When does charisma matter for top-level leaders?

Effect of attributional ambiguity

PHILIPPE JACQUART
EMLYON Business School, France
jacquart@em-lyon.com

JOHN ANTONAKIS
Faculty of Business and Economics
University of Lausanne, Switzerland
(HEC Lausanne)
john.antonakis@unil.ch

in press

Academy of Management Journal

The first version of this manuscript was submitted to this journal on 18 September 2012 (i.e., 7 weeks before the U.S. presidential election took place) and included the forecast that President Obama would be reelected. We are grateful for the assistance of Inna Karapetiants and Yannick Rohrbasser in coding the acceptance speeches and for the help of Molly Molloy and Samson Zadmehran in collecting these speeches. We also appreciate helpful comments received from Scott J. Armstrong, Jean-Philippe Bonardi, Romain Brixtel, Fabrizio Butera, Marius Brulhart, Marina Fiori, Deanne Den Hartog, Joerg Dietz, Philippe Monin, Dominic Rohner, Nicolas Roulin, Mathias Thoenig, Niels Van Quaquebeke, and Christian Zehnder on earlier versions of this manuscript. We are also very grateful to the associate editor, Raymond Sparrowe for his helpful editorial guidance and also appreciate the constructive comments of the reviewers.

When does charisma matter for top-level leaders?

Effect of attributional ambiguity

One stream of leadership theory suggests leaders are evaluated via *inferential* observer processes that compare the fit of the target to a prototype of an ideal (charismatic) leader. *Attributional* theories of leadership suggest that evaluations depend on knowledge of past organizational performance, which is attributed to the leader's skills. We develop a novel theory showing how inferential and attributional processes simultaneously explain top-level leader evaluation and ultimately leader retention and selection. We argue that observers will mostly rely on attributional mechanisms when performance signals clearly indicate good or poor performance outcomes. However, under conditions of attributional ambiguity (i.e., when performance signals are unclear), observers will mostly rely on inferential processes. In Study 1 we tested our theory in an unconventional context—the U.S. presidential election—and found that the two processes, due to the leader's charisma and country economic performance, interact in predicting whether a leader is selected. Using a business context and an experimental design, in Study 2 we show that CEO charisma and firm performance interact in predicting leader retention, confirming the results we found in Study 1. Our results suggest that this phenomenon is quite general and can apply to various performance domains.

Keywords: Inferences; Attributions; Performance signals; Charisma; Econometric models; Attributional ambiguity; Heuristics; Leadership; Leader selection; U.S. Presidents; Obama; Romney; Elections; Forecasting; Upper echelons.

Broadly defined, organizations—whether entities or institutions like firms or nation states (Hodgson, 2006)—have leaders at their helms. What psychological processes explain how these top-level leaders are evaluated and selected? We advance a novel theory by combining two psychological explanations: Inferential and attributional. Top-level leaders are selected because of (a) how leader-like they seem, and (b) the performance of their organization, which is causally attributed to them. Via these two signaling channels, selectors have some stereotypical cues about future organizational performance. Selectors, particularly in limited information conditions, intuitively believe that the more an individual is leader-like and/or the better the performance of their organization, the more likely the leader will engender good future performance.

The two psychological processes are well explained by current leadership theories. From an inferential perspective, the degree to which a leader's characteristics resemble a prototypical leader imbues the target with leader-like qualities (Lord, Foti, & De Vader, 1984). One key quality, charisma—an attribute idealized across cultures (Den Hartog, House, Hanges, & Ruiz-Quintanilla, 1999)—strongly predicts leader prototypicality per se (Antonakis, Fenley, & Liechti, 2011) and objective outcomes (Lowe, Kroeck, & Sivasubramaniam, 1996). Another explanatory route is attribution theory: It suggests that organizational performance—a presumed indicator of the leader's competence—determines how leaders will be evaluated (Calder, 1977; Lord, Binning, Rush, & Thomas, 1978; Meindl & Ehrlich, 1987). Thus, good performance is thought to be an outcome of effective leadership. This tendency is so deeply rooted that such attributions are made even when they are not warranted (Weber, Camerer, Rottenstreich, & Knez, 2001).

