

# CEO Selection and Executive Appearance

Douglas O. Cook  
University of Alabama Culverhouse College of Business  
dcook@cba.ua.edu

Shawn Mobbs   
University of Alabama Culverhouse College of Business  
smobbs@cba.ua.edu (corresponding author)

## Abstract

Survey assessments have found limited evidence of benefits of executive attractiveness. We use an objective measure of facial attractiveness that is correlated with survey assessments but less noisy and identify several benefits from executive facial attractiveness previously found in the general population but heretofore empirically elusive among executives. We examine the effect of both measures on executive compensation, promotion to CEO and the corresponding shareholder reaction, and promotion to board chair. The objective measure identifies significantly positive labor market effects for executive attractiveness in all outcomes in contrast to survey assessments of attractiveness that do not correlate with any outcome.

For the Lord sees not as man sees: *man looks at the outward appearance*  
*[when selecting a King]*, but the Lord looks on the heart.

1 Samuel 16:7 circa 1000 B.C.

## I. Introduction

The importance of quantitative individual executive characteristics in influencing firm policy decisions and outcomes is well established in the literature<sup>1</sup> but qualitative characteristics, though more difficult to measure, can perhaps be even more influential. Appearance is one such characteristic that is known to significantly influence workers' labor market outcomes (Hamermesh and Biddle (1994)) as well as impact financial contracting decisions (Duarte, Siegel, and Young

---

We are grateful to James Bates, Kathleen Centilli, Claiborn (Mac) Fountain, Kellie Hensley, and Katy Wallace for valuable research assistance. We have also benefited from conversations with and comments from Renée Adams, Anup Agrawal, Bruce Barrett, David Becher, Wesley Channell, Tony Cookson, Henrik Cronqvist, Jefferson Duarte (the referee), Daniel Greene, Daniel Hamermesh, Paul Malatesta (the editor), Bill Mayew, Amos Nadler, Micah Officer, Tijana Rajkovic, Jesus Salas, Geoffrey Tate, and seminar participants at The 2014 Financial Management Association Annual Conference, Georgia State University, Louisiana State University, the 2014 Mississippi State University Magnolia Finance Conference, the 2014 Northern Finance Association Annual Meeting, the 2015 European Finance Association Annual Meeting, and the 2016 Society of Financial Studies Cavalcade. Cook gratefully acknowledges financial support from the Ehney A. Camp, Jr., Chair of Finance and Investments.

<sup>1</sup>See Bertrand (2009) for a thorough review of the literature on CEO characteristics.

(2012), Brooks, Huang, Kearney, and Murray (2014), Blankespoor, Hendricks, and Miller (2017), and Ravina (2019)). In addition, as the above quote reveals, appearance has played a role, whether explicitly or implicitly, in leadership choices for centuries. Yet, empirical evidence on the role of appearance, particularly attractiveness, of executives is scarce or mixed at best.

One reason for this lack of definitive evidence on the influence of executive attractiveness is the reliance upon survey assessments to measure attractiveness. At first glance, survey assessments of attractiveness seem to be the most natural way to assess or measure beauty. Studies in psychology and economics have long relied on survey assessments of attractiveness. However, in these studies the subjects, decision makers in the experiment, are the ones influenced by attractiveness (e.g., Dipboye, Fromkin, and Wiback (1975), Dipboye, Arvey, and Terpstra (1977), and Heilman and Saruwatari (1979)). In contrast, when studying the effects of executive attractiveness, those making decisions about the executives are not the ones being surveyed. Since it is the board of directors who evaluate executives, set pay and appoint a CEO or chair, the ideal survey sample to study the influence of attractiveness on these decisions would be the directors, but this is not feasible. As a result, previous researchers have relied on survey results from a different sample of individuals to obtain a subjective assessment of executive attractiveness or other characteristics. Unfortunately, this approach is fraught with noise. First, demographic differences between survey participants and those for whom their assessment serves as a proxy can contribute to the noise. For example, Foo and Clark (2011) find that young adults rate younger faces as more attractive than older faces, but older adults do not make such a distinction. Thus, young survey participants may rate an older face as less attractive, while an older person may evaluate the same face as relatively more attractive. Second, any familiarity with the assessed subjects by the survey participants can also create noise (Zajonc (1968), Rhodes, Halberstadt, and Brajkovich (2001), and Peskin and Newell (2004)). These issues make replicating findings difficult and inhibit our understanding of the role of appearance in executive labor markets.

For example, using surveys obtained through Amazon's Mechanical Turk service, Duarte et al. (2012) find a significant link between the appearance of trustworthiness and actual individual creditworthiness, but no significant relation between attractiveness and creditworthiness. However, using a different sample of surveyors Ravina (2019) finds evidence of a significant relation between attractiveness and creditworthiness. Relatedly, Graham, Harvey, and Puri (2017) use survey assessments of executive appearance by graduate students of business and find a significant association between competence and executives who are CEOs but only find a weak relation between attractiveness and CEOs. Thus, while fundamental elements of executive attractiveness may indeed be influential, survey assessments of attractiveness from one group used to proxy for those of another have not definitively identified strong economic relations.

We address these concerns by introducing a new nonsurvey-based measure of fundamental attractiveness. Specifically, we use an objective measure of facial attractiveness that is a scientifically based and practitioner proven facial mask. We use the sum of absolute deviations from 25 key facial nodes on the mask to the corresponding nodes on an executive's face, as an objective measure of facial

attractiveness (smaller total deviations correspond to more attractive faces). For comparison, we also collect survey assessments of executive attractiveness in our sample. In these surveys, we find substantial variation and range of rater scores, reflecting the noise inherent in this measure of attractiveness. The mask-based measure correlates with the survey-based measure but is less noisy.

Using a sample of 237 executives most likely competing to be CEO in 100 firms, we examine the effect of attractiveness, using both the objective and survey measures, on several executive outcomes. First, we examine compensation and find that, similar to Graham et al. (2017), even after adjusting for the errors-in-variables (EIV) inherent in survey measures, we are unable to identify a strong significant relation between survey-based measures of attractiveness and compensation. However, we do find a strong relation using our objective mask-based measure that is economically meaningful. A 1-standard-deviation increase in executive attractiveness (smaller deviations) is associated with a 20.5% increase in compensation.

Since the objective measure is able to identify the heretofore-elusive economic link between attractiveness and executive pay, we analyze other executive labor market events to see whether the objective mask-based measure provides additional insights beyond those provided by survey-based measures. We find no association between the survey-based measure of attractiveness and the likelihood of an executive's promotion to CEO. However, we find strong evidence that facial attractiveness, measured using the objective mask-based measure, is significantly associated with an executive's promotion to CEO after controlling for titles, board membership, and executive pay ranking.

Focusing on the subsample of 100 newly selected CEOs, we examine whether the survey-based measure or the mask-based measure of attractiveness can identify if the new CEO's attractiveness influences shareholders. While both measures of attractiveness are associated with a positive shareholder reaction, only the mask-based measure's relation is significant. Lastly, consistent with attractiveness influencing decision makers (e.g., Mobius and Roensblat (2006)), we find that newly appointed CEOs who are more attractive according to the mask-based measure, are appointed to the board chair significantly faster. Again, we are unable to detect a significant relation when using the survey-based measure.

Our findings make several important contributions to the literature. First, researchers in the fields of psychology (e.g., Aronson and Mills (1965)), politics (e.g., Todorov, Mandisodza, Goren, and Hall (2005)), economics (e.g., Hamermesh and Biddle (1994)), and finance (e.g., Duarte et al. (2012)) have long been interested in the role of physical appearance in a variety of outcomes (Hatfield and Sprecher (1986)) including the job market. Why appearances matter in the job market and what relevant information appearances convey about job performance are important and largely unanswered questions in this literature.

One reason for the remaining unanswered questions despite this vast history of research is, in part, the difficulty of replicability. Most of the related prior research has relied on survey assessments of appearance. Replicating a study based on survey assessments requires utilizing the original survey participants, which is usually not practical nor feasible. Another reason for the inconclusiveness of prior research, particularly that in the executive labor market, is that survey assessments are only a proxy for the assessments of appearance held by the actual decision

makers (e.g., board of directors), which introduces another level of measurement error. Our findings highlight the difficulty in replicating prior findings on the role of appearance in the executive labor market when using a survey-based measure while showing that an objective-based measure is less susceptible to measurement error. Thus, our findings demonstrate that *how* appearance is measured matters. Having an objective measure of appearance that is less susceptible to measurement error and that facilitates replicability helps us to understand better the role of appearance in the executive labor market.

Using the objective measure of attractiveness, our results complement and extend earlier findings on the benefits of attractiveness widely documented in the general population to the executive population. Given the vast evidence of the influence of attractiveness among the population (e.g., Aronson and Mills (1965), Widgery and Webster (1965), Goldman and Lewis (1977), Guise, Pollans, and Turkat (1982), and Warner and Sugarman (1986)), it is reasonable to expect executive attractiveness to play a role in executive labor market outcomes. By using an objective measure, we confirm this intuitive but empirically elusive relation. Being able to identify associations between fundamental elements of executive attractiveness also helps to shed light on prior findings. For example, our finding that attractive CEOs receive the board chair title faster suggests attractiveness can facilitate entrenchment and thus weaken any positive relation between CEO attractiveness and firm performance. This, in addition to the EIV problem, may help further explain why Graham et al. (2017) do not find a significant relation between survey assessments of CEO attractiveness and firm performance.

Our use of an objective measure of fundamental facial attractiveness also extends the recent literature that is beginning to use objective measures of qualitative executive traits. Hsieh, Kim, Wang, and Wang (2020) use a machine learning based facial-feature-point detection methodology to measure CFO trustworthiness and the corresponding association with audit fees. They apply their measure, which they base on an index related to four facial features: angle of inner eyebrow ridges, face roundness, chin width, and nose-to-lip distance, to Google Images of CEOs and CFOs and find more trustworthy-appearing executives to be associated with lower audit fees. Halford and Hsu (2020) measure CEO attractiveness using a web-based facial analysis application on Anaface.com to study the relation between CEO attractiveness and shareholder reaction to their appointment and merger and acquisition decisions. To implement this, they sample each CEO six times through Anaface and take the mean of the measures. Unfortunately, the authors are not privy to the computational algorithm used by Anaface. Both studies are able to identify significant relations between their respective measures of facial appearance and the outcomes they examine. Similar to our approach, both studies base their measure of either trustworthiness or attractiveness on measurements from key nodes on the executives' faces. However, since the exact algorithms used are not disclosed, it is unclear on what basis attractiveness is assessed. Conversely, our approach provides a clear objective standard from which to measure deviations, providing a simple, replicable and systematic means of measuring attractiveness.

Our evidence also makes a significant contribution to the literature by uncovering a new and important executive trait, attractiveness. Unlike other characteristics, such as education or prior experience, with the exception of elective cosmetic

surgery, which is certainly feasible, elemental attractiveness is not alterable through experience or choices. As such, it represents an innate trait that executives receive at birth that can influence their executive labor market outcomes later in life, much like one's potential IQ or height (Adams, Keloharju, and Snupfer (2018)). More generally, our findings that beauty plays a role in executive labor markets complements studies of the value of beauty in other settings like education (Hatfield and Sprecher (1986)) and political contests (Todorov et al. (2005)).

