermesh and Biddle 1994) have also established the presence of a beauty bias however, they largely ignore time dynamics.

We attempt to track beauty rewards extracted over time. Early career decisions resemble a scenario where the employer has limited information therefore may unwittingly absorb perceptions from appearance cues. Mobius and Rosenblat (2005) show in an experimental labor market, employers consider attractive workers as more able even with equal performance measures. Similar outcomes are shown in public goods experimental setup by Andreoni and Petrie (2006). However, unlike an experiment, in the real job market as employers get time to learn and are able to differentiate true qualities of job seekers any beauty reward should vanish. While we should still see an overall premium of beauty of an individual's lifetime, most of it should be extracted very early on.

While this rationale should hold for men, a more compounded affect has been argued for attractive women. Buss and Heselton (2005) suggest that women are significantly threatened by other attractive women, and therefore may respond negatively to attractive female job seekers. Sagarin et al (2003) show that this feeling of jealousy could impact women more than men. Ruffle and Shtudiner (2014) propose the "dumb-blonde" stereotype as yet another potential reason why attractive women may be penalized. This hypothesis roots from a misplaced belief that young attractive women have easier options available and therefore unfit for serious tasks. Heilman and Saruwatari (1979) find attractive women have a better shot at clerical jobs compared to plain looking counterparts, but a worse chance at more desirable important positions. This and other negative penalties of envy and jealousy are more likely early in the career for women. Ruffle and Shtudiner (2014) show that young single women are most likely be associated with jealous response when responsible for evaluating job seeking attractive women. Therefore, older women may escape from such jealous reactions.

Unlike the social psychology literature which specifically points out settings where attractive women are penalized, majority of labor economics studies assert a lifetime positive beauty premium for both men and women. Beside the perception of positive personality traits, Biddle and Hamermesh (1998) suggest a base desire to be around attractive people as a possible reason for this premium. Valuation of attractive women for their presentability and greater instances of sexual harassment for women (Estrada and Harbke, 2008, Arani and Mobarakeh 2011) suggest another, perhaps perverse, reward for attractive women.

In tracking the impact of beauty over time, our work undertakes resolution of these conflicting findings. Early in the career these two counteracting effects are likely to diminish any significant reward or penalty for women. Once the first effect of envy, jealousy and "dumb-blonde" stereotype goes away at age 30 and beyond, the remaining effect of base desires should result in a beauty premium.

3 Data Description

We collect 7436 user profiles from one of the largest US professional social network. This online platform is primarily used for professional networking and by employers to post jobs and job seekers to post their CV's. A typical user profile on this platform consists of curriculum vitae describing work experience, education and a personal photo among other details. Members on this platform typically make connections with each other which may represent real world associations. While members self-report their professional information, the social nature of the platform ensures veracity of the data.

The platform allows us to select individuals based on university or schools attended and companies worked for. For our main analysis, we select 7436 individuals who graduated from a Top

Table 1: List of variables for each profile.

Variables	Sample Values
Attractiveness Score	1,2,3,4,5,6,7
Gender	Male (72.7%), Female (27.3%)
Ethnicity Category	East Asian (3.4%), West-Asian and North african (19.4%), Others (77.2%)
University Name	Harvard University, Stanford University etc.
Degree Name	M.S., M.F.E, B.A., J.D. etc
Undergraduate Area	Science (66.6%), Arts (17.0%), Business (6.0%), Others (10.4%).
Undergraduate Ranking	Top 10 (19.2%), Top 50 (22.5%), Post 50 (58.3%)
Title	Software Developer, Consultant, Intern, CEO.
Employer Size	Very Small, Small, Medium, Large, Very Large.
Employer Industry	Computers (31.8%), Management Consulting (23.2%), Finance (18.2%) etc.
Job	"Accountant at Large Finance firm", "Founder at Very Large internet firm" etc.
Job Type	Technical, Management, Finance, Law enforcement etc.

Some of the variables like attractiveness and university ranking are not directly available instead calculated, as described in subsequent sections.

50 MBA programs in the 10-year period between 2004 and 2014. This sample selection criteria serves two purposes. First, the criteria ensure that the individuals in the sample are comparable because of their background (top 50 MBA). They would most likely compete for similar kind of jobs and should see similar career progressions. Second, individual looks could change over a long time. Since, we do not have access to profile pictures of individuals over a long-time period, we would not know how attractiveness of an individual changed over time. And hence a longer sample could potentially lead to biased results. Individual looks change minimally within a 10-year period and hence the attractiveness scores will be more accurate corresponding to the career milestones. We collect complete profile information for these individuals.

Table1 list some of the key profile details captured. Education background and a series of professional roles held by the individual over their career are of key importance. We typically find 2-5 career steps and 2 degrees (e.g. high school, undergraduate, masters, MBA, PhD etc.) listed by the users.

Figure 1a (left): Shows a histogram of number of career milestones posted by individuals. **Figure 1b** (right) shows number of career milestones in top 5 industry categories broken down by gender. **Figure 1c** (bottom) shows attractiveness ratings for profile pictures broken by gender. Women are fewer in number but rated higher on attractiveness on an average.