Thus, leaders are positively evaluated if they act leader-like and obtain good performance. A leader doing a "good" job will be re-appointed; a leader doing a "poor" job will be replaced (Hilger, Mankel, & Richter, 2013). Of course, observers, particularly distant ones, cannot know

everything about the leader, if organizational performance is good or poor, or whether the leader was responsible for the performance. Observers lack full information; yet they have to make a heuristic decision under conditions of uncertainty (cf. Tversky & Kahneman, 1974).

The two psychological processes reflect different mechanisms and profoundly different approaches to understanding leadership. The first, which is leader-centric, argues that leaders, particularly charismatic ones are highly influential and affect organizational outcomes. The second, a follower-centric perspective, states that leadership is a social construction; observers use organizational outcomes, whether or not caused by the leader, to decide if a leader is effective (Meindl, 1990). Current theory has not reconciled these two processes (cf. Day & Lord, 1988; Meindl & Ehrlich, 1987). Do both processes matter for top-level leadership? Might the processes "work together" in explaining leader selection? We seek to answer these questions by advancing a hybrid theory of leader evaluation and consequently selection. We focus on top-level leaders and how they are evaluated by selectors, because of the impact such leaders can have on collectives (House, Spangler, & Woycke, 1991; Jones & Olken, 2005).

Our contribution is threefold. First, we will reconcile (a) *inferences* made about leaders' characteristics, providing indications of how effective they may be, with (b) *attributions* about the performance of the leader's organization, providing indications of how effective the leaders have been (cf. Erickson & Krull, 1999)[1]. Although inferences or attributions about target leaders have been examined extensively in laboratory settings (e.g., Awamleh & Gardner, 1999; Howell & Frost, 1989; Rush, Thomas, & Lord, 1977), they have not been studied simultaneously or in terms of how leaders are selected in consequential roles. The CEO selection and succession

[1] Note, we use the terms "inferences" and "attributions" as defined by Erickson and Krull (1999) who have argued that the terms refer to distinct decision-making processes: Inferences concern understanding the nature of an individual whereas attributions concern the cause of an outcome (see also Lord and Maher, 1994, who have used the terms recognition- and inference-based processes respectively).

literature, has focused on factors like firm performance and demographic characteristics (Datta & Guthrie, 1994, 1997), board composition and ownership (Boeker & Goodstein, 1993), board power (Zajac & Westphal, 1996), or on understanding how new CEOs affect strategic change in succession events (Hutzschenreuter, Kleindienst, & Greger, 2012). Both laboratory and field studies, however, have not modeled how inferential and attributional processes simultaneously affect selection decisions when attributional information is unclear.

Second, apart from bridging two disparate theories, our key contribution is on the moderating effect of clarity of attributional information. When performance is clearly good (or bad) leader evaluations will ultimately be positive (or negative) largely irrespective of the leader's charisma. Simply put leader charisma matters little if outcomes attributed to the leader send clear signals. However, in conditions of attributional ambiguity, where performance signals are inconclusive, this uncertainty will spur selectors to decide on the basis of inferential processes. In other words, the charismatic leader will shine when attributional information is nebulous.

Third, in addition to testing our theory in a more conventional setting—the decision to retain a CEO—we use an unconventional context to demonstrate the general nature of our theory: The selection of the U.S. president. This extraordinary setting provides a controlled environment allowing for a direct test of our theory, as we explain in detail later.

With respect to our contributions, it is important to briefly differentiate attributional ambiguity from Weberian (1947) notions of crisis, which can refer to: (a) bad organizational performance, or (b) times of turbulence or environmental volatility (Waldman, Ramirez, House, & Puranam, 2001) and this irrespective of organizational performance signals. Our focus is on the clarity of performance signals, which concerns the absence of a clear negative or positive organizational performance signal and does not necessarily reflect environmental volatility.