## II. Measuring Attractiveness and Related Research

### A. An Objective Measure of Attractiveness

Recently, researchers have begun to exploit new technologies to help quantify and assess investor responses to various types of previously hard-to-measure qualitative information. For example, Mayew and Venkatachalam (2012) use voice analysis software to detect and quantify qualitative information in the variation of emotion in executives' voices during earnings conference calls. Similarly, Mayew, Parsons, and Venkatachalam (2013) find evidence that male CEOs with a deeper pitch voice head larger companies, make more money, and are forced out less often. While these new technologies do not capture every aspect of the qualitative characteristics they are measuring, for example, voice analysis software does not capture hand gestures, facial expressions, or body language used by the speaker that also convey information about executive emotion, they do capture important information related to these characteristics.

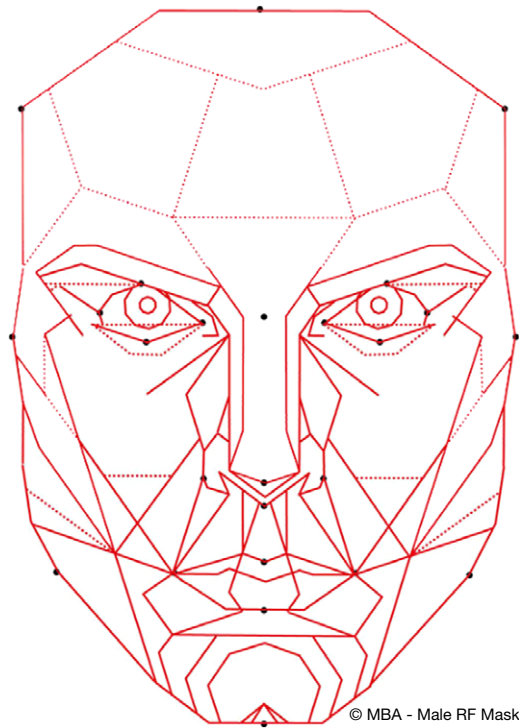
Similar to these technologies, an objective mask-based measure of facial attractiveness provides a simple means of systematically quantifying a previously difficult-to-measure source of qualitative information – elemental components of facial attractiveness. Scientists in the field of psychology have long studied what people find attractive or esthetically pleasing. Throughout this research, which dates as far back as Fechner (1871), scientists have documented consistent evidence of strong preferences for objects or appearances in proportions of 1.618:1, often referred to as the golden ratio, or phi. The presence of phi as a measure of esthetic appeal is well known in art (*The Sacrament of the Last Supper* by Salvador Dali), music (*Sonneries de la Rose + Croix* by Erik Satie), architecture (Egyptian pyramids), design (shape of postcards), nature (geometry of crystals) and mathematics (shape of pentagon). The value of phi ( $\varphi$ ), derived as: quantities  $a$  and  $b$  are in the golden ratio  $\varphi$  if:

$$\frac{a+b}{a} = \frac{a}{b} = \varphi \Rightarrow \varphi^2 - \varphi - 1 = 0 \Rightarrow \varphi = \frac{1 + \sqrt{5}}{2} = 1.618034.$$

While the preference for the golden ratio, over others such as unity, is still being studied (Davis and Jahnke (1991)), more recent practitioner reports reveal additional support for the use of the golden ratio, specifically as an objective standard of beauty for the human face. Dr. Stephen Marquardt developed a proprietary facial mask based on the golden ratio for the purpose of cosmetic surgery (Marquardt (2002)). While studying photographs of faces around the world

FIGURE 1  
Male Phi-Mask

Figure 1 displays the male Phi-Mask and the 25 nodes used to measure deviations from the mask to the corresponding points on the executive's face.

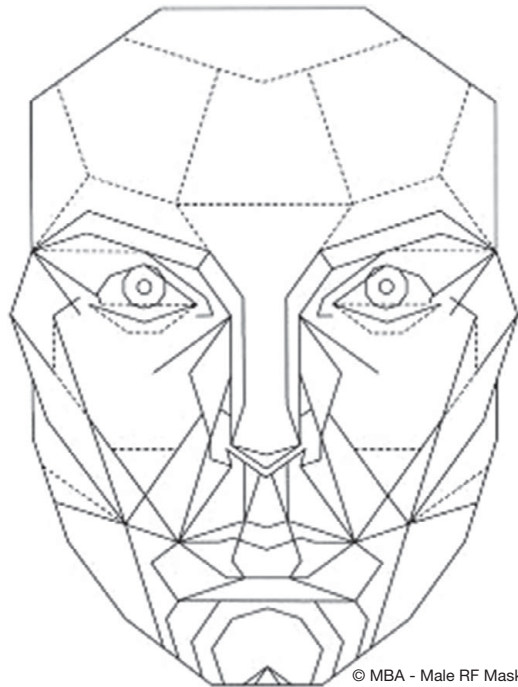


considered to be beautiful, Dr. Marquardt found that the golden ratio was prominently displayed in two dimensions. Based on these observations, he developed a facial mask known as the Phi-Mask. Male and female Phi-Masks are displayed in Figures 1 and 2, respectively. In an interview with the Journal of Clinical Orthodontics, Marquardt says “The simplest configuration that describes the Golden Ratio in two dimensions is an acute Golden Triangle with sides of 1.618 and a base of 1, or an obtuse Golden Triangle with a base of 1.618 and sides of 1. Together these elements form a Golden regular pentagon, and the regular pentagon itself, if duplicated, inverted, and superimposed on itself, forms the Golden Decagon – a regular vertex radial decagon.” The details of the derivation and properties of this Phi-Mask are available in the patent documents.<sup>2</sup>

<sup>2</sup>Quoting from Bashour (2006b): “Marquardt uses the primary pentagonal complex to form the basic framework of the mask, using specific lines, line segments, and points to construct the component lines and points of the mask. Various sized secondary decagons or pentagonal complexes related mathematically to the primary complex (the size of the subsidiary pentagon complexes is derived using the formula  $PC[\text{subject}] = PC[\text{reference}] \times (1/\phi)^n \times Z$  where  $PC[\text{subject}]$  stands for the main pentagon complex shown in Figure 2 [In Bashour’s article, Figure 2 refers to the Phi-Mask framed by the reference pentagonal complex.],  $n = 0$ ,  $Z = 1$ ,  $\phi$  is the golden ratio, and  $n$  and  $Z$  are variables used) are used to

FIGURE 2  
Female Phi-Mask

Figure 2 displays the female Phi-Mask used to measure deviations from the mask to the corresponding points on the executive's face.



© MBA - Male RF Mask

The details of the mask make it a useful measure of attractiveness since it captures more aspects of facial appeal than prior measures that rely on simple characteristics. Consistent with this intuition, Schmid, Marx, and Samal (2008) find that measures of facial attractiveness based on the golden ratio better capture beauty than do simple measures of symmetry. Since its creation, there have been several studies evaluating the use of the mask as an objective measure of facial beauty. For example, the American Society of Plastic Surgeons published two articles, Bashour (2006a), (2006b), in the journal of *Plastic and Reconstructive*

---

derive the remaining component lines and points of the mask:  $[n = 6; Z = 1]$  is used twice for the two iris complexes;  $[n = 5; Z = 1]$  is used 3 times, for the nasal tip complex, the internal lip complex, and the internal nares complex;  $[n = 5; Z = \phi^{1/3}]$  is used once as the inner nasal tip halo complex;  $[n = 5; Z = 2/\phi]$  is used once as the outer nasal tip halo complex;  $[n = 4, Z = 1]$  is used 14 times, for the nose/mouth complex, the mouth/chin complex, chin inferior border complex, chin complex, right and left sided chin complexes, right and left eyebrow complexes; right and left cheek complexes, and right and left nose/mouth complexes;  $[n = 3, Z = \phi^{1/3}]$  is used twice for the right and left eyebrow/cheek complexes;  $[n = 2, Z = 1]$  is used once for the frontal repose smile complex; and  $[n = 1, Z = 1]$  is used once for the internal facial pentagon system.”



*Surgery* on the use of the golden ratio as an objective measure useful in reconstructing facial features to enhance appearance.<sup>3</sup>

Just as voice analysis software quantifies important information conveyed in executive emotions, the facial mask technology allows us to quantify important aspects of attractiveness, which can be informative to shareholders or directors as they formulate their expectations of executives' abilities. Also, like these other technologies, it is not perfect. This two-dimensional objective measure does not fully capture all aspects of attractiveness, such as height or body type. Nor does it capture all dimensions of facial attractiveness such as smiles or skin tones.<sup>4</sup> Furthermore, we cannot be certain that our measure does not correlate with other important factors associated with executive ability such as perceived trustworthiness that may be reflected in appearance through one's genotype. Thus, the Sum of Mask Deviations from the Phi-Mask could capture personality traits that are important in a CEO. Nonetheless, this objective measure provides a systematic way to quantify fundamental features of facial attractiveness that is not subject to the varied biases of survey participants. Such a measure can contribute significantly to research in labor economics on this previously difficult to assess, but important, individual characteristic.

## B. Challenges With Survey-Based Measurements of Attractiveness

If beauty is indeed in the "eye of the beholder," then using survey assessments of attractiveness from one population, say Amazon's Mechanical Turk workers, as a proxy for how another population, say the board of directors, assesses attractiveness can be problematic. Differences in age, psychological, economic, cultural, and social status between survey participants and the board of directors, whose assessment of attractiveness they are to represent, can make this proxy very noisy (Foo and Clark (2011)). Furthermore, the familiarity associated with the assessments of highly visible executives involved in a CEO selection tournament can introduce additional noise into the survey assessments (Zajonc (1968), Rhodes et al. (2001), and Peskin and Newell (2004)).

These various sources of noise result in an attenuation bias when trying to identify any economic impact from fundamental elements of executive attractiveness. Empirically, this is an EIV problem, which results from an imprecise measurement of the variable of interests (attractiveness). We illustrate this using a modified illustration from Greene (1993). Assume that the model

<sup>3</sup>Bashour (2006a), (2006b) documents that deviation from the Phi-Mask explains 25% to 75% (depending on methodology) of the variation in attractiveness judgments formed by Internet and direct survey judges for 37 male, 35 female, and 31 composite faces. Using 40 cases, Kim (2007) also finds the Phi-Mask to be a useful analytical tool for facial surgery. The Phi-Mask is not a perfect measure. Although Marquardt studied faces from around the world in developing the mask to ensure it captures global elements of attractiveness, various ethnic groups may give different weights to different aspects of the face when determining attractiveness. Holland (2008) suggests the female Phi-Mask captures masculinized European females. Despite these shortcomings, the mask does capture some rudimentary elements of attractiveness common to numerous ethnicities and thus allows us to quantify and objectively measure attractiveness.