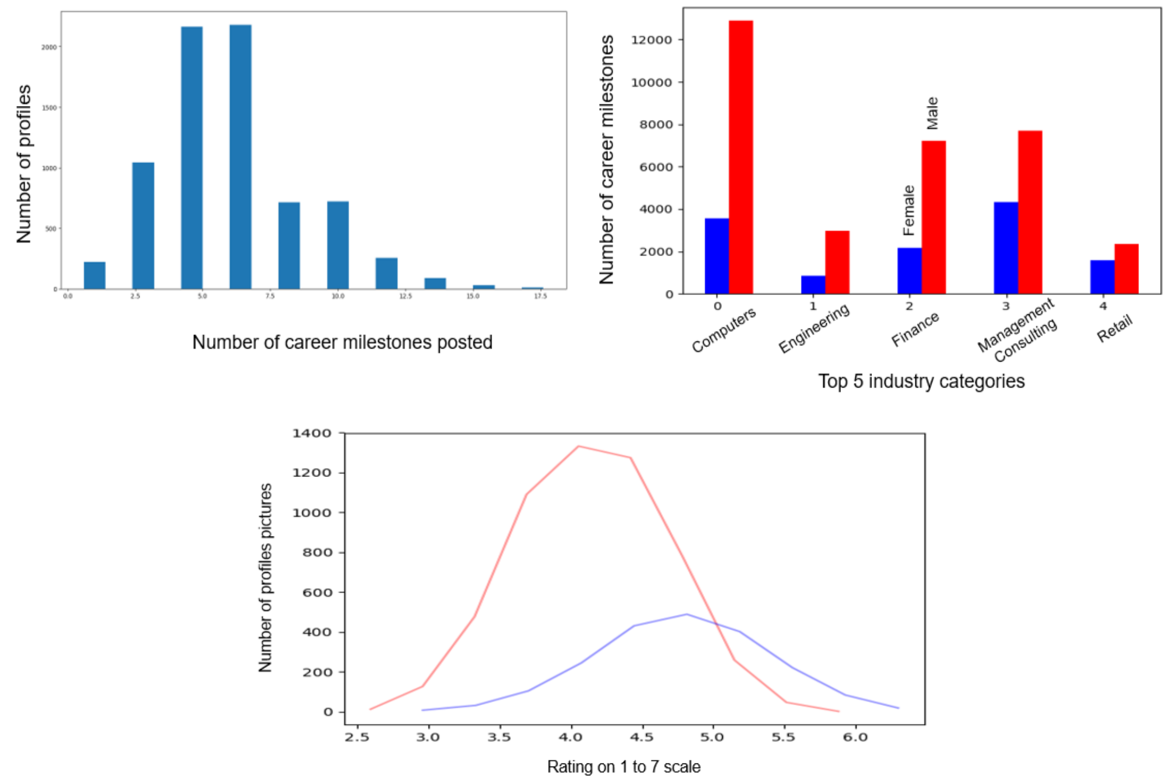


Table 1 also reports a few additional fields we calculate for the purpose of our analysis. These include gender and ethnicity groups (Ambekar et al 2009) derived from individuals name, area of undergraduate studies derived from name of the degree posted on the profile and ranking of university attended obtained from usnews.com. Our dependent variable of interest is career success. We borrow a process of hierarchy construction of jobs and titles across industries from the literature in labor economics (Baker et al 1994 and Gayle et al 2012). Specifically, we use data on pairwise job switches to construct this hierarchy. For example, at an intuitive level if more people switch from job A to job B and few people switch from job B to job A then job B is higher in hierarchy than job A. The accuracy of this method depends on having a dataset that has a large number of switches over time. As a result, we augment our data with an additional sample of 92,450 user profiles that have attended a top 50 university or performed the roles of: consultant at top 20 management consulting firms, software developer at top 20 technology firms and sales at top 20 Investment Banking firms. The information on these additional users is only used for constructing job hierarchies and not for our main analysis on beauty bias. We discuss more details on job hierarchy construction in next section.

4 Methods

The profile images and career data we have collected are in a relatively unstructured form. In this section, we describe how we go from a profile picture to tagging an individual with attractiveness ratings. Second, we detail our approach for construction of job hierarchy used for quantifying success.

4.1 Attractiveness Ratings

Our source platform provides us a single image per person. Given the professional nature of the profiles, people often post clear head shots. We detect and remove profiles from our sample for which the pictures either do not capture a face or capture more than one face. Assuming at least 10 human raters per images it would require 74,360 ratings (7436*10). As a solution to manage our costs, we get only a small sample of images rated by individuals and use it as training data. We then train a machine learning model to learn relationship between a low-level image feature and attractiveness label on our training data. As a result, our machine learning model learns to mimic evaluating attractiveness like a human rater. We explain the different steps used in building and testing out machine learning classifier next.

4.1.1 Training Data

We set up a small-scale experiment to get raters to judge a random set of 659 out of our full sample of 99,976 (7436 primary and 92540 auxiliary) pictures on attractiveness. These ratings would act as training labels to our machine learning model. This experiment was executed on AMT - Amazon Mechanical Turk (Paolacci et al 2010). The raters were selected to participate if they qualify three criteria: 1) geographical location in United States, 2) completed at least 500 human intelligence tasks (HITs) in past on the AMT platform and 3) have received 95% or higher approvals on all their completed HITs. These conditions were applied to ensure the raters understand our requirements in English, they are experienced in using the AMT platform and have showcased good quality of work in past. Additionally, limiting the raters to United States is likely to help in mimicking how a subject's appearance would be judged in a typical US professional work environment. Attractiveness is a subjective measurement and multiple raters are unlikely to exactly agree in their ratings. We calculate a standardized Cronbach alpha (Santos 1999) value of 0.73, which is typical psychometric standard used for inter rater consensus between the raters. The relatively high value (0.7 typically being an acceptable standard) gives us confidence that using an average would capture at least a coarse variance in attractiveness of profile pictures.