Charisma has been shown to matter most for firm performance in conditions of perceived environmental turbulence (Waldman et al., 2001); thus, the charismatic leader is seen as a savior for an organization operating in a risky environment. This phenomenon is different from what our theory explains. Our model suggests that charisma will matter most for leader selection when organizational performance is ambiguous, and this irrespective of environment volatility.

In the next section, we build a general theory of leader selection—independent of the context in which we test it—by discussing how leaders are evaluated from a social-cognitive perspective. We focus on evaluations resulting from attributions on performance cues and from inferences based on prototypes of ideal leaders. As concerns the latter, we argue that charisma is a key characteristic of such prototypes and that charisma will likely matter a great deal in situations of attributional ambiguity. Then we highlight how our theory explains leader evaluation (i.e., retention and/or selection) in a political (Study 1) and a business (Study 2) context.

EVALUATION OF LEADERS

Attributions of Leadership and Performance Cues

Building on Jones and Davis' (1965) correspondence inference theory, Calder (1977) argued that attributions begin with observations of (a) the leader's behaviors and (b) their outcomes (e.g., organizational performance). The outcome is attributed to the leader if behaviors are distinctive and not situational dependent. Thus, attributions are likely made when both behaviors and their effects are observable; however, departing from Jones-Davis (1965), Calder (1977) suggests that "individuals may also rely heavily on knowledge about effects which are associated with a person, even though the requisite behavior for these effects has not been observed" (p. 197).

Theoretically, observers reason in a probabilistic way; they observe an event and attribute a cause to it if the event is representative of the cause (Tversky & Kahneman, 1974). For example,

organizational performance is representative of the leader's competence; good organizational performance is attributed to effective leadership (Binning, Zaba, & Whattam, 1986; Calder, 1977; Lord & Maher, 1994). Observing unambiguous effects is sufficient to make attributional judgments (Calder, 1977), and this in a heuristic way. In the case of close leaders (information rich situation), judgments are probably made in a more controlled way especially when the context of the judgment is seen as important (cf. Martinko, Harvey, & Douglas, 2007).

This attribution phenomenon is general. In a study of U.S. oil companies, CEO compensation increased following rises in firm profit, even though profits depended oil price fluctuations, which are mostly exogenous to the actions of any single CEO (Bertrand & Mullainathan, 2001). Business press articles about a U.S. airline deconstructed and reconstructed the CEO's image to match the evolution of the firm's performance (Chen & Meindl, 1991). These attributions apply too in the selection of CEOs where selectors scrutinize the organizational performance of the presiding CEO to see whether he or she should remain in office. If organizational performance is below par, outside candidates will be considered (Hilger et al., 2013), who in turn will have the performance of their companies scrutinized (Khurana, 2002). This process parallels what occurs in professional sports, where coaches are often fired following prolonged poor team performance.

Interestingly, organizational performance is attributed to leaders even when performance is known to depend on exogenous factors (Weber et al., 2001). Thus, some scholars (Meindl & Ehrlich, 1987) state that leadership might simply be socially constructed and does not matter for organizational outcomes. Of course, although performance signals bias leaders evaluations, leaders still impact organizational performance (Jones & Olken, 2005; Lowe et al., 1996).

In summary, the extent to which organizational performance is positive (or negative) will determine whether selectors will be positively (or negatively) predisposed to the leader; this

evaluation turn affect whether the leader is reselected. Thus, an incumbent leader will only be retained if performance has been good (Hilger et al., 2013).

Hypothesis 1: Organizational performance signals will predict leader selection

Note, we use the term "selection" in a broad sense to include leader selection and reselection (for incumbents) in cases of tournaments or retention at the helm of an organization or institution.