<sup>4</sup>We explore the impact of smiles further in the Supplementary Material.



$$(1) \quad y = \beta x^* + \varepsilon$$

fits the classical normal model, where  $y$  is an executive outcome and  $x^*$  is fundamental executive attractiveness. Since executive attractiveness is not measured precisely, assume that  $x$  is the survey-based measure of executive attractiveness that measures fundamental attractiveness with some error  $u$ :

$$(2) \quad x = x^* + u \text{ with } u \sim N[0, \sigma_u^2].$$

Assume that  $u$  is independent of  $y$  and  $x^*$  and  $\sigma_u^2$  is positive. Substituting equation (2) into (1) yields:

$$(3) \quad y = \beta x + [\varepsilon - \beta u] = \beta x + w.$$

Since  $x$  equals  $x^* + u$ , the regressor in (3) is correlated with the disturbance:

$$(4) \quad \text{cov}[x, w] = \text{cov}[x^* + u, \varepsilon - \beta u] = -\beta \sigma_u^2.$$

Since this is a violation of a central assumption of the classical model, the least squares estimate of  $\beta$  is inconsistent and biased toward zero. This bias is worse the greater the variability in the measurement error, and is referred to as attenuation. The intuition is that when we regress  $y$  on  $x$ , we have omitted a variable  $u$ , from the regression. If  $u$  were observed, then there would be no identification problem. Thus, the key to estimating  $\beta$  is to produce information about the degree of measurement error. Unfortunately, for survey assessments of attractiveness, it is impossible to determine the error with which fundamental attractiveness is measured. In Section IV.A, we discuss how to estimate a minimum feasible level of reliability for the survey measure of attractiveness and from that estimate how to implement this estimation. However, an objective measure of attractiveness alleviates this concern.

### III. Data

To compare the ability of these two measures of attractiveness, survey assessments and our mask-based measure, to identify associations between attractiveness and executive labor market outcomes, we focus on executives who are likely candidates for promotion to CEO. We start by identifying 351 succession events at S&P 500 firms between 2000 and 2009, excluding transitions due to mergers or acquisitions or when the new CEO is an interim. We also determine whether the new CEO appointment followed a forced or voluntary departure of the outgoing CEO. To obtain an attractiveness score using the facial mask, we first used Google Images to extract multiple pictures of each of the newly appointed CEOs from annual reports and other internet sources around the time of the succession event. We then provided these images to a graphical designer who filtered out the cases where she could not apply the mask. For example, the designer discarded images when CEO pictures were in profile or where the picture pixels were limited, thereby, preventing facial node identification. This process yielded a mask score for 100 new CEO

pictures. We then applied a similar screening strategy when collecting pictures for the non-CEO executives of these 100 firms. In total, we obtained facial attractiveness measurements for 255 executives in 100 firms that appointed a new CEO between 2000 and 2009.

We identified executives in these 100 firms most likely to be competing to be CEO by considering only those whom each firm employed during the appointment year and the prior year. We focus on these candidates since most CEO successions are internal (over 80% of our sample firms) and acquiring pictures and mask measurements for all possible external executives would be prohibitively expensive and provide little additional insight.

Among the internal executives, we use board seats, relative compensation and age to identify those most likely to be competing to be CEO. Executives competing to be CEO are often appointed to the board of directors during the evaluation process (Hermalin and Weisbach (1988)) and their compensation is likely greater than other executives (Murphy (1985), Bognanno (2001)). Thus, we consider an executive to be a CEO candidate if the executive is also on the board or if the executive's compensation ranking is in the top three within the firm. Finally, we exclude executives older than 65 since they are not likely to be long-term CEO candidates. In our sample, we find the average (median) number of executives competing to be CEO is 2.37 (2). Three firms had zero internal candidates using our criteria. Thirteen had only one, presumably an heir apparent. (Interestingly, some of these heirs were not selected as the firm went with an outside CEO.) Forty-two had 2 competitors, 31 firms had 3, 8 had 4, and 3 had 5.

The final sample includes 255 executives, 100 of which became CEO while the remaining 155 were runner-ups. The sample size is similar to those used in the various tests of Graham et al. (2017). Of the new CEO appointments, 18 are external and 82 are internal appointments. Excluding the externally appointed CEOs results in 237 internal candidates competing in CEO tournaments across the 100 sample firms. Of the internal candidates, 82 are winners and in these firms, there are 123 runner-ups. Among firms selecting an external CEO, there are 32 internal candidates not chosen.

## A. Appearance Score: Phi-Mask and Survey Assessments

We measure elemental facial attractiveness using deviations of 25 key points on executive faces from the corresponding nodes of the Phi-Mask. Specifically, for each sample executive we search annual reports and use Google Images to locate a facial photo taken as near to the CEO announcement date as possible. We impose this restriction to best capture the image evaluated by the board and do not use images more than a year after appointment to avoid changes in executive appearance due to age. We hire graphic designers<sup>5</sup> to align each facial image with that of the Phi-Mask and then measure absolute deviations from the 25 nodes on the Phi-Mask to the corresponding nodes on the executive's face in units of

<sup>5</sup>In the Supplementary Material, we discuss the hiring of a second graphic designer and compare Phi-Mask measurements. We also decompose the Phi-Mask to examine the explanatory power of various components of the face.

points.<sup>6</sup> The sum of these 25 deviations represents our measure of executive facial attractiveness where faces with larger (smaller) deviations are less (more) attractive. We compute this measure for all 255 executives.

The second measure of attractiveness is based on survey assessments. Recent papers on appearance have used survey assessments from MBA students (Graham et al. (2017)) or workers from Amazon's Mechanical Turk service (Duarte et al. (2012)). We follow Duarte et al. and employ Amazon's Mechanical Turk service to obtain 25 independent workers to rate each executive photograph on the attribute of facial attractiveness, with 10 being very attractive and 1 being very unattractive. We average the ratings for each executive to derive a Survey Attractiveness Score. For ease of comparison of these two different measures of executive facial attractiveness, in the regression analyses, we report results using the standardized values of these two measure of attractiveness as well as all control variables.

## B. Descriptive Statistics

To ensure that the 100 CEO transitions we examine are not systematically different from the other CEO transitions, we identify, collect, report, and compare key descriptive statistics on all 351 newly selected CEOs in S&P 500 firms within our sample period. Panel A of Table 1 reports these statistics. Nineteen percent of the new CEOs are from outside the firm. Nineteen percent also have prior CEO experience and 13% replace a forcefully removed CEO.

Next, we compare these characteristics for the subsample of CEOs with a mask deviation measure to the subsample of CEOs without a mask deviation measure. The last two columns in Panel A of Table 1 report *t*-tests of the differences in the means and the corresponding *p*-values. These columns show that the two subsamples of CEOs are not statistically different based on prior experience, or age, or whether or not they are outsiders. They differ only in the fraction of females and the fraction following a forced CEO departure. There are 4 female CEOs in our sample with facial measurements and none in the remaining sample. In our sample, 18% of the CEOs followed a forced CEO departure, compared to 10% of the remaining sample.

Panel B of Table 1 reports descriptive statistics for the 237 internal executives competing to be the next CEO for these 100 CEO transitions. The average age of the executive pool is 53 and females represent about 8% of the executive sample. The mean Sum of Mask Deviations is 381.3 points and the mean Survey Attractiveness Score is 4.47.

## C. Variation in Survey Assessments of Attractiveness

Using multiple raters and taking the average reduces noise but if there is a wide range of attractiveness assessments among the raters, the resulting proxy may still be a very noisy proxy for the effects of attractiveness on the decision agents, in our case, the board of directors or shareholders. While the issue of rater variation is known, it is difficult to quantify. For example, coefficient alpha (or Cronbach's

<sup>6</sup>A point is a unit of length commonly used to measure the height of a font equal to 1/72th of an inch. The distance in points is invariant to the screen size of the measured image.

TABLE 1  
New CEO and Executive Characteristics

Table 1 reports descriptive statistics for 351 newly appointed CEOs and the non-CEO executive CEO candidates during our sample period from 2000 to 2009. Panel A reports statistics for the subsample of 100 new CEOs for which we obtain an attractiveness measurement using the Phi-Mask and the subsample of 251 for which we do not calculate Phi-Mask deviations. OUTSIDER equals 0 if the new CEO was not a current employee of the firm. PRIOR\_CEO\_EXPERIENCE equals 0 if the new CEO was a CEO at least once before and 0 otherwise. AGE is the age of the executive at the time of the CEO selection. FEMALE equals 0 if the executive is female and 0 if male. SUM\_OF\_MASK\_DEVIATIONS is the sum of the deviations from 25 nodes on the Phi-Mask to the actual point on the executive's face as measured in Adobe Photoshop in units of points. Panel B reports descriptive statistics for the executives at the 100 CEO transitions in our sample of CEOs with mask measurements who are most likely candidates to be the next CEO. This sample consists of 82 executives selected to be the next CEO and 155 who were not selected. SURVEY\_ATTRACTIVENESS\_SCORE is measured as the average score for each executive from survey participants on Amazon's Mechanical Turk, where each executive is ranked on a scale of 1 to 10, with 10 being the most attractive looking. Panel C reports the estimate of interrater variability (square root of the average variance for the 237 internal executive candidates) for each executive by quartile of survey based attractiveness. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

	Full Sample				CEOs With Mask				CEOs Without Mask				(With-Without)	
	<u>N</u>	<u>Mean</u>	<u>Median</u>	<u>Std. Dev.</u>	<u>N</u>	<u>Mean</u>	<u>Median</u>	<u>Std. Dev.</u>	<u>N</u>	<u>Mean</u>	<u>Median</u>	<u>Std. Dev.</u>	<u>Mean</u>	<u>p-Value</u>
OUTSIDER	351	0.19	0.00	0.39	100	0.18	0.00	0.39	251	0.19	0.00	0.39	-0.008	(0.88)
PRIOR_CEO_EXPERIENCE	351	0.19	0.00	0.39	100	0.15	0.00	0.36	251	0.21	0.00	0.41	-0.0572	(0.23)
AGE	351	50.55	51.00	6.20	100	50.21	51.00	5.79	251	50.68	50.00	6.36	-0.4713	(0.52)
FEMALE	351	0.01	0.00	0.10	100	0.04	0.00	0.20	251	0.00	0.00	0.00	0.03***	(0.01)
FOLLOWED_FORCED_CEO_DEPART	351	0.13	0.00	0.33	100	0.18	0.00	0.39	251	0.10	0.00	0.31	0.0764*	(0.05)

Panel B. Internal Executive Candidate Characteristics

	<u>N</u>	<u>Mean</u>	<u>Median</u>	<u>Std. Dev.</u>
SUM_OF_MASK_DEVIATIONS	237	381.3	367.5	162.4
AGE	237	53.14	54.00	6.47
FEMALE	237	0.08	0.00	0.27
SURVEY_ATTRACTIVENESS_SCORE	237	4.47	4.36	1.11

Panel C. Survey Measure Interrater Variability of 25 Mechanical-Turk Raters for Each Executive

	<u>Q0</u>	<u>Q1</u>	<u>Q2</u>	<u>Q3</u>	<u>F-Test (p-Values)</u>			
					<u>Q1-Q0</u>	<u>Q2-Q0</u>	<u>Q1-Q3</u>	<u>Q2-Q3</u>
Interrater variability	1.51	1.66	1.70	1.56	(<0.01)***	(<0.01)***	(0.02)**	(<0.01)***

Alpha (Cronbach (1947)) is a measure of interrater reliability reported by several papers using survey assessments of attractiveness (e.g., Biddle and Hamermesh (1998), Ravina (2019)), but it is hampered by the fact that it increases with the number of survey participants (Cortina (1993)).<sup>7</sup> Thus, more raters and a higher coefficient alpha do not imply less noise. Duarte et al. (2012) illustrate in a simulation exercise reported in their Supplementary Material that the EIV problem arising from variation in survey assessments of attractiveness and other surveyed characteristics can still obfuscate true relations. Given the concern inherent with using multiple raters to generate survey assessments of attractiveness, we next closely examine the variation within our Mechanical Turk rater scores.