4.1.2 Image Feature Extraction

We begin by pre-processing images this step involves removing extraneous details such as picture quality, picture background, clothing, hairstyle and facial expressions from images. This correction is necessary to ensure any machine-driven face classification or rating task picks up actual appearance differences instead of superficial differences in how the photograph was taken. We first use dlib library (dlib.net/python) to get a bounding box for the face on profile picture. The color histogram and overall pixel intensity values are normalized across images. The images raise further challenge in that the pose in the picture is not standardized, i.e. the head poses may differ across profiles. For our machine learning algorithm to work better, we need to adjust the head pose. We investigate two approaches: Localized PCA (Ahonen et al 2006) and 2D affine transformation (Zhang and Gao 2009) to align face images that differ due to head rotation in three dimensions i.e. roll, tilt and yaw (figure 2). We find the latter method to perform better.

Figure 2: Correction of rotated head.
(A) shows a simple 2D translation which accounts for tilt. (B) shows a 2D affine transformation which accounts for roll, tilt and yaw.

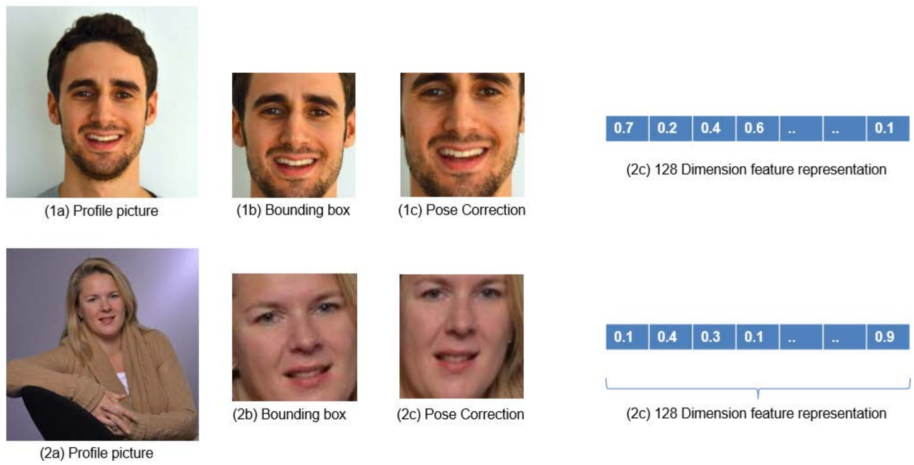
Once front facing comparable face images are obtained, we compress a typical 400x400x3 (RGB) image into a handful of key features that capture a majority of differences between appearances. We investigate three different approaches from literature on machine rated facial beauty: Golden ratios and neoclassical canons (Schmid et al 2008), Eigen faces (Turk and Pentland 1991) and deep neural networks (Lawrence et al 1997). Eventually in testing phase the deep neural network method easily outperforms the other two methods. We reuse Deep Neural network implementation by the Open Face project (Amos et al 2016). This architecture, while trained for face recognition task, provides a 128-dimensional intermediate layer. This layer represents a low dimensional embedding of any face image. Once we have the low dimensional features generated using the deep neural network implementation, we verify that they have predictive power on attractiveness. Figure 3 lays out a summary of various transformations performed on profile pictures to arrive at a set of 128 features for every face.

4.1.3 Model Taring and Testing

Our training data consists of images labeled as attractive on a 1 to 7 scale where 1 represents lowest value of attractiveness and 7 represents highest value of attractiveness. Each image is labeled by 5 coders and we use the average of the five ratings for an image as its true rating. Because the resultant ratings are continuous, we train a Support Vector regression model that learns relationship between 128 image features and the attractiveness label. The model achieves high prediction accuracy with a

Figure 3: Two example of profile pictures and subsequent transformation to retrieve 128-dimension feature representation.

Notice the removal of pose, hairstyle, background and clothing.



Root mean square prediction error of 0.92 on a 5-fold cross validation. Figure 4 reports the prediction range for our model with varying attractiveness levels of actual ratings. The histogram of ratings shows that a large number of faces are rated in the 3.5 - 5.5 region, as a result our model also does a relatively good job of mimicking user ratings in this region.

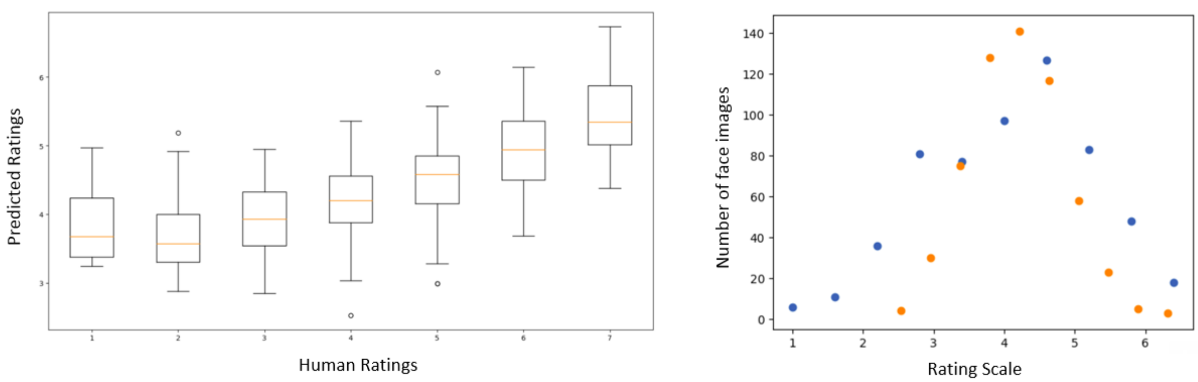


Figure 4a (left): Variance of predicted ratings against human ratings. Figure 4b (right): Shows histogram of human ratings in dark blue and histogram of predicted ratings in light red. Note the small number of observations available at the extremes.