Inferences of Leadership and Person Perception

Similar to the intuitive attribution processes described above, individuals also judge leaders by relying on heuristics. The foundations of these inferential judgments are schemas—cognitive knowledge structures representing a concept along with its attributes and of the relations between these attributes and other concepts (Fiske, 1995). Schemas are developed with repeated exposure to common attributes that become indicative of a prototype. Once triggered schemas require little mental effort and convey configural information quickly (Fiske, 1995). Hence, individuals use slivers of information representative of a prototype to classify a target (cf. Tversky & Kahneman, 1974). Only a few indicators (e.g., of charisma) are needed to be classified under a particular label (e.g., charismatic). Once classified, other indicators of the prototype, even though not observed directly, will be associated with the target (Cantor & Mischel, 1977). This process makes for a stable memory structure about the target on the particular label, directs attention, guides information encoding and memory, and influences judgments and attitudes (Fiske, 1995).

Hence, individuals have implicit prototypes of leaders (Lord et al., 1984). For leadership at the upper echelons where there is a lack of information about what leaders are like, observers perceive leaders in idealized ways and are susceptible to classifying the leader following leader's image building efforts (Antonakis & Atwater, 2002). Charisma is an important element used in leader image building; from a dramaturgical perspective, leaders construct their charisma via

impression management techniques and by packaging and communicating their message to the target audience in an attractive way (Gardner & Avolio, 1998). Such leaders are especially good communicators and use framing and scripting techniques to project vision (Gardner & Avolio, 1998); these leaders typify what followers expect of them (Hogg, 2001), as we discuss next.

The Importance of Charisma for the Leadership Prototype: Charisma is "symbolic leader influence rooted in emotional and ideological foundations" (Antonakis et al., 2011: 376). Such leaders are exceptionally expressive and inspiring (Gardner & Avolio, 1998) and "manifest and symbolize desired collective values" (Shamir, 1995: 40). They typify the prototypicality that followers seek; they are liked and influence followers because followers identify with them (Hogg, 2001). These leaders are able to articulate "an ideological vision . . . [that engenders] a sense of identity with the collectivity" (Shamir, House, & Arthur, 1993: 585). Charismatic leaders make extensive use of articulation and impression management skills (Conger & Kanungo, 1987). These leaders are persuasive and use image-building techniques to imbue themselves with charisma (House, 1977); the use of rhetorical strategies, is the key that "shapes the charismatic relationship" (Gardner & Avolio, 1998: 42). Charismatic leaders affect their followers through the message they deliver (Shamir et al., 1993), which stems "from nonverbal and verbal influencing tactics that reify the leader's vision" (Antonakis et al., 2011: 376).

Charismatic leaders use specific strategies in terms of what they say and how they say it. To create emotional links with their followers charismatic leaders state their moral conviction, sort wrong from right, communicate high and ambitious goals, as well as confidence that these can be achieved; they do this using rich but simple descriptions that trigger a vivid vision (Antonakis et al., 2011; Antonakis & House, 2002; House, 1977; House & Shamir, 1993; Shamir, Arthur, & House, 1994; Shamir et al., 1993). Doing so requires the leader to use rhetorical techniques

including metaphors, contrasts, lists, stories, and so forth (Den Hartog & Verburg, 1997; Frese, Beimel, & Schoenborn, 2003; House, 1977; Shamir et al., 1994), termed "charismatic leadership tactics," which predict leader emergence and prototypicality (Antonakis et al., 2011).

For example, the use of metaphors by U.S. presidents is correlated with ratings of charisma (Mio, Riggio, Levin, & Reese, 2005) and with ratings of their greatness by historians (Emrich, Brower, Feldman, & Garland, 2001). In experimental settings, researchers have found that perceptions of leader charisma and effectiveness are strongly influenced by the leader displaying a strong delivery style—both in terms of vocal fluency and of non-verbal behaviors (Awamleh & Gardner, 1999). These charismatic tactics can be experimentally manipulated and have strong effects on many outcomes (e.g., trust in the leader, affect for the leader, Antonakis et al., 2011).

Thus, we expect that the extent to which a leader is perceived as charismatic will affect the extent to which this leader is perceived as leader-like and selected, because observers have well developed schemas that suggests that charismatic leaders are highly effective.