As a first estimate of the variation among the Mechanical Turk raters of each executive's attractiveness, we calculate an estimate of the interrater variability among the rating scores of each executive's picture. To derive an estimate of interrater variability for the sample, we first compute the survey rating variance among the 25 raters for each of our 237 executives. We first calculate the pooled estimate of variance for our 237 executives by taking the average of these 237 variances. We then take the square root of this pooled estimate of variance as our measure of interrater variability, which for the full sample is 1.61. This degree of variation reflects the noise inherent in survey assessments of attractiveness. However, it is quite possible that this variation is not uniform across the sample. For example, since survey scores are within a finite range, truncation on either end of the scale may produce variation that is greater in the middle of the distribution and smaller in the two tails. It is also reasonable to expect that there would be less variation among survey participants when rating the most or least attractive executives.

To test this conjecture, we divide our sample of executives into quartiles based on the mean Mechanical Turk rater scores, with the least (most) attractive being in the lowest (highest) quartile. As a first step, a 1-way ANOVA analysis of these four groups reveals that the variation across these four groups are not equal. In Panel C of Table 1, we report the estimate of the interrater variability for each quartile. The greatest variation is in the middle two quartiles, which is where we expect survey participants to vary the most, among the sample of "average" looking executives. Next, we conduct pairwise *F*-tests for the equality of variances across each pair of quartiles. We find that the variation across the middle two quartiles is not significantly different. However, the variation in each of these middle two quartiles is significantly greater than the variation in both the top and bottom quartiles at the 5% level or better. Lastly, there is less variation, more agreement among raters, in the two extreme quartiles, the least and most attractive executives. However, even within the top and bottom quartiles, the estimate of interrater variability is still over 1.5 points out of a possible 10. The results in Panel C of Table 1 highlight two important aspects of survey assessments of attractiveness. First, there is considerable interrater variability, which reflects the noise inherent in surveys. Second, this

---

<sup>7</sup>The Cronbach Alpha for our sample is 0.915, which reflects high interrater reliability, when using all 25 raters. However, if we only consider 6 raters (randomly selecting the first 3 and the last 3), the Cronbach Alpha drops to 0.701, which is very similar to that reported in Ravina (2019) (0.766), which also uses only 6 raters.

interrater variability also varies with executive attractiveness. These multiple degrees of variation underscore a fundamental challenge with survey assessments of attractiveness. To further explore this variability we consider rater range.

Across the full sample, we find the average rater range is 6.6 and 55% of our observations have a range greater than 6. However, since ranges are driven by outliers, we next focus on the ratings that are in agreement by looking at the differences between the most common rater score (mode) and the second most common score. If these two differ by only 1 it suggests greater agreement among the raters in the executive's attractiveness. However, if there is greater disagreement among raters, then this difference is likely greater than 1. The difference between the most common and the second most common score within an executive's set of ratings is greater than 1 for 37.6% of our sample (28.3% differ by 2, 7.2% differ by 3, and 1.7% differ by 4). These statistics reveal that the noise inherent in the survey data on attractiveness creates an EIV problem that contributes to an attenuation bias, which obfuscates any possible links between attractiveness and the studied outcomes.

#### D. A Comparison of Survey Assessments and Objective Assessments of Executive Attractiveness

Next, we examine sources of variation among survey assessments. We expect the mask-based measure to be correlated with survey assessments (Bashour (2006a), (2006b)), but we also consider several other sources of variation, such as age, gender, race, and familiarity. Age is of particular importance in our study given that Mechanical Turk workers are much younger than directors in large firms. Duarte et al. (2012) surveyed a subsample of Mechanical Turk workers used in their analysis and found that 52% were less than 35 years of age, while 85% were less than 50 years of age. Furthermore, the Mechanical Turk workers' population is not as large or as diverse as it appears (Stewart, Ungemach, Harris, Bartels, Newell, Paolacci, and Chandler (2015)).

We also control for firm size (market capitalization) to capture the visibility of firms as a proxy for the visibility of their executives. Finally, we use Google Trends data from 2013 to June 2015, the end of our survey period, to proxy for survey participants' familiarity with the executives in our sample. The Google Trend score is indexed and scaled so that a high score indicates the search term was searched more frequently throughout the trend period and a low score reflects relatively infrequent search activity for the term. We use the firm name as the search term, under the premise that survey participants are more likely to be familiar with the executives of firms with which they are more familiar.<sup>8</sup>

In Table 2, model 1, we regress the survey assessment measure on the Sum of Mask Deviations and other executive and firm characteristics. Since we have multiple executives per firm, we report robust standard errors clustered by firm. We find that the masked-based objective measure of attractiveness correlates significantly with survey assessments of attractiveness at the 5% level. However,

<sup>8</sup>Although we also find a positive association when we use the executive names as the search terms, many of the names had insufficient search volume, which resulted in a trend score of 0 for a majority of the executives and a coefficient estimate not statistically significant at traditional levels.

TABLE 2  
Determinants of Measures of Executive Attractiveness

Table 2 reports the results of OLS regression analysis of executive attractiveness measures. The dependent variable in model 1 is a measure of the SURVEY\_ASSESSMENT\_OF\_ATTRACTIVENESS, which is the average rating from Mechanical Turk workers who rate each executive's attractiveness on a scale of 1 to 10, with 10 being the most attractive. The dependent variable in model 2 is the SUM\_OF\_MASK\_DEVIATIONS. Sum of Mask Deviations is the absolute sum of deviations from each of the 25 nodes on the Phi-Mask to the corresponding point on the executive's face. Standard errors are robust and clustered by firm. *p*-values are reported in parentheses beneath the coefficient estimates. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable	
	SURVEY_ASSESSMENT_OF_ATTRACTIVENESS	SUM_OF_MASK_DEVIATIONS
	Model 1	Model 2
SUM_OF_MASK_DEVIATIONS	-0.105** (0.04)	
AGE	-0.3*** (<0.01)	-0.164** (0.02)
FEMALE	0.17** (0.01)	-0.01 (0.82)
BALD	-0.25*** (<0.01)	0.03 (0.59)
ASIAN	-0.138** (0.02)	-0.05 (0.42)
AFRICAN_AMERICAN	0.16*** (<0.01)	-0.039 (0.6)
GLASSES	-0.191*** (<0.01)	-0.04 (0.56)
WHITE_HAIR	-0.12** (0.04)	-0.036 (0.56)
FACIAL_HAIR	-0.077 (0.15)	0.16** (0.03)
FIRM_GOOGLE_TREND_SEARCH_SCORE_2013-2015	0.117* (0.09)	
ln(MARKET_CAPITALIZATION)	-0.08 (0.22)	
OPEN_SMILE		0.0938 (0.15)
CONSTANT	-0.022 (0.71)	0.024 (0.69)
No. of obs.	237	237
R <sup>2</sup>	33.99%	6.47%

we also find strong evidence of significant age-related effects. Specifically, executives who are older, bald, wear glasses, or have white hair are associated with significantly lower survey assessments. Those survey assessments are significantly and negatively related to age is important because not only does it reveal an age bias, but this bias can vary greatly across survey participants. If different people age differently, this phenotypical effect can increase variation in survey assessments of attractiveness due to survey assessors' varying age-related biases.<sup>9</sup> We also find that females and African Americans are associated with higher assessment scores and Asians are associated with lower scores. The Google Trends Search Score measure of familiarity relates positively to the survey assessment of attractiveness and is significant at the 10% level, which is consistent with familiarity bias being embedded in the survey measure of attractiveness. Together, these results suggest that

<sup>9</sup>We thank an anonymous referee for providing this insight.



subjective assessments of the attractiveness of senior executives are very noisy. Thus, the objective mask-based measure, which is correlated with subjective assessments, but is not subject to these biases can be a good alternative.

To assess the effect of these same characteristics on the Phi-Mask measure, we run a similar regression and report the results in model 2, with the dependent variable being the Sum of Mask Deviations. The only characteristics with significant coefficient estimates are age and facial hair. Interestingly, both of these traits can affect the precision with which any objective measure of facial appearances, such as mask deviations, is measured. For example, facial hair can obscure key facial nodes. Thus, it is not surprising that executives with facial hair are associated with greater mask deviations. Conversely, the coefficient on age is negative and significant. While the mask measure is not prone to the same phenotype effects as survey-based measures, it may capture a different phenotype effect. For example, if age-related facial changes make the underlying bone structure, or fundamental elements of facial attractiveness, more prominent, perhaps due to thinning of the skin, this can reduce measurement noise in the mask deviation measurements. Thus, while an older face with thinner skin may be viewed by survey participants as relatively less attractive (model 1), the face may actually have smaller mask deviations because one can better capture fundamental elements of attractiveness with less noise.

These two results suggest that other executive traits may also affect the precision of mask deviation measurements. For example, if more fit executives have less facial fat their mask deviations can also be measured with less noise. Thus, to the degree that executive characteristics, such as age or fitness, affect the precision of mask deviation measurements, the mask may capture these traits. Despite these shortcomings, the effects appear to be small. A 1-year increase in age is associated with only a 4.15 point reduction in the Sum of Mask Deviations, which is only 1.1% of the sample mean and median. Moreover, the decrease in deviations from a 10-year increase represents only one-fourth of a standard deviation of the Sum of Mask Deviations. The coefficient for facial hair suggests a similar magnitude. Thus, while mask measurements may be affected by executive age, fitness, or grooming, these effects appear to be very small.

#### IV. Attractiveness and Executive Labor Market Outcomes

In this section, we use both measures of executive facial attractiveness and examine several economic outcomes for which attractiveness is likely to play an important role.

##### A. Executive Facial Attractiveness and Compensation

While numerous studies and books (e.g., Hamermesh and Biddle (1994), Hamermesh (2011)) document that attractiveness is associated with higher compensation in the labor market, there is little evidence that “Beauty Pays” among corporate executives. Graham et al. (2017) find weak evidence that subjective assessments of attractiveness are associated with higher compensation among

CEOs but they find a stronger association with subjective assessments of competence. In this section, we explore whether our objective measure of facial attractiveness, which is less noisy than survey assessments, provides empirical evidence of this connection.

In Table 3, we report results replicating the Graham et al. (2017) findings and regress the natural log of total compensation on firm size, operating performance, and industry and year fixed effects using our pool of executive CEO candidates.<sup>10</sup> We also report robust standard errors clustered by firm since we have multiple executive observations in each firm year. In model 1, we find a positive association between the survey-based measure of executive attractiveness and pay, but the relation is not statistically significant. In model 2, we use the survey-based measure of competence and find a positive and significant relation between competence and pay. A 1-standard-deviation increase in survey competence score is associated with a 14.1% increase in compensation (or \$0.923 million relative to our sample mean). This is similar to Graham et al. (2017) who document that a 1-standard-deviation in survey-assessed competence is associated with 11%–14% increase in total pay. In model 3, we include both survey-based measures and continue to find that competence, but not attractiveness, is associated with higher executive pay.