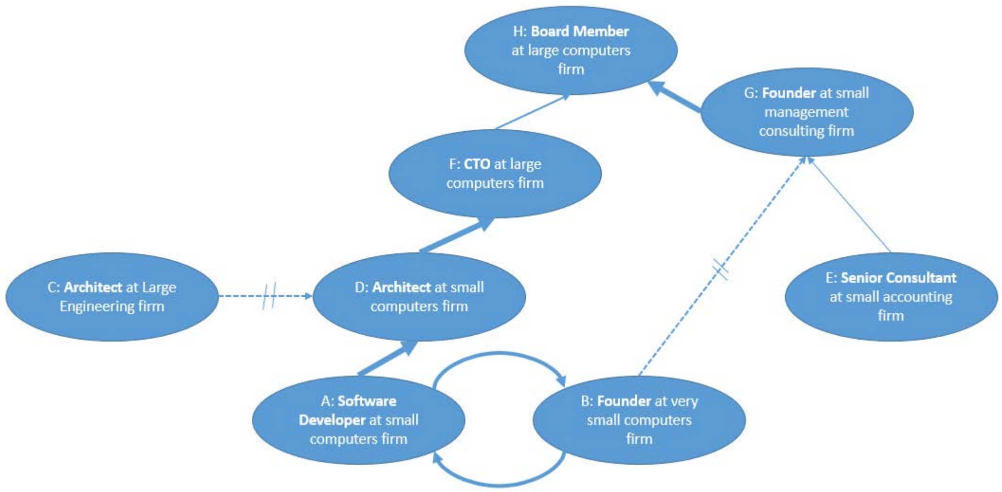
4.2 Job Rank

Any study on success in professional roles presupposes a preference order of desired roles. We utilize approach suggested by Baker et al (1994) and (Gayle, Golan and Miller 2012) which relies on observed job switches to establish a preference order. This formulation presumes that a large number of transitions from title A to title B relative to the number of backward transitions typically imply that title B is more desirable and therefore ranked higher. This approach works well when there is a clear hierarchy, which is the case for late career positions (CEO, CFO, principal consultant, board member) in firms of comparable size. However, majority of milestones in our data corresponds to early and mid-career roles in significantly dissimilar firms and industries, therefore fewer pairwise transitions are observed.

We make three significant changes to extend the original approach. We observe more than 100,000 distinct titles in our data because of the unstructured nature of the profiles. A majority of these titles appear in 5 or fewer profiles, therefore not affording enough information on pairwise transitions for us to accurately infer their preference order. Instead we can more effectively build a preference ranking for the 1000 most frequent titles and then approximate the ranking for the remaining titles based on text similarity with the 1000 ranked frequent titles. As an example, "principal consultant" and "software developer" are two common titles which can be used to infer ranking for "principal consultant of energy-oil industry" and "Agile software developer". Therefore, as a first transformation, we convert title strings into a tf-idf (Mihalcia 2006) representation. We chose 2000 as the size of the vocabulary, therefore describing every title as 2000-dimensional vector. This choice of vocabulary size keeps words informative of rank such as "vice", "associate", "chief", while eliminating words such as: "rockstar", "energy-oil" etc. Once we have a ranking of the 1000 frequent titles, this tf-idf representation allows us to calculate distance between any two titles. We use weighted k-NN (Cover et al 1967, k = 10) to approximate ranking for the less frequent titles.

$$rank_{jobA} = \sum_{10-NN} cosine(tfidf_{jobA}, tfidf_{jobN}) * rank_{jobN} \qquad (1)$$

Figure 5: A small set of jobs with thickness and direction of arrows representing the frequency and direction of typical job switches. The dashed line represents transitions that almost never occur even though the titles are similar.



Secondly, we define a "job" as a combination of Title, employer size and employer industry. This ensures that examples job C vs job D and job B vs job G are not grouped together simply because the title remains the same in otherwise very dissimilar roles. Once we have a set of "jobs" that occur relatively frequently in our dataset, we count all directed transitions between any two pair of jobs. The matrix M represents these transitions. We initialize all jobs to have equal rank (r_o), and then run our variant of PageRank algorithm (reference) to arrive at a differentiated rank score for each job (r). The Page Rank algorithm is traditionally used to determine importance of web pages based on in-links and out-links. In our case, jobs are equivalent to web pages and links are equivalent to job transitions. A top position is not likely to have the largest number of incoming transitions instead it gets incoming transitions from relatively senior positions (example: CFO to CEO, Principal to CEO). This algorithm iteratively propagates rank along the directed edges, therefore eventually transferring weight from a junior position (intern, analyst) to top positions (CEO, board member). The rank (r) corresponds to the likelihood of ending up in a given job if an individual spends infinite time in the job market, switching between the titles with probability of each step drawn from the transitions matrix (M).

r_0: initial rank = 0.0001, r_f: final rank, M: transition matrix, E: teleportation matrix, x = 0.9

$$r_f = (x*M + (1-x)*E)r_0$$

Table 2: Initial rank assigned to each job and final rank obtained after convergence of page rank formula.