Hypothesis 2: Charisma will predict leader selection.

The Ambiguity of Performance Signals and the Rise of the Charismatic Leader: Both Hypothesis 1 and 2 depend on two information signals: Performance (i.e., outcomes) and behavioral (i.e., charismatic). We believe that these two signals interact with each other. The clearer the performance signal, whether positive or negative, the less likely charisma will matter for leader evaluation and selection. Charisma will matter less because selectors have what they believe is concrete evidence of the leader's competence (or incompetence)

Theoretically, strong performance signals are a "litmus test"; if the test is conclusive, other factors will not matter much in the selection decision. Whether the leader is charismatic or not is not of issue; the only thing that matters is if the leader has "proven" to be successful (or

unsuccessful) and deserves appointment (or dismissal). However, in conditions of attributional ambiguity (i.e., when the signal is not clearly positive or negative), selectors have to decide how to weight performance signals and charismatic leadership signals. It is in these situations, where the litmus test is inconclusive, that charisma will matter much. As mentioned by Calder (1977) situations where clear performance signals are unavailable are problematic for making attributions. Theoretically, in such situations selectors will look to the character of the leader, because they believe, intuitively, that charismatic leaders usually engender positive outcome. Thus, observers will use this information to make a judgment regarding how good the leader probably is, or with time will be. In this way, judgments of selectors will be influenced by the extent to which the leader is charismatic and this inferential process is heavily weighted in decision-making when performance signals are fuzzy.

Our theory of attributional ambiguity diverts from theories of crisis. According to Max Weber (1947) crisis situations engender psychological distress and followers seek a leader who can reassure them and allay their fears; in this way "charismatic leadership has a salvationistic or messianic quality" (Kets de Vries, 1988: 238). However, crisis is not a needed antecedent of charismatic leadership (Shamir & Howell, 1999). Also, a crisis can be either indicative of (a) a clearly negative signal (organizational performance is bad), or (b) a time of economic or political volatility, or an environment that is dynamic, risky, and stressful (Tosi, Misangyi, Fanelli, Waldman, & Yammarino, 2004; Waldman et al., 2001). These two descriptions of crisis are not isomorphic; the former is triggered by a negative performance signal and the latter by turbulence (which may be accompanied by clear, usually negative, or ambiguous performance signals).

Key to understanding how we differ from Weberian ideas is the role of performance signals, which will clarify *who* will be selected in a time of crisis. Contrary to Weberian notions, our

theory suggests that crisis—a clear signal of bad organizational performance—is detrimental to a charismatic leader if the leader is thought to be causally responsible for the crisis. Thus, the leader will be replaced by another leader. However, a charismatic leader will be selected in conditions of attributional ambiguity. Although we agree that crisis, in some situations, may ease the passage for a charismatic leader, our theory focuses on what will occur when organizational performance signals are clear or ambiguous.

Hypothesis 3: The effect of charisma on leader selection will be moderated by the clarity of performance signals such that the effect of charisma (a) will be positive and significant when performance signals are ambiguous and (b) non-significant when performance signals are clear.

TOP-LEVEL LEADERSHIP IN THE CORNER AND THE OVAL OFFICE

Via the inferential-attribution framework, our goal is to provide an integrative theoretical account and empirical test of how top-level leaders are selected. We expect our theory to operate in and be bounded by conditions of high leader distance, where there is limited information on the leader because the leader is physically distant from the followers, because of high status differences, and/or simply because of infrequent contact between the leader and followers (Antonakis & Atwater, 2002). We chose two such contexts in which to test out theory: (a) the U.S. presidency and (b) a conventional CEO setting, using field and laboratory data respectively.

Both political and business leadership requires social influence and organizational skills, which are inherent characteristics of management; in addition, both contexts reflect leadership at top hierarchical levels. However, at the outset they seem to be qualitatively dissimilar and entail different dynamics with respect to leader selection or retention. Thus, there is an important question to address: Is leadership in the U.S. presidential context relevant to management theory?