These results suggest that the appearance of competence, rather than attractiveness, influences executive pay. Alternatively, it could be that the EIV problems arising from the noise inherent in survey assessments of attractiveness inhibit these models from identifying any significant relation between executive attractiveness and pay. The EIV concern, as illustrated earlier, is that the greater noise induced in the measurement of executive attractiveness can cause standard OLS regressions to bias its coefficient estimate. We can adjust for this bias if we know the reliability of the measurement. Unfortunately, we do not know the reliability with which our multiple survey participants accurately measure fundamental attractiveness. However, we can estimate the minimum reliability of the survey assessments of attractiveness required to make the variable useful. The  $R^2$  of a regression of survey assessments on all other variables, including pay, represents how well all other variables explain survey assessment variability. If the reliability of the survey measure of attractiveness is not at least at this level, then it does not provide any additional explanatory power for the regression and need not be included. In our case, the  $R^2$  from this regression is 0.2475, which therefore represents the lowest reasonable reliability for our survey measure of attractiveness. Thus, to account for the EIV concern we run EIV-adjusted OLS using a reliability ranging from 1.0 (standard OLS) to as low as 0.2475.<sup>11</sup> As we decrease the reliability of the survey

<sup>10</sup>We selected the set of controls used in Table 2 based on Graham et al. (2017) because the focus of this analysis is to replicate their earlier findings. In addition, we use a limited set of controls similar to Graham et al. (2017), because like them we have a relatively small sample size. In unreported results, including the additional controls yielded similar results though slightly weaker statistically.

<sup>11</sup>To account for the negative bias stemming from the measurement error of survey assessment of attractiveness, we need the reliability ratio. While we do not know the precise reliability for our survey assessments of attractiveness, we can estimate a lower bound of this reliability as described above. Given the lower bound of reliability we can estimate the errors-in-variables adjusted OLS estimates following Treiman (2009). The reliability ratio,  $r$ , is equal to  $(1 - \text{variance of the measurement noise} / \text{variance of observed survey assessment of attractiveness})$  or equivalently  $(\text{variance of true assessment of attractiveness} / \text{variance of observed survey assessment of attractiveness})$ .

TABLE 3  
Executive Compensation

Table 3 reports regression results of executive total compensation. The dependent variable is the  $\ln(\text{tdc1} + 1)$ , where  $\text{tdc1}$  is the executive's total annual compensation from Execucomp. Each model includes industry (Fama–French 49) and year fixed effects. All models are OLS regressions. Model 4 incorporates an errors-in-variables estimation that makes corresponding adjustments because the SURVEY\_ATTRACTIVENESS\_SCORE variable is measured with error. The measurement reliability is unknown, but we assume it to be at the lowest possible value. We estimate the assumed reliability to be just higher than the  $R^2$  from a regression of the Sum of Mask Deviations on all other variables, including total compensation. Since this regression had an  $R^2$  of 0.247, we assumed a measurement reliability of this variable to be 0.25. SUM\_OF\_MASK\_DEVIATIONS is the absolute sum of deviations from each of the 25 nodes on the Phi-Mask to the corresponding point on the executive's face. SURVEY\_ASSESSMENT\_OF\_ATTRACTIVENESS (COMPETENCE) is the average rating from Mechanical Turk workers who rate each executive's attractiveness (competence) on a scale of 1 to 10, with 10 being the most attractive (competent).  $p$ -values are in parentheses beneath each coefficient estimate. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	$\ln(\text{TOTAL COMPENSATION})$	$\ln(\text{TOTAL COMPENSATION})$	$\ln(\text{TOTAL COMPENSATION})$	$\ln(\text{TOTAL COMPENSATION})$	$\ln(\text{TOTAL COMPENSATION})$	$\ln(\text{TOTAL COMPENSATION})$
	OLS	OLS	OLS	EIV-OLS	OLS	OLS
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
SURVEY_ATTRACTIVENESS_SCORE	0.05 (0.53)		-0.061 (0.54)	8.405 (0.41)	0.033 (0.67)	-0.072 (0.46)
SURVEY_COMPETENCE_SCORE		0.141* (0.070)	0.18* (0.07)			0.17* (0.08)
SUM_OF_MASK_DEVIATIONS					-0.205*** (<0.01)	-0.201*** (<0.01)
$\ln(\text{SALES})$	0.42*** (<0.01)	0.424*** (<0.01)	0.423*** (<0.01)	0.617** (0.02)	0.417*** (<0.01)	0.42*** (<0.01)
OPERATING_PERFORMANCE	0.41*** (<0.01)	0.44*** (<0.01)	0.44*** (<0.01)	0.89 (0.13)	0.419*** (<0.01)	0.449*** (<0.01)
No. of obs.	203	203	203	203	203	203
Fixed Effects	Ind./Yr.	Ind./Yr.	Ind./Yr.	Ind./Yr.	Ind./Yr.	Ind./Yr.
Assumed reliability of Survey Score				0.25		
$R^2$	41.29%	42.33%	42.47%	66.20%	43.96%	45.03%

measure of attractiveness, its coefficient estimate becomes economically larger and positive. We report the results for the lowest level of reliability in model 4 but we still do not observe a significantly positive relation between survey-based measures of attractiveness and executive pay. Thus, either there is no statistically significant relation between attractiveness and pay amongst the executive ranks as there is among lower level workers (Hamermesh and Biddle (1994), Hamermesh (2011)) or we are unable to overcome the EIV problem inherent in survey assessments of attractiveness. To address this latter possibility, we next turn to our objective measure of executive attractiveness.

In model 5, we include the Sum of Mask Deviations and find this measure to be statistically and significantly associated with total compensation. A 1-standard-deviation decrease (more attractive) in mask deviations (162.4 points or about 2.25 inches) is associated with a 20.5% increase in executive compensation or about \$1.3 million higher than the average executive. This magnitude is similar to other studies that attempt to quantify the effects of difficult-to-measure CEO characteristics on pay. For example, it is similar to the 19% greater (\$0.85 million) pay for a generalist CEO relative to that of a specialist CEO documented in Custodio, Ferreira, and Matos (2013).<sup>12</sup> Falato, Li, and Milbourn (2015) document that an improvement in their credential measures of press coverage or in a career fast track is associated with approximately \$2.8 million higher pay. Similarly, an improvement in their college credential is associated with approximately \$1 million higher pay (about 20% of their sample mean). These results suggest that the objective mask-based measure better captures fundamental aspects of executive attractiveness with less error than a subjective survey assessment.

Next, in model 6, we include both the survey-based and the mask-based attractiveness measures. We continue to find no significant relation between the survey-based measure of attractiveness and compensation, whereas we do find a positive and significant coefficient estimate for the survey-based measure of competence and a negative and significant coefficient estimate for the Sum of Mask Deviations. Because both the survey-based measure of competence and the mask-based measure of attractiveness are both measures of executive appearance, it is difficult to attribute the association with greater CEO pay to either characteristic. For example, if both measures relate to aspects of appearance that affect CEO pay, such as personality traits stemming from the executive's genotype and reflected in the executive's appearance, then we cannot attribute the fact that appearance does affect CEO pay to either characteristic. In an effort to isolate the effect of

---

attractiveness/variance of observed survey assessments of attractiveness) (Fuller (1987)). In a simple linear regression, the OLS estimator,  $b$ , of  $\beta$  is given by  $S_{xy}/S_{xx}$  where  $S_{xy}$  and  $S_{xx}$  are the sample covariance and sample variance, respectively. However, as discussed earlier, when  $X$  is measured with error, the OLS estimator is biased downward because  $S_{xx}$  is larger than its true value due to the greater deviations introduced by the measurement errors. If we know the variance of the measurement noise, we can adjust this estimate to remove the bias. Thus, a consistent estimator is given by  $S_{xy}/(S_{xx} - \sigma_u^2)$ , where  $\sigma_u^2$  is the variance of the measurement noise (Kmenta (1971)). If we know the reliability we can estimate  $\sigma_u^2$ ,  $((1 - r) \times \text{variance of observed survey assessments of attractiveness})$  (Lockwood and McCaffrey (2020)), and thus obtain a consistent estimate of  $\beta$ . We estimate the variance-covariance matrix following Buonaccorsi (2010).

<sup>12</sup>The  $R^2$  results are also similar as various models are in the 40%–50% range.

attractiveness from the appearance of competence, in unreported results, we repeat this regression replacing the Sum of Mask Deviations with the residuals from a regression of the Sum of Mask Deviations on the Survey Competence Score and we find the results are unchanged. While we cannot rule out that the mask may be correlated with other personality traits that could be reflected in an executive's genotype, these unreported results suggest that the effect of attractiveness on CEO pay is distinct from that of the appearance of competence on CEO pay.

In summary, due to the EIV problem associated with survey assessments of executive attractiveness, we are unable to identify a significant association between survey-assessed executive attractiveness and pay. In contrast, our objective measure of attractiveness, which is uninhibited by EIV concerns, provides a strong link between executive attractiveness and compensation. Together, these results reveal that our objective measure of facial appearance provides additional explanatory power beyond that provided by survey assessments of either attractiveness or competence.

## B. Tournaments and Facial Attractiveness

In [Table 4](#), we use regression models to explore determinants of an executive being selected as CEO. Models 1–5 report results from probit model regressions while models 6–7 report results using an OLS regression (linear probability model), with industry fixed effects. The dependent variable is 1 if the executive is appointed as CEO and 0 otherwise. Since we have multiple executives per firm-year observation we incorporate robust standard errors clustered by firm. We control for board membership, compensation rank (Mobbs and Raheja (2012)), having the title of President (Naveen (2006)), ownership (excluding options), age, gender, and race.

In model 1, we find that, consistent with Naveen (2006) and Mobbs and Raheja (2012), board membership, compensation rank, and holding the title of President each significantly increase the likelihood of winning the tournament. These characteristics are important because they are also measures of human capital accumulated by executives throughout their careers. It is noteworthy that, in unreported analysis, these three variables alone explain over 27% of the variation in the likelihood of selection to be the next CEO. We also find that female executives are associated with a lower likelihood of becoming the next CEO and Asian executives have a positive association with becoming the next CEO.

Our main question is whether either the survey-based or mask-based measure of executive attractiveness provides any incremental influence in executive promotion to CEO. We begin with the survey-based measure of attractiveness and report the results in model 2. The survey assessment does not relate significantly to the likelihood of an executive's appointment as CEO. Given the purported benefits of attractiveness in the labor market, this result is puzzling. However, as noted previously, the noise in this measurement and the associated EIV concerns may prevent this measure from identifying an important significant relationship with the likelihood of an executive being promoted to CEO. In model 3, we include the mask-based measure of attractiveness, which does not have the same concerns. Here we find a significant relation between executive attractiveness (lower mask deviations) and the likelihood of being selected as CEO.