Job	Log (initial rank)	Log (final rank)
H: **Board Member** at large computers firm	-4	-2.5
F: **CTO** at large computers firm	-4	-2.9
G: **Founder** at small management consulting firm	-4	-2.7
A: **Software Developer** at large computers firm	-4	-5.6
B: **Founder** at very small computers firm	-4	-5.9

4.3 Job Type

How fast an individual climbs the hierarchy we have developed may inherently differ in career verticals. As an example person A moves across titles such as “Researcher”, “Assistant Professor”, “Professor” while person B follows: “Intern”, “Analyst”, “Senior consultant”, “Chief principal consultant”. A and B even if considered equally successful may be incomparable in terms of how fast they progress at different phases of their career. We therefore construct groups of jobs i.e. job type using an unsupervised hierarchical clustering. This clustering divides career verticals if relatively few individuals are observed to switch between those verticals. The six clusters obtained are tagged manually purely for ease of reference as shown in figure 6.

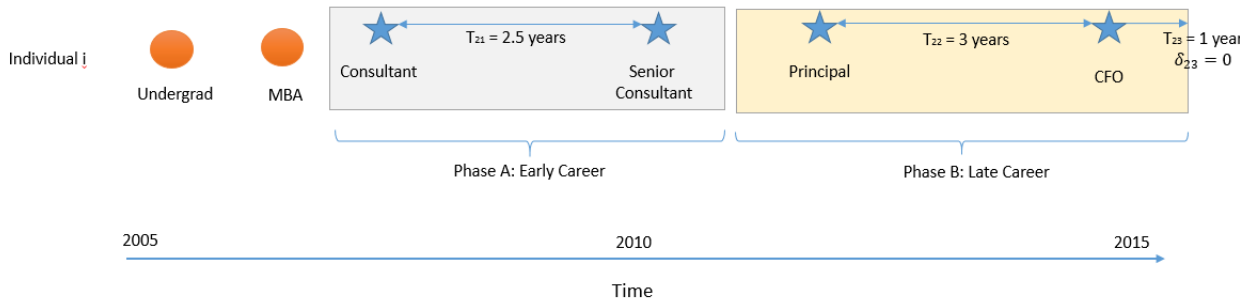
Figure 6: Two sample cluster of job titles obtained.

5. Econometric Model

Our objective is to estimate if attractiveness has a significant impact on an individual's success in different career phases. Figure 7 shows two individuals who complete their educations and attain career milestones at different times. While they have different paths, the first individual reaches a comparable position in a much shorter time overall. We would like to identify if attractiveness played a role in speeding up their career jumps. We perform survival analysis to model expected duration of time it takes an individual to transition to a higher ranked job. Among individuals of same ability, starting at the same job we would expect an individual that gets preferential treatment because of their good looks to take shorter time to make a career jump to a more desirable job.

Unlike a linear regression, survival analysis provides three critical advantages. Firstly, survival time in our case is a positive number. This restriction is not automatically implied in linear regression. Secondly, our observation are censored in scenarios where an individual hasn't left their current job, as shown in figure 7. We can either choose to entirely ignore current job milestone observations or assume that the individual transitions out of their last job in the year we took snapshot of their profile data. Both choices potentially result in a biased estimation. Survival analysis instead allows us to model likelihood of not observing a transition out of a given job. Finally, the ordinary linear regression enforces a normal distribution on the survival time. Instead survival analysis allows parametric methods (Weibull, exponential or lognormal) or semi-parametric methods (Cox proportional hazard) that allow greater flexibility in choice of underlying baseline survival function.

Figure 7: Sample of individual with different education and career timelines.



We select Weibull distribution as the baseline hazard function ($h_0(t)$) parameterized by λ and p as shown in eq 2. This baseline distribution describes the inherent behavior of job switches assuming a homogeneous population. The corresponding survival function ($S_0(t)$) at t is the probability that an individual will spend at least t years in a given job before moving up. Figure 7 shows an example of multiple survival times for an individual across different jobs over their career. Individual i who reports k distinct career milestones will have k corresponding survival time observations. The last of those observation is censored i.e. the duration after which the individual jumps off their current job is unknown. We use δ_{ik} as a binary indictor of whether or not the survival time observation for the individual is censored or not. Assuming a homogeneous population, the likelihood of observed survival times would be calculated as per eq 3. However, we need to adjust the survival time for population characteristics such as education, job sector, size of their jump, attractiveness and individual level heterogeneity. Equation 4 describes inclusion of these covariates similar to standard regression model, differentiated only by inclusion of censored observations and linear relationship of covariates with log of survival time. T_{0ik} is drawn from our choice of Weibull distribution. The quantity $\exp\{x_{ik}*B\}$ alters the baseline survival time based on differentiating characteristic of a specific individual. Any characteristic or covariate in x with positive coefficient indicates contribution to a larger survival time. Such a characteristic can be interpreted as hurting the individual's ability to move up to desirable roles quickly.

$$S_0(t) = \exp\{-(\lambda t)^p\}. \qquad h_0(t) = p\lambda(\lambda t)^{p-1} \tag{2}$$

$$L = \prod_{ik} L_{ik} = \prod_{ik} f(t_{ik})^{\delta_{ik}} * \left(1 - F(t_{ik})\right)^{1-\delta_{ik}} = \prod_{ik} h(t_{ik})^{\delta_{ik}} * S(t_{ik}) \tag{4}$$

$$T_{ik} = \exp\{x'_{ik} * \beta\} * T_{0ik} \tag{4}$$

We also need to account for unobserved differences between individuals. As an example, two individuals matched on ability due to similar university education may still differ in their ability in the job market. We incorporate an unobserved individual specific random effect drawn from a Gaussian distribution. The likelihood of the observed transitions would now be calculated similar to equation 3, with an adjustment to the survival function as described above.