There are specificities, explained in the next section, which can be factored-out from this context (i.e., modeling incumbency and political party affiliation). Therefore, we highlight differences and similarities between these contexts, and explain how it is precisely because of these differences that we can ensure a clean test of our theory.

The major difference between the leader selection process is that unlike the U.S. presidency, CEOs are not chosen through an explicit tournament (i.e., election); of course, when more than one candidate is being considered it becomes a tournament (Connelly, Tihanyi, Crook, & Gangloff, 2014). Next, even if in a tournament, CEOs are not voted in by far-removed evaluators. Also, in political settings selectors (i.e., voters) are generally not held accountable to anyone; however, in business settings, selectors (i.e., boards and search committees) are accountable to a number of stakeholders including investors, and selectors may have other preferences or obligations that may constrain their selection decision. Finally, CEOs are not obliged to campaign for reelection at set times during their tenure. Succession events in business settings are usually planned or triggered because of bad organizational performance or other factors.

Despite these differences we argue that selectors will, at a basic psychological level, still use the same broad decision-making mechanisms we highlight in our model; that is, selectors care about whether the leader is leader-like (charismatic) and whether the leader (or his/her political party) has presided over positive performance. Both sets of selectors have to make probabilistic judgments in the absence of full behavioral or attributional information. Even though it is hard to know precisely if the CEO or the U.S. president is fully responsible for organizational outcomes the leaders are nonetheless seen as responsible for them.

We acknowledge a big difference regarding the "accountability" of selectors. Boards of directors are affected by personal and professional obligations from sitting on multiple boards,

having common social networks, or from other factors. However, we think it reasonable to believe that boards of directors also seek to select a competent leader who acts leader-like and has a good track record; thus directors also reason via the inferential and attributional paths. The context of the U.S. presidency does not involve such accountability dynamics—they are held constant in this context—thus allowing for an unconfounded test of the effects of our theory. For instance, factors such as seniority or one's position within a given business network matter; however, these factors and others cannot be easily controlled for in a CEO context and gaining direct access to board members for a selection study is difficult. In addition, board members might not be fully transparent about their preferences and obligations for the selection decisions; and, indirect measures would require uncovering factors that may not leave any archival traces. Thus, both contexts are fair game for our model, and concern the *initial decision* of whether an incumbent leader (or in the case of a second-term president, his party) will be re-selected.

In addition, the presidential election is exogenously determined; every four years selectors must judge the incumbent's (or the incumbent's party) economic record and then decide whether the leader (or the incumbent's party) should be reappointed. Because the event trigger is exogenous (i.e., it follows a strict cycle), there is no possibility that unexplained variability in the trigger is correlated with unexplained variability in economic performance factors; therefore, the problem of endogeneity is avoided. Hence, data from this context will allow for consistent estimation of the attributional and inferential factors that predict leader selection success. In addition, the "organization" is held constant and we have the same performance indicators available over time for the organization (the U.S. economy). Such a research setting has an advantage over observing two (or more) candidates each associated with prior organizational performance involving distinct organizations, making a direct comparison of difficult to model.

Finally, the trigger to CEO succession usually follows a script that has strong parallels to political leadership: It begins with an observable decline in performance for which the current CEO is blamed and ousted (Khurana, 2002); then comes the search for "an individual who has served as a CEO or president of at a *high-performing and well-regarded company*," who is ideally charismatic too (Khurana, 2002: 20, italics added). Thus, although not a tournament, the incumbent is jettisoned on the basis of bad organizational performance and the savior is sought as a function of performance- and behavioral-based signals. This process has strong similitudes to leader selection in political contexts.