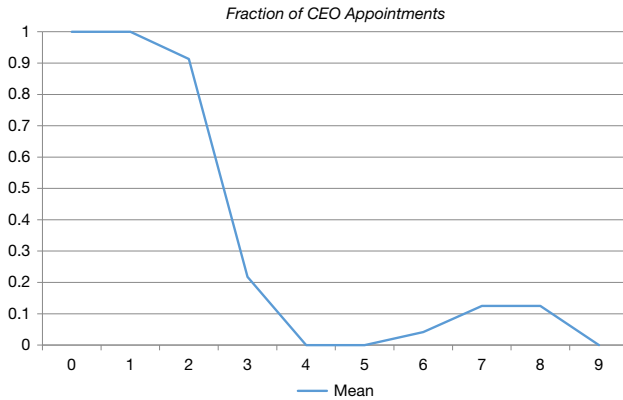
TABLE 4  
Do Looks Matter in CEO Selection?

Table 4 reports results from regression analyses of the likelihood of an executive being selected to be the next CEO. The dependent variable equals 0 if the board selects the executive to be the CEO in the following year and 0 otherwise. SUM\_OF\_MASK\_DEVIATIONS is the absolute sum of deviations from each of the 25 nodes on the Phi-Mask to the corresponding point on the executive's face. SURVEY\_ASSESSMENT\_OF\_ATTRACTIVENESS is the average rating from Mechanical Turk workers who rate each executive's attractiveness on a scale of 1 to 10, with 10 being the most attractive. The identified most likely CEO candidates are those who are younger than 65 and either on the board or ranked in the top 3 in pay. Models 1 to 5 are probit models. Models 6 and 7 are linear probability models and include Fama-French defined industry fixed effects. EXECUTIVE\_IS\_A\_DIRECTOR is an indicator variable that equals 0 if the executive is also on the board during the year. COMPENSATION\_RANK is the rank of the executive based on total compensation with 1 being the highest ranking (i.e., highest paid) executive. PRESIDENT is an indicator variable that equals 0 if the executive held the title of President in the prior year and 0 otherwise. OWNERSHIP is the percentage of shares outstanding owned by the executive, excluding options. AGE is the executive's age in the year the board selects the CEO. Standard errors are robust and clustered by firm. *p*-values are reported in parentheses beneath each coefficient estimate. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: Selected as CEO (1/0)									
				Exclude Lowest 2 Deciles of SUM_OF_MASK DEVIATIONS		Exclude Lowest 3 Deciles of SUM_OF_MASK DEVIATIONS		Exclude Lowest 2 Deciles of SMD		
	Prob(CEO = 1) 0.18			Prob(CEO = 1) 0.04		Prob(CEO = 1) 0.03		OLS	OLS	
	Model 1	Model 2	Model 3	<i>dy/dx</i>	Model 4	<i>dy/dx</i>	Model 5	<i>dy/dx</i>	Model 6	Model 7
SURVEY_ASSESSMENT_OF_ATTRACTIVENESS		-0.059 (0.58)	-0.137 (0.42)	-0.036	-0.11 (0.55)	-0.009	0.059 (0.77)	0.004	-0.024 (0.33)	-0.013 (0.64)
SUM_OF_MASK_DEVIATIONS			-1.621*** (<0.01)	-0.423	-0.8*** (<0.01)	-0.068	-0.691*** (<0.01)	-0.052	-0.268*** (<0.01)	-0.092*** (<0.01)
EXECUTIVE_IS_A_DIRECTOR	0.413*** (<0.01)	0.42*** (<0.01)	0.458** (0.02)	0.120	0.447** (0.04)	0.038	0.429 (0.12)	0.032	0.09*** (<0.01)	0.07** (0.02)
COMPENSATION_RANK	-0.741*** (<0.01)	-0.743*** (<0.01)	-0.715*** (<0.01)	-0.187	-0.525** (0.02)	-0.045	-0.33 (0.14)	-0.025	-0.111*** (<0.01)	-0.035 (0.29)
PRESIDENT	0.377*** (<0.01)	0.381*** (<0.01)	0.259** (0.04)	0.068	0.279* (0.06)	0.024	0.304* (0.06)	0.023	0.07** (0.01)	0.06* (0.08)
OWNERSHIP	0.079 (0.39)	0.078 (0.4)	0.22* (0.09)	0.057	0.16 (0.13)	0.014	0.13 (0.23)	0.010	0.02 (0.52)	0.01 (0.61)
AGE	-0.021 (0.83)	-0.046 (0.68)	-0.357 (0.12)	-0.093	-0.366 (0.1)	-0.031	-0.333 (0.25)	-0.025	-0.028 (0.36)	-0.053** (0.03)
FEMALE	-0.356*** (<0.01)	-0.34*** (<0.01)	-0.234 (0.16)	-0.061	-0.111 (0.44)	-0.009	0.03 (0.85)	0.002	-0.06*** (<0.01)	-0.02 (0.21)
ASIAN	0.342* (0.05)	0.335* (0.06)	0.343*** (<0.01)	0.090	0.258** (0.02)	0.022	0.258*** (<0.01)	0.019	0.07*** (<0.01)	0.06* (0.06)
AFRICAN_AMERICAN	0.047 (0.48)	0.048 (0.49)	0.225*** (<0.01)	0.059	0.195*** (<0.01)	0.017	0.248*** (<0.01)	0.018	0.03*** (<0.01)	0.04*** (<0.01)
CONSTANT	-0.593*** (<0.01)	-0.595*** (<0.01)	-0.865*** (<0.01)		-1.142*** (<0.01)		-1.22*** (<0.01)		1.38*** (<0.01)	0.18*** (<0.01)
No. of obs.	237	237	237		171		148		237	171
Pseudo- <i>R</i> <sup>2</sup> /Adjusted- <i>R</i> <sup>2</sup>	34.51%	34.59%	65.54%		31.36%		31.11%		56.39%	10.36%

FIGURE 3  
 Fraction of Executives Selected to Be the Next CEO by Phi-Mask Deviation Decile

Figure 3 shows the fraction of executives selected as CEO within each Phi-Mask deviation decile, with decile 9 being those with the greatest deviation, and decile 0 being those with the least deviation.



In the column to the right of the coefficient estimates of model 3, we report the marginal effects for our sample. A 1-standard-deviation decline (more attractive) in the Sum of Mask Deviations is associated with a 0.423 increase in the probability of the executive being selected as CEO. This economic impact is quite large as the unconditional probability of an executive being appointed CEO is 0.18, or about 1 of 5 executives. The effect of attractiveness is also large relative to other characteristics. For example, it represents an effect more than twice that of a 1-standard-deviation increase in compensation rank (lower number).

Introducing this measure of attractiveness to the model also notably increased the explanatory power from a pseudo- $R^2$  of 34.59% in model 2 to 65.54%. To explore this relation further, in Figure 3, we plot the fraction of executives appointed to the CEO position in each Sum of Mask Deviation decile. A 1-standard-deviation increase in attractiveness (decrease in Sum of Mask Deviations) is equivalent to a drop of two and a half to three deciles. Greater fractions of the executives in the more attractive (smaller) deciles become the next CEO. Interestingly, the most attractive executives, those in the lowest two deciles, all become CEOs. To be sure these most attractive executives are not driving our results, we repeat the analysis excluding the executives in the bottom two deciles (most attractive). We report the results in model 4, with the corresponding marginal effects in the column to the right of the coefficient estimates. In this smaller subsample, the unconditional probability of an executive being appointed CEO is only 0.04, or 1 in 20 executives. We find results similar to those for the full sample in model 3. The Sum of Mask Deviations is still negative and significant at the 1% level, though the magnitude of the marginal effect drops to a 0.068 increase in the probability of becoming CEO for a 1-standard-deviation increase in attractiveness. This economic effect is also closer to that of being a director or having a higher compensation rank, albeit still slightly stronger. The marginal effect for being a director indicates that if the executive is a director the likelihood of becoming CEO increases by 0.038, almost doubling the unconditional probability. The marginal effect for a 1-standard-deviation improvement in compensation ranking



is associated with an increase of 0.045 in the likelihood of being selected as CEO. The coefficient for holding the title of President suggests an economic effect of about half that of a compensation rank increase. In model 5, we further exclude the executives in the bottom three deciles of Sum of Mask Deviations and find very similar results to those in model 4. Thus, though the full sample results reported in model 3, may bias the coefficient estimate upward due to the concentration of many attractive executives who are appointed CEO, the results are not solely driven by this skewness. In summary, facial beauty has a significant statistical and economic impact on tournament succession outcomes.

Next, in model 6 we repeat model 3 using a linear probability model to be sure that the probit specification is not driving this finding (Caudill (1988)). We find consistent results with similar economic magnitudes. In model 7, we exclude the bottom two deciles of Sum of Mask Deviations and we continue to find similar results, but a lower economic impact from the Sum of Mask Deviations. As for the additional executive characteristics, we find that ownership relates positively to the likelihood of selection in one of the models. Executive age does not affect the likelihood of being selected CEO. The Female indicator is negative and significant in 3 of the models. Finally, the Asian and African American indicators are positive in all models. However, the African American indicator is only significant in four models whereas the Asian indicator is significant in all models.<sup>13,14</sup>

### C. Shareholder Reaction

Next, we consider how shareholders respond to the attractiveness of the newly appointed CEOs. In Table 5, we examine the determinants of the 3-day cumulative abnormal return (CAR) surrounding the announcement of the 100 newly appointed CEOs. Since we have only one observation per firm and several firms per industry, we use robust standard errors clustered by the Fama–French-49 industry classifications. In model 1, the only explanatory variable is the survey-based measure of the newly appointed CEO's attractiveness. We find that it is positive, but insignificantly related to shareholders' 3-day CAR. In model 2, we add the mask-based

<sup>13</sup>In unreported analysis, we also include a control for the intensity of the tournament incentives, measured as the compensation gap between the CEO and the median non-CEO executive. Greater tournament incentives could increase the value of attractiveness in winning the tournament. However, we find no evidence of tournament incentives strengthening or weakening the effect of executive attractiveness on succession likelihood.

<sup>14</sup>Kamiya, Kim, and Park (2018) find that the facial width to height ratio (fWHR) of the CEO predicts the riskiness of the firm. To control for this effect, we use the fWHR of the Phi Mask as our baseline. For the Phi Mask, the fWHR is the inverse of the golden ratio (0.618). For each executive in our sample we determine if their fWHR is greater or lesser than that for the Phi Mask. We measure the (x,y) deviations from the far left/right nodes and the top/bottom nodes of the executive's face and create an indicator variable that equals one if the fWHR is greater than that of the Phi Mask as a possible determinant for the likelihood of being selected as CEO. We find that 81% of the 237 executives in our sample have a fWHR greater than the Phi Mask and 70% of the 100 newly selected CEOs have a fWHR greater than the Phi Mask. However, we find that this variable is negatively related to the likelihood of being selected as CEO, but it is not quite significant ( $p$ -value = 0.11). The coefficient estimate for mask deviations remains significantly negative at the 1% level. This result is consistent with Kamiya et al. (2018), who find that executives with high fWHRs are less likely to be appointed as the CEO after the dismissals of CEOs with financial misconduct.