Finally, as shown in Figure 7 we split the observed transitions of an individual between two phases. Phase A includes job switches made within 3 years of graduation from the MBA program. While Phase B includes remaining transitions after 3 years of graduation. This split allows us to track and compare rewards from attractiveness early and late in the career. We chose 3 years as the cutoff for early vs late career as 3 years is a sufficient time to learn about an individual's ability. We also tried a few other cutoff points and found that the beauty bias completely shifts after year 3. Equation 5 represent the individual observed and unobserved characteristics that we attempt to estimate. We use ranking of undergraduate and the MBA program as well as area of undergraduate studies as a substitute for individual's ability and training received. An individual's inherent ability is also controlled by individual specific random effects. We run separate analysis for genders while controlling for ethnicity to account for any demographic differences. Job rank decile indicates the rank of the job the individual is jumping from and allows us to control for inherently different switching behavior in senior positions. Similarly, job type dummies account for differences between career verticals. Each parameter is estimated for four scenarios i.e. early career male, late career male, early career female and late career female.

$$x'_{ik} * \beta = \beta_0 * attractiveScore_i + \beta_1 * ethinicity_i + \beta_2 * jobRank_{ik} + \beta_3 * jobType_{ik} + \beta_4 * job\ Transition\ Jump\ Size_{ik} + \beta_5 * undergraduate\ Area_i + \beta_6 * undergraduate\ University\ Ranking_i + \beta_7 * mba\ Ranking_i + \varepsilon_i \quad (5)$$

$$\varepsilon_i \sim N(0, \sigma^2) \quad (6)$$

6. Results

We have 7436 individuals in our dataset who have completed their MBA in a period of 10 years between 2004 and 2014. We eliminate any profiles without a picture, including only those profiles that were collected with a search on top 50 business school name, which allows us to control for skills and ability of individuals in our dataset. As described in the previous section we estimate by maximizing the log likelihood of observed job switches in four situations split by gender and career phase.

Table 3(a) and 3(b): Distribution of continuous and nominal covariates in the four independent survival regression models

Model	**Covariate**	**Mean**	**Variance**
Early career Male	Job Rank	-7.95	0.47
Early career Male	Job Rank jump	0.069	0.66
Early career Male	Attractiveness Score	4.33	0.27
Late career Male	Job Rank	-7.81	0.38
Late career Male	Job Rank jump	0.052	0.47
Late career Male	Attractiveness Score	4.31	0.26
Early career Female	Job Rank	-8.02	0.38
Early career Female	Job Rank jump	0.055	0.61
Early career Female	Attractiveness Score	4.96	0.34
Late career Female	Job Rank	-7.915	0.25
Late career Female	Job Rank jump	0.059	0.38
Late career Female	Attractiveness Score	4.89	0.34

Model	**Covariate**	**Value Distribution**
Early career Male	Ethnicity_category	Other(75.14%), Indian(21.26%), As(3.59%)
Early career Male	Undergraduate_Area	Science(80.89%), Arts(13.63%), Business(5.48%)
Early career Male	Undergraduate_university_ranking	post50(60.59%), top50(21.49%), top10(17.93%)
Early career Male	Postgraduate_university_ranking	top10(61.40%), top50(38.60%)
Late career Male	Ethnicity_category	Other (76.76%), Indian (20.13%), As(3.11%)
Late career Male	Undergraduate_Area	Science(80.70%), Arts(14.36%), Business(4.94%)
Late career Male	Undergraduate_university_ranking	post50(59.81%), top50(21.22%), top10(18.97%)
Late career Male	Postgraduate_university_ranking	top10(61.54%), top50(38.46%)

Early career Female	Ethnicity_category	Other (82.81%), Indian (14.19%), As(3.00%)
Early career Female	Undergraduate_Area	Science(56.13%), Arts(33.80%), Business(10.06%)
Early career Female	Undergraduate_university_ranking	post50(51.92%), top50(25.29%), top10(22.79%)
Early career Female	Postgraduate_university_ranking	top10(65.42%), top50(34.58%)
Late career Female	Ethnicity_category	Other (84.40%), Indian (13.08%), As(2.51%)
Late career Female	Undergraduate_Area	Science(54.41%), Arts(36.18%), Business(9.41%)
Late career Female	Undergraduate_university_ranking	post50(51.30%), top50(26.52%), top10(22.18%)
Late career Female	Postgraduate_university_ranking	top10(65.60%), top50(34.40%)

We find that coefficient for attractiveness shows significance for the two genders in different career phases (table 3). Firstly, for men we find significant (at 5% level) and positive impact of attractiveness on the hazard function in the early career phase however an insignificant impact in the late career stage. More specifically two standard deviation difference in attractiveness increases the probability of getting promoted in a given period by 22.6% within the first 3 years. It indicates that attractive men are able to extract reward from their looks in attaining more desirable roles only early in the career. This offers evidence for our argument that employers learn about the individual over time, and able to differentiate between them on true quality. Once this happens they are less likely to rely on or be swayed by superficial perceptions of ability, social skills or competence driven by looks.