To conclude, the common denominators that the two contexts share are amenable for testing our hypotheses. Both CEOs and political leaders are selected by individuals that use heuristic processes in limited information conditions. Both sets of leaders are conscious of their image because they are closely scrutinized by stakeholders: (a) U.S. presidents by voters, elected politicians in both houses, other officials, and administrators; and (b) CEOs by board members, investors, employees, and customers. Like Presidents, CEOs need to be concerned about being reappointed; although this concern is not cyclical, it will be prompted if organizational performance is below par. The initial reappointment assessment of an incumbent leader is influenced both by whether the leader is charismatic and by prior performance outcomes associated with the incumbent; when performance signals are ambiguous selectors will base their decision on how charismatic the leader is.

STUDY 1

We used archival data to model the outcomes of U.S. presidential elections from 1916 to 2008. We combined country level performance and incumbency data from a well-established econometric model with objectively measured indicators of candidates' charisma. The setting is

such that voters lack full information on the actual competence and dispositions of candidates and all the important economic-level factors associated with the candidates and their parties. Yet, they must use all available information when they cast their vote (Kelley & Mirer, 1974).

We will model the decisions of selectors, which theoretically stem from inferential and attribution processes via the share of popular vote received by the candidates. That is, we assume that evaluations due to inferential and attribution processes reflect in voter choices. We model only contenders from the Democratic and Republican parties (we will exclude third party candidates from the model because they usually receive a very small percentage of the total vote).

Modeling the Specificities of the U.S. Presidency to Test the Theory

From an attribution perspective and parallel to what we have theorized, econometric models of voting assume that presidential candidates are evaluated on the state of the economy (Lewis-Beck & Stegmaier, 2000). If the economy is healthy (e.g., high GDP growth, low inflation, which impact the labor market), electors reward the incumbent *or his or her party*, in the case of an incumbent not running again. If the state of economy is poor, voters punish the incumbent, *or his or her party*, and vote for the challenger. Thus, given party affiliation, term limits, or other factors, which can make for a choice between two non-incumbents, non-incumbents will be "made guilty" or "haloed" by complicity (i.e., party affiliation). As a striking example, and using only country-level economic data and incumbency data from 1916-2008, one of the most well-established econometric model, the Fair's presidential model (2009, 2010) explains 90.1% of the variance of vote outcomes (i.e., the two-party vote share considering only the Democratic and Republican contenders). This result provides strong support to the proposition that attributions of leadership can be made solely from observing the effects of leadership (Calder, 1977).

As for incumbency per se, the Fair (1978, 2009, 2010) model suggests that its effect cuts both ways. An incumbent has an advantage over the challenger because of more access to "airtime," and is more familiar, and thus more recognizable (Gaissmaier & Marewski, 2011). However, the longer a particular party has been in power, the more weary voters will be of this party: The party has "overstayed its welcome." In other words, incumbency status will predict (a) a higher share of the popular vote for the incumbent nominee and (b) a lower share of the popular vote for the incumbent party to the extent that the nominee's party has been in power for a long period of time. Thus, given that that we can partial-out the idiosyncrasies of the context, we are able to leverage our context and sample in service of testing our theory (rather than forcefully testing our theory on ill-suited context and sample).

Of course, partisanship matters a great deal too (Bartels, 2000) and most voters will consistently vote in line with their party affiliation. Consequently, galvanizing the party base and appealing to "swing voters"—those who do not identify strongly with a party—also matters (Mayer, 2008). Swing voters will, of course, also vote depending on the state of the economy; in addition, in situations of attributional ambiguity, which the econometric model ignores, we surmise that voters will compare candidates on charisma. In these situations, those who do not have strong party identification will be unsure whether to reward or blame the incumbent on the basis of country-level economic performance and will be more likely to consider questions like: "who of the two is likely to have the better leadership skills?" "who'd make a better commander in chief?" or "who do I like more?" Our theory suggests that in such situations (of attributional ambiguity), weakly or unaffiliated voters will be swayed by the more charismatic candidate.

To recap, when performance signals are very clear, voters will not care much about the incumbent's or the challenger's charisma; they will simply reward or punish the incumbent (or

the incumbent's party if the incumbent is not running again) on the basis of the economic fundamentals. Incumbency provides an advantage for sitting president only if the president's party has not been in power in the preceding election. Most importantly, charisma will matter only when the signals about the state of the economy are ambiguous. In these situations the contender who is more charismatic—whether challenger or incumbent—will win.