TABLE 5  
Shareholder Reaction to CEO Appointment Announcement

Table 5 reports regression results of the 3-day cumulative abnormal return (CAR) for the  $[-1,1]$  day window around the announcement of a new CEO appointment for the 100 CEO tournament winners. We estimate the market model using the value-weighted CRSP index as a proxy for the market returns over days  $[-201, -10]$ . We calculate the abnormal return for each day in the event window by subtracting the expected return (market model) from the actual return. SUM\_OF\_MASK\_DEVIATIONS is the sum of the deviations from the 25 nodes on the Phi-Mask to the actual point on the CEO's face in points as measured using Adobe Photoshop. NUMBER\_OF\_PRESS\_RELEASE\_STATEMENTS is the number of times the executive is quoted in a press release in the 3 years prior to being appointed as CEO. Standard errors are robust and clustered by the Fama-French 49 defined industries.  $p$ -values are in parentheses beneath each coefficient estimate. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels respectively.

	Dependent Variable: 3-Day CAR			
	Model 1	Model 2	Model 3	Model 4
SURVEY_ASSESSMENT_OF_ATTRACTIVENESS	0.00328 (0.45)	0.00358 (0.39)	0.00358 (0.38)	0.00578 (0.11)
SUM_OF_MASK_DEVIATIONS		-0.00582*** ( <i>&lt;0.01</i> )	-0.0058** (0.01)	-0.0051** (0.04)
NUMBER_OF_PRESS_RELEASE_STATEMENTS			0.0045* (0.08)	0.0044* (0.06)
NEW_CEO_AGE_AT_APPOINTMENT				0.0064** (0.04)
CONSTANT	-0.002 (0.67)	-0.0017 (0.66)	-0.002 (0.64)	-0.001 (0.72)
No. of obs.	100	100	100	100
$R^2$	0.88%	3.69%	5.86%	8.35%

measure of attractiveness and find a significant relation between this measure and shareholders' reactions. A decrease in the Sum of Mask Deviations (improvement in facial attractiveness) by 1-standard-deviation is associated with a 0.58 percentage point increase in shareholder reaction. This finding underscores the importance of having an objective measure less encumbered by noise that is able to detect evidence that the market indeed incorporates this qualitative information.

In model 3, we include the number of press releases the executive was responsible for in the 3 years prior to the CEO appointment announcement (Greene and Smith (2021)). The coefficient estimate is positive and significant at the 10% level. While this association is consistent with shareholders valuing the experience of publicly representing a firm, it is also consistent with shareholders being more familiar with these candidates and thus being more certain of their ability at their appointment announcement. If firms are more inclined to select more attractive candidates to speak to the press the shareholder reaction could be due to greater familiarity rather than attractiveness. However, in unreported results, we find there is no significant difference between the Sum of Mask Deviations for CEOs in the top and bottom quartiles of press releases. Furthermore, even after controlling for the number of press releases, the coefficient estimate for the Phi-Mask deviations remains essentially unchanged from that in model 2.

In model 4, we also control for the age of the newly selected CEO and find a positive and significant coefficient estimate, consistent with shareholders responding more favorably to the selection of older more experienced CEOs. Again, the masked-based measure of executive facial attractiveness remains statistically significant at the 5% level. In addition, the survey assessment of facial attractiveness, which is subject to age-related biases and greater noise in its measurement, is now

only marginally insignificant at traditional levels ( $p$ -value = 0.11). Nonetheless, the flaws with a survey-based measure of attractiveness that lead to the EIV concern continue to inhibit its ability to identify significant economic relations between attractiveness and shareholder value. In summary, the evidence in Table 5 reveals that shareholders attribute value to CEO attractiveness but this value is difficult to detect when using noisy survey-based assessments.

#### D. New CEO Chair Appointment

Next, we use our measures of CEO attractiveness to examine another important promotion, the appointment of the CEO to board chair. Prior literature in economics and psychology suggests that attractiveness may facilitate faster promotions as attractive individuals are effective persuaders (e.g., Aronson and Mills (1965)). It is not clear whether CEO-Chair duality is beneficial for the firm, but it is clear that CEOs benefit from the combined role by having greater control over the board. Thus, it is reasonable to expect that all CEOs would desire this duality.

In addition to our measures of attractiveness, we also control for firm size, abnormal stock performance over the CEO's appointment year, an indicator if the new CEO is over 55 years old, an indicator if the new CEO was previously the firm President and a measure of tournament strength. Larger firms can have greater information asymmetry between managers and the board, which makes combining the CEO and chair role informationally efficient (Brickley, Coles, and Jarrell (1997)). Conversely, boards of larger firms may require a longer "evaluation" period of new CEOs before appointing them as board chair and outgoing CEOs may retain the chair title for an extended period (Mobbs (2015)). First-year performance reflects the new CEO's talent and ability in the position. A poor showing can delay the joint CEO-chair appointment, whereas a strong first-year performance can expedite the appointment as CEOs prove their ability to the board. Older CEOs are likely to have more experience and, thus, be better prepared to handle the additional responsibilities associated with the chair position, whereas younger CEOs may require greater time to prove their ability to the board. Finally, because the type of succession can influence the time to chair appointment, we include two succession controls. First, since "passing the baton" successions can have expectations of when the new CEO also becoming the chair (Vancil (1987)), we control for whether the chosen CEO held the title of president prior to the succession. In addition, we control for the strength of the tournament incentives with the percentage gap between the prior CEO's compensation and the median compensation of the tournament competitors.<sup>15</sup>

In Table 6, we report results from Cox Proportional Hazard models where the dependent variable equals the number of days until the new CEO receives the chair position of the board. Since we only have one executive observation for each firm year but multiple firms per industry, we use robust standard errors clustered by industry. In model 1, we report the hazard ratios from the parameter estimates for a model using only the control variables. We find that stock performance and the

<sup>15</sup>When we include the percentage of independent directors on the board, this variable is not significant, but the Sum of Mask Deviations remains significant.

TABLE 6  
Likelihood of the CEO Being Awarded the Chair Near CEO Appointment

Table 6 reports the results of regression analyses on the likelihood of the new CEO being appointed as chair of the board using Cox Proportional models. The dependent variable, DAYS\_TO\_CHAIR, is the number of days until the CEO is appointed as Chair. SUM\_OF\_MASK\_DEVIATIONS is the absolute sum of deviations from each of the 25 nodes on the Phi-Mask to the corresponding point on the executive's face. ABNORMAL\_STOCK\_PERFORMANCE is the 12-month compounded return for the fiscal year less the 12-month compounded return of the market. PRESIDENT is an indicator variable that equals 0 if the executive held the title of President in the prior year and 0 otherwise. Standard errors are robust and clustered by industry. We report *p*-values in parentheses beneath each coefficient estimate. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels respectively.

	Dependent Variable: DAYS_TO_CHAIR		
	Model 1	Model 2	Model 3
SURVEY_ATTRACTIVENESS_SCORE		0.887 (0.3)	0.885 (0.35)
SUM_OF_MASK_DEVIATIONS			0.823* (0.08)
ln(SALES)	1.182 (0.14)	1.197 (0.12)	1.203 (0.1)
ABNORMAL_STOCK_PERFORMANCE	1.41*** ( <i>&lt;.01</i> )	1.456** (0.01)	1.54** (0.01)
OVER_55	1.363 (0.11)	1.29 (0.19)	1.305 (0.16)
PREVIOUS_PRESIDENT	0.8** (0.03)	0.808** (0.04)	0.79** (0.03)
TOURNAMENT_STRENGTH	1.002 (0.99)	0.98 (0.85)	0.995 (0.96)
No. of obs.	100	100	100
Prob > $\chi^2$	0.00	0.01	0.04

Previous President indicator both relate significantly to the likelihood of becoming chair. In Cox Proportional Hazard models, explanatory variables with coefficient estimate greater than 1 imply a greater “risk” or “hazard,” whereas coefficient estimates less than 1 imply a less likely “hazard.” In our context, the “hazard” is appointment to board chair. The coefficient estimate for stock returns, which is greater than 1, indicates that better stock returns hasten the “hazard” or increases the likelihood the CEO receives board chair sooner. Conversely, the coefficient estimate for the Previous President indicator, which is less than 1, indicates that CEOs holding the title of President before their CEO appointment face a lower “hazard” rate or a lower likelihood they will receive the board chair appointment on a given date. Firms implementing a “pass the baton” succession plan, where the heir receives the title of President before becoming the CEO, often allow the departing CEO to hold the board chair position for an extended period after the CEO transition occurs (Vancil (1987)).

In model 2, we include the survey-based measure of CEO attractiveness and find no evidence that attractiveness affects the hazard rate of the CEO reaching the board chair position. In model 3, we incorporate the mask-based objective measure of CEO attractiveness. Here we find evidence that attractiveness does affect the timeline of the CEO's appointment to the board chair. The significant coefficient estimate on Sum of Mask Deviations of 0.823 indicates that unattractive CEOs will have a significantly lower “hazard” rate for becoming board chair, implying that CEOs that are more attractive will have a significantly higher “hazard” rate for becoming board chair. Economically, a 1-standard-deviation increase in mask

deviations (less attractive) is associated with a 17.7% ( $1-0.823$ ) reduction in the expected hazard, in other words, a reduction in the likelihood of becoming chair on a given date since CEO appointment.<sup>16</sup> Lastly, we compute the Schoenfeld residuals to test the proportional-hazards assumption in this model. We find the  $p$ -values of all the coefficients and for the global test to be 0.15 or higher. Thus, there is no evidence to support a nonproportional hazard. In other words, there is no evidence that the hazard rate varies over time, which could limit the effectiveness of the model and thus our inferences.

In summary, using the objective mask-based measure of CEO attractiveness, we find that CEOs that are more attractive gain the board chair significantly faster. This outcome improves our understanding of the CEO-Chair duality decision and is consistent with attractiveness enabling the CEO to assert greater influence over the board of directors. However, these results are valid only when using the objective measure of attractiveness and not when using the survey-based measure.

## V. Conclusions

Survey assessments of physical characteristics such as facial attractiveness are subject to measurement error noise, which biases estimates of any association between the measured characteristic and related outcomes. We address this limitation of survey-based measures of facial attractiveness by introducing and utilizing a scientifically based measure of facial beauty to study the role of executive facial attractiveness. We use the Phi-Mask to measure the facial attractiveness of over 230 executives competing for CEO positions in 100 large publicly traded companies. The Phi-Mask is based on the esthetically pleasing aspects of the golden ratio, which has historical evidence in a variety of contexts that date back over 2,000 years and scientific research cited in our paper over 100. Since this scientifically based measure is free from the noise that affects subjective survey assessments (e.g., differences in age, culture, or familiarity between survey participants and the board of directors), we are better able to identify empirically the positive link between executive attractiveness and labor market outcomes.

Similar to the benefits attributed to attractiveness in the general population, using this new technology we are able to identify that more attractive executives receive higher compensation and are more likely to be promoted to CEO. Furthermore, CEOs that are more attractive are associated with a more favorable shareholder reaction to the news of their appointment and are associated with being appointed as chair of the board faster. Survey assessments of executive attractiveness, which are much noisier due to inherent biases, are not able to identify these same relations.

---

<sup>16</sup>Conversely, when we use the negative of the sum of mask deviations such that the more attractive CEOs have higher values (less negative), the coefficient estimate of the hazard ratio is 1.22. This indicates that a 1-standard-deviation increase in attractiveness is associated with a 22% higher hazard rate, or likelihood of being appointed to chair on a given date. Alternatively, when we repeat the analysis using an indicator for being in the bottom quartile of mask deviations (most attractive quartile) the coefficient estimate is 1.46, which suggests that a CEO in the top quartile of attractiveness has a 46% greater hazard rate compared to other CEOs.