For women however we observed exactly the opposite. We find that beauty effectively plays no role early in the career. This happens as a result of counteracting stereotypes and perceptions particularly about attractive women. Their career prospects are hurt by envy and jealousy on one hand, while assisted by personality perception and the value of their presentability. As we would expect, in later phase of their career the envious and jealous response to their career prospects diminish. Further any potential stereotype of “dumb-blonde” is proven wrong with their serious continued participation in the work force. Only a positive contribution of their beauty should persist, as we find with a positive coefficient significant at 5% level for women in late career phase.

Table 4: Parameter estimates for four models based on gender and career phase.

VARIABLES	(1) Model 1 – Early Career Male	(2) Model 2 - Late Career Male	(3) Model 3 – Early Career Female	(4) Model 4 – Late Career Female
Beauty	0.102**	-0.00725	0.00450	0.156**
	(0.0518)	(0.0483)	(0.0695)	(0.0662)
Rank Jump	-0.0928**	-0.0715*	0.0248	-0.196***
	(0.0417)	(0.0370)	(0.0700)	(0.0689)
Technical	-0.146	-0.309	-0.299	-0.219
	(0.547)	(0.425)	(0.594)	(0.701)
Consultant	-0.129*	0.318***	-0.0926	0.102
	(0.0741)	(0.0855)	(0.109)	(0.124)
Finance	-0.174	-0.158	-0.00379	-0.237
	(0.134)	(0.120)	(0.245)	(0.263)
Management	0.685***	-0.0280	-0.516	0.172
	(0.235)	(0.224)	(0.414)	(0.447)
Asian	0.301**	0.00628	-0.202	0.275
	(0.148)	(0.151)	(0.231)	(0.245)
Indian	0.0933	0.0923	0.216*	0.00129
	(0.0672)	(0.0646)	(0.119)	(0.120)
Undergrad – Top 10	-0.0118	-0.0603	0.254**	-0.171*
	(0.0640)	(0.0617)	(0.0986)	(0.0921)
Undergrad – Top 10-50	0.00149	-0.0610	0.118	-0.0416
	(0.0775)	(0.0732)	(0.109)	(0.106)
MBA – Top 10	0.108**	-0.0341	0.0898	-0.0432
	(0.0547)	(0.0508)	(0.0871)	(0.0817)
Undergrad – Arts	-0.0322	0.0296	-0.0293	-0.0669
	(0.0753)	(0.0710)	(0.0889)	(0.0828)
Undergrad –Business	0.00735	-0.0587	0.0140	0.0135
	(0.114)	(0.117)	(0.142)	(0.135)
Employer Switch	-0.0822	0.197***	-0.0655	-0.00198
	(0.0527)	(0.0475)	(0.0781)	(0.0737)
Constant	-2.961***	-1.922***	-2.698***	-2.603***
	(0.258)	(0.235)	(0.407)	(0.392)
Job Rank Dummies	Yes	Yes	Yes	Yes
Log(p)	1.2746***	0.8536***	1.2893***	0.8496***
	(0.0234)	(0.0175)	(0.0341)	(0.0293)
σ_u^2	0.9959***	0.5009***	1.0200**	0.3764***
	(0.0985)	(0.0497)	(0.1434)	(0.07416)
Observations	3,990	3,875	1,699	1,413
Number of Individuals	2,921	2,037	1,206	742

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Two standard deviation difference in attractiveness increases the probability of getting promoted in a given period by 36.6% for women beyond the first 3 years after graduation.

Beside attractiveness, our results also show a few more obvious characteristics that influence how fast an individual moves up. The estimated parameter for our chosen baseline hazard function (Weibull distribution) suggests that compared to late career, early in the career an individual is 50% more likely to jump upward within 1 year at their role. The baseline hazard function and the constant coefficient (both significant at 1% level) together describe the gender gap when compared across models for men and women. The gap is significantly wide late in the career, we find that men are 80% more likely to move up in a given period. Further as expected job switches where the starting job is already a senior position i.e. a high job rank are likely to be slower. While a switch from say an intern position would be much faster as indicated by the significance of the job type = intern indicator. Finally the coefficient for size of the jump has a positive coefficient which matches with the simple explanation that bigger jumps will take longer. These ancillary outcomes provide further confidence to our survival analysis.

7. Conclusion

From the point of employers in US labor market looking for personality fit among candidates is a very justifiable policy, however stereotypes swaying this judgement is a very likely outcome. Moreover any discernible bias purely in favor of beauty could cost employers both a highest quality workforce as well as a public relations nightmare. Literature in this area has typically shown long term premium from beauty for men. Our work uniquely establishes that beauty impact vanishes away for men as employers learn and are become informed of individuals true qualities. This also implies that eely rewards of beauty originate from visual perceptions. Any real correlation of beauty and ability would have fetched consistent lifetime gains, our results dismiss this possibility. In light of this, employers would do well to focus on ensuring fairness of early career hiring process. Policies such as oral only interviews to build an initial unbiased opinion could help.

For women, previous work shows both a lifetime premium as well as experimentally observed instances of penalties. We find that this confusion likely arises from different period of study. While an overall beauty premium exists, women in particular may suffer from compounded effects of multiple stereotypes at different stages of the career. Practitioners need to be lot more careful when designing policy, accounting for penalizing impact of jealousy as well as discriminating rewards arising from valuation of their presentability in currently male dominated work places.