In summary, using a more objective standard of attractiveness reveals that executive facial attractiveness can be a distinguishing trait that is important in executive labor market advancement. Moreover, the objectivity of the new method of measuring elemental aspects of beauty used in this study opens avenues for further research in other diverse fields of scientific inquiry including investments, economics, psychology, politics, and medical studies of reconstructive surgery.

## Supplementary Material

Supplementary Material for this article is available at <https://doi.org/10.1017/S0022109022000461>.

## References

- Adams, R.; M. Keloharju; and S. Knupfer. "Are CEOs Board Leaders? Lessons from Traits of a Million Individuals." *Journal of Financial Economics*, 130 (2018), 392–408.
- Aronson, E., and J. Mills. "Opinion Change as a Function of the Communicator's Attractiveness and Desire to Influence." *Journal of Personality and Social Psychology*, 1 (1965), 73–77.
- Bashour, M. "History and Current Concepts in the Analysis of Facial Attractiveness." *Plastic and Reconstructive Surgery*, 118 (2006a), 741–756.
- Bashour, M. "An Objective System for Measuring Facial Attractiveness." *Plastic and Reconstructive Surgery*, 118 (2006b), 757–774.
- Bertrand, M. "CEOs." *Annual Review of Economics*, 1 (2009), 1–29.
- Biddle, J., and D. Hamermesh. "Beauty, Productivity and Discrimination: Lawyers, Looks and Lucre." *Journal of Labor Economics*, 16 (1998), 172–201.
- Blankespoor, E.; B. E. Hendricks; and G. S. Miller. "Perceptions and Price: Evidence from CEO Presentations at IPO Roadshows." *Journal of Accounting Research*, 55 (2017), 275–327.
- Bognanno, M. "Corporate Tournaments." *Journal of Labor Economics*, 19 (2001), 290–315.
- Brickley, J.; J. Coles; and G. Jarrell. "Leadership Structure: Separating the CEO and Chairman of the Board." *Journal of Corporate Finance*, 3 (1997), 189–220.
- Brooks, A. W.; L. Huang; S. W. Kearney; and F. E. Murray. "Investors Prefer Entrepreneurial Ventures Pitched by Attractive Men." *Proceedings of the National Academy of Sciences*, 111 (2014), 4427–4431.
- Buonaccorsi, J. P. *Measurement Error: Models, Methods, and Applications*. Boca Raton, FL: CRC Press (2010).
- Caudill, S. "An Advantage of the Linear Probability Model over the Probit or Logit." *Oxford Bulletin of Economics and Statistics*, 50 (1988), 425–427.
- Cortina, J. M. "What is Coefficient Alpha? An Examination of Theory and Application." *Journal of Applied Psychology*, 78 (1993), 98–104.
- Cronbach, L. J. "Test 'Reliability': Its Meaning and Determination." *Psychometrika*, 12 (1947), 1–16.
- Custodio, C.; M. A. Ferreira; and P. Matos. "Generalist versus Specialist: Lifetime Work Experience and Chief Executive Officer Pay." *Journal of Financial Economics*, 108 (2013), 471–492.
- Davis, S. T., and J. C. Jahnke. "Unity and the Golden Section: Rules for Aesthetic Choice?" *American Journal of Psychology*, 104 (1991), 257–277.
- Dipboye, R. L.; R. D. Arvey; and D. E. Terpstra. "Sex and Physical Attractiveness of Raters and Applicants as Determinants of Resumé Evaluations." *Journal of Applied Psychology*, 62 (1977), 288–294.
- Dipboye, R. L.; H. L. Fromkin; and K. Wiback. "Relative Importance of Applicant Sex, Attractiveness, and Scholastic Standing in Evaluation of Job Applicant Resumes." *Journal of Applied Psychology*, 60 (1975), 39–43.
- Duarte, J.; S. Siegel; and L. Young. "Trust and Credit: The Role of Appearance in Peer-to-Peer Lending." *Review of Financial Studies*, 25 (2012), 2455–2483.
- Falato, A.; D. Li; and T. Milbourn. "Which Skills Matter in the Market for CEOs? Evidence from Pay for CEO Credentials." *Management Science*, 61 (2015), 2845–2869.
- Fechner, G. T. *Zur Experimentellen Aesthetik [Toward an experimental aesthetic]*. Leipzig: S. Hirzel (1871).



- Foo, P. W., and M. C. Clark. "Adult Age and Gender Differences in Perceptions of Facial Attractiveness: Beauty is in the Eye of the Older Beholder." *Journal of Genetic Psychology*, 172 (2011), 162–175.
- Fuller, W. A. *Measurement Error Models*. New York: Wiley (1987).
- Goldman, W., and P. Lewis. "Beautiful is Good: Evidence that the Physically Attractive are More Socially Skillful." *Journal of Experimental Social Psychology*, 13 (1977), 125–130.
- Graham, J. R.; C. R. Harvey; and M. Puri. "A Corporate Beauty Contest." *Management Science* 63 (2017), 3044–3056.
- Greene, W. *Econometric Analysis*. 2nd ed. New York, NY: Macmillan (1993).
- Greene, D., and J. Smith. "How Does Acquisition Experience Affect Managerial Career Outcomes?" *Journal of Financial and Quantitative Analysis*, 56 (2021), 1381–1407.
- Guise, B.; C. Pollans; and I. D. Turkat. "Effects of Physical Attractiveness on Perception of Social Skill." *Perceptual and Motor Skills*, 54 (1982), 1039–1042.
- Halford, J., and H. S. Hsu. "Beauty is Wealth: CEO Attractiveness and Firm Value." *Financial Review*, 55 (2020), 529–556.
- Hamermesh, D. S. *Beauty Pays: Why Attractive People are More Successful*. Princeton, NJ, Princeton University Press (2011)
- Hamermesh, D. S., and J. Biddle. "Beauty and the Labor Market." *American Economic Review*, 84 (1994), 1174–1194.
- Hatfield, E., and S. Sprecher. *Mirror, Mirror... The Importance of Looks in Everyday Life*. Albany, NY: State University of New York Press (1986).
- Heilman, M. E., and L. R. Saruwatari. "When Beauty is Beastly: The Effects of Appearance and Sex on Evaluations of Job Applicants for Managerial and Nonmanagerial jobs." *Organizational Behavior and Human Performance*, 23 (1979), 360–372.
- Hermalin, B., and M. Weisbach. "The Determinants of Board Composition." *Rand Journal of Economics*, 19 (1988), 589–606.
- Holland, E. "Marquardt's Phi Mask: Pitfalls of Relying on Fashion Models and the Golden Ratio to Describe a Beautiful Face." *Aesthetic Plastic Surgery*, 32 (2008), 200–208.
- Hsieh, T. S.; J. B. Kim; R. R. Wang; and Z. Wang. "Seeing is Believing? Executives' Facial Trustworthiness, Auditor Tenure and Audit Fees." *Journal of Accounting and Economics*, 69 (2020), 1–23.
- Kamiya, S.; Y.-H. Kim; and S. Park. "The Face of Risk: CEO Facial Masculinity and Firm Risk." *European Financial Management*, 25 (2018), 239–270.
- Kim, Y.-H. "Easy Facial Analysis using the Facial Golden Mask." *Journal of Craniofacial Surgery*, 18 (2007), 643–649.
- Kmenta, J. *Elements of Econometrics*. New York, NY: MacMillan (1971).
- Little, A. C.; B. C. Jones; and L. M. DeBruine. "Facial Attractiveness: Evolutionary Based Research." *Philosophical Transactions of the Royal Society B*, 366 (2011), 1638–1659.
- Lockwood, J. R., and D. F. McCaffrey. "Recommendations about Estimating Errors-in-Variables Regression in Stata." *Stata Journal*, 20 (2020), 116–130.
- Marquardt, S. R. "Dr. Stephen R. Marquardt on the Golden Decagon and Human Facial Beauty. Interview by Dr. Gottlieb." *Journal of Clinical Orthodontics*, 36 (2002), 339–347.
- Mayew, W.; C. Parsons; and M. Venkatachalam. "Voice Pitch and the Labor Market Success of Male Chief Executive Officers." *Evolution and Human Behavior*, 34 (2013), 243–248.
- Mayew, W., and M. Venkatachalam. "The Power of Voice: Managerial Affective States and Future Firm Performance." *Journal of Finance*, 67 (2012), 1–43.
- Mobbs, H. S. "Is an Outside Chair Always Better? The Role of Non-CEO Inside Chairs on Corporate Boards." *Financial Review*, 50 (2015), 547–574.
- Mobbs, H. S., and C. Raheja. "Internal Managerial Promotions: Insider Incentives and CEO Succession." *Journal of Corporate Finance*, 18 (2012), 1337–1353.
- Mobius, M. M., and T. S. Rosenblat. "Why Beauty Matters." *American Economic Review*, 96 (2006), 222–235.
- Murphy, K. "Corporate Performance and Managerial Remuneration: An Empirical Analysis." *Journal of Accounting and Economics*, 7 (1985), 11–42.
- Naveen, L. "Organizational Complexity and Succession Planning." *Journal of Financial and Quantitative Analysis*, 41 (2006), 661–683.
- Peskin, M., and F. Newell. "Familiarity Breeds Attraction: Effects of Exposure on the Attractiveness of Typical and Distinctive Faces." *Perception*, 33 (2004), 147–157.
- Ravina, E. "Love & Loans: The Effect of Beauty and Personal Characteristics in Credit Markets." Working Paper, Columbia University (2019).
- Rhodes, G.; J. Halberstadt; and G. Brajkovich. "Generalization of Mere Exposure Effects to Averaged Composite Faces." *Social Cognition*, 19 (2001), 57–70.
- Schmid, K.; D. Marx; and A. Samal. "Computation of a Face Attractiveness Index Based on Neoclassical Canons, Symmetry, and Golden Ratios." *Pattern Recognition*, 41 (2008), 2710–2717.



- Stewart, N.; C. Ungemach; A. J. L. Harris; D. M. Bartels; B. R. Newell; G. Paolacci; and J. Chandler. "The Average Laboratory Samples a Population of 7,300 Amazon Mechanical Turk Workers." *Judgement and Decision Making*, 10 (2015), 479–491.
- Todorov, A.; A. N. Mandisodza; A. Goren; and C. C. Hall. "Inferences of Competence from Faces Predict Election Outcomes." *Science*, 308 (2005), 1623–1626.
- Treiman, D. J. *Quantitative Data Analysis: Doing Social Research to Test Ideas*. San Francisco: Jossey-Bass (2009).
- Vancil, R. *Passing the Baton: Managing the Process of CEO Succession*. Boston: Harvard Business School Press, (1987).
- Warner, R., and D. Sugarman. "Attributions of Personality based on Physical Appearance, Speech and Handwriting." *Journal of Personality and Social Psychology*, 50 (1986), 792–799.
- Widgery, R. N., and B. Webster. "The Effects of Physical Attractiveness upon Perceived Initial Credibility." *Michigan Speech Journal*, 4 (1965), 9–15.
- Zajonc, R. B. "Attitudinal Effects of Mere Exposure." *Journal of Personality and Social Psychology Monograph Supplement*, 9 (1968), 1–27.