References

1. Ahonen, Timo, Abdenour Hadid, and Matti Pietikainen. "Face description with local binary patterns: Application to face recognition." IEEE transactions on pattern analysis and machine intelligence 28.12 (2006): 2037-2041.
2. Ambekar, A., Ward, C., Mohammed, J., Male, S., and Skiena, S. 2009. Name-ethnicity classification from open sources. In Proceedings of the 15th ACM SIGKDD international Conference on Knowledge Discovery and Data Mining.
3. Amos, Brandon, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. OpenFace: A general-purpose face recognition library with mobile applications. Technical report, CMU-CS-16-118, CMU School of Computer Science, 2016.
4. Andreoni, James, and Ragan Petrie. "Beauty, gender and stereotypes: Evidence from laboratory experiments." Journal of Economic Psychology 29.1 (2008): 73-93.
5. Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994a. "The internal economics of the firm: Evidence from personnel data". Quarterly Journal of Economics 109, no. 4:881–919.

6. Biddle, J. E., D. S. Hamermesh. 1998. Beauty, productivity, and discrimination: Lawyers' looks and lucre. Journal of Labor Economics 16(1) 172–201.
7. Bradley J. Ruffle and Ze'ev Shtudiner. 2014. "Are Good-Looking People More Employable?". Management Science. 1760 – 1776.
8. Buss, D. M., M. Haselton. 2005. The evolution of jealousy. Trends in Cognitive Sciences 9(11) 506–507.
9. Caryl Rivers and Rosalind C. Barnett. 2016. "8 big problems for women in workplace". Chicago Tribune.
10. Christopher Y. Olivola and Alexander Todorov. 2010. "Elected in 100 milliseconds: Appearance-Based Trait Inferences and Voting". Journal of non verbal behavior. 34:83–110.
11. Cover, Thomas, and Peter Hart. "Nearest neighbor pattern classification." IEEE transactions on information theory 13.1 (1967): 21-27.
12. Daniel S Hamermesh. 2011. "Beauty Pays: Why attractive people are more succesful". Princeton University Press.
13. Dion, K., E. Berscheid, E. Walster. 1972. What is beautiful is good. Journal of Personality and Social Psychology 24 285–290.
14. Fruhen, L., Watkins, C. D. and Jones, B. C. (2015) Perceptions of facial attractiveness, dominance and trustworthiness predict managerial pay awarded in experimental tasks. Leadership Quarterly, 26(6), pp. 1005-1016.
15. George-Levi Gayle, Limor Golan and Robert A. Miller. 2012. "Gender Differences in Executive Compensation and Job Mobility". Journal of Labor Economic. Vol. 30, No. 4 (October 2012), pp. 829-872.
16. Hamermesh D. S., J. E. Biddle. 1994. Beauty and the labor market. American Economic Review 84(5) 1174–1194.

17. Heilman, M. E., Saruwatari, L. R. 1979. When beauty is beastly: The effects of appearance and sex on evaluations of job applicants for managerial and non-managerial jobs. Organizational Behavior and Human Performance 23 360–372.
18. Hossein K. Arani and Sajad D. Mobarakeh. 2011. “Sociolinguistics: Education, Women, Beauty, Discrimination & Exploitation: Investigating the Ugly Reality”. International Journal of Humanities and Social Science. Vol. 1 No. 21.
19. Kulka, Richard A., and Joan B. Kessler. "Is Justice Really Blind?–The Influence of Litigant Physical Attractiveness on Juridical Judgment." Journal of Applied Social Psychology 8.4 (1978): 366-381.
20. Lawrence, Steve, et al. "Face recognition: A convolutional neural-network approach." IEEE transactions on neural networks 8.1 (1997): 98-113.
21. Leventhal, Gloria, and Ronald Krate. "Physical attractiveness and severity of sentencing. "Psychological Reports 40.1 (1977): 315-318.
22. Mihalcea, Rada, Courtney Corley, and Carlo Strapparava. "Corpus-based and knowledge-based measures of text semantic similarity." AAAI. Vol. 6. 2006.
23. Mobius, Markus M. and Tanya S. Rosenblat. 2006. “Why beauty matters”. American Economic Review 96, no. 1: 222-235.
24. Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis. "Running experiments on amazon mechanical turk." (2010).
25. Richard B. Cohen. 2017. “Pretty people always win: Beauty bias in the workplace”. Abovethelaw.com
26. Richard Feloni. 2016. “Google's unconscious bias training presentation”. Business Insider.

27. Sagarin, B., D. V. Becker, R. E. Guadagno, L. D. Nicastle, A. Millevoi. 2003. Sex differences (and similarities) in jealousy: The moderating influence of infidelity experience and sexual orientation of the infidelity. Evolution and Human Behavior 24 17–23.

28. Santos, J. Reynaldo A. "Cronbach's alpha: A tool for assessing the reliability of scales." Journal of extension 37.2 (1999): 1-5.

29. Schmid, Kendra, David Marx, and Ashok Samal. "Computation of a face attractiveness index based on neoclassical canons, symmetry, and golden ratios." Pattern Recognition 41.8 (2008): 2710-2717.

30. Tanzina Vega. 2015. "Working while brown: What discrimination looks like now". CNN Money.

31. Turk, Matthew A., and Alex P. Pentland. "Face recognition using eigenfaces." Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on. IEEE, 1991.

32. Zebrowitz, Leslie A., and Susan M. McDonald. "The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts." Law and human behavior 15.6 (1991): 603.

33. Zhang, Xiaozheng, and Yongsheng Gao. "Face recognition across pose: A review." Pattern Recognition 42.11 (2009): 2876-2896.