Data and Sample

We built our model on the most recent Fair (2009) econometric model. This model captures the attributional effects in which we are interested—while modeling incumbency—which we extend by adding the inferential (charisma) effects. Thus, the key variables in our full model are (a) share of votes received (indicating leader or party (re)selection), (b) performance signal (country-level economic performance), (c) clarity of performance signal, and (d) leader charisma.

All data are from 1916-2008. We derived a charisma score from acceptance speeches (or acceptance letters before the advent of such speeches) by Democrat and Republican presidential candidates at their party's national convention. We obtained the speeches from *The American Presidency Project* and from the *Official Report of the Proceedings* (1832-1952). We use these speeches because they are delivered in similar settings, allowing for a measure of comparison that would not have been possible to attain using other texts (e.g., debates or stump speeches). Also, the nomination acceptance speech has to move to the "center" of the political spectrum and appeal to the middle-of-the-road swing voters. Acceptance speeches arguably constitute the single most important speech delivered by the candidates during the election period and should be the speech heard or read by the greatest number of voters (either directly or indirectly through various media outlets); indeed, the *Los Angeles Times* reported that the final nights of both conventions each drew over 30 million live TV viewers in the U.S. (James, 2012).

The Fair Presidential Voting Equation

Fair's (1978) theory of economic voting, and its extension (Fair, 2009) models the vote share of the two-party U.S. presidential vote (i.e., considering only the Democratic and Republican contenders and excluding third parties) based on the performance of the U.S. economy and presidential incumbency. Predicting the two-party vote share thus simplifies the testing of the theory because we only need to model voter choices between two candidates. For a review of the theoretical framework of the model—which is beyond the scope of this paper—refer to Fair (1978, 2009). The latest specification of the presidential voting equation is the following model for election year *t* (for the sake of consistency, we use the same variable names as Fair):

$$V_t = \alpha_0 + \alpha_1 G_t . I_t + \alpha_2 P_t . I_t + \alpha_3 Z_t . I_t + \alpha_4 DPER_t + \alpha_5 DUR_t + \alpha_6 I_t + \alpha_7 WAR_t + \epsilon_t \qquad (1)$$

The dependent variable is the Democrat vote-share, *V*. Regarding the economic variables, there is one a short horizon variable *G*, which is the growth rate (at an annual rate) of real per capita GDP in the first three quarters of the election year; two variables, *P* and *Z*, cover the entire period of the current administration up to the election. *P* is the absolute value of the GDP deflator (inflation rate at an annual rate). *Z* measures the number of quarters in which the growth rate of per capita GDP exceeded 3.2% at an annual rate; in this respect *Z* can be thought of as a "good news" variable. *DPER* and *DUR* capture two antagonistic effects of incumbency: *DUR* captures the voters' weariness from having the same party in power (coded 0 if either party has been in the White House for one term, 1 [−1] if the Democratic [Republican] party has been in the White House for two consecutive terms, 1.25 [−1.25] if the Democratic [Republican] party has been in the White House for three consecutive terms, and so on, i.e., addition of .25 or -.25 for each term added). *DPER* captures the advantage the incumbent president has as consequence being a familiar figure (coded 1 if a Democratic presidential incumbent is running again, -1 if a

Republican presidential incumbent is running again, and 0 otherwise). Because the dependent variable is the Democratic vote-share, the variable *I* is used to determine the direction of the effect of the three economic variables (with which it is multiplied): *I* = 1 if the presidential incumbent at the time of the election is a Democrat and *I* = -1 if the incumbent is a Republican. Finally, a dummy variable, *WAR*, captures the short and long-term effects specific to both world wars on the national economy. No other effects of wars are considered by Fair because no other conflicts have affected the U.S. economy in the same way. Thus, to allow for direct comparison with the Fair presidential voting equation, we maintain its original specification. We obtained the data for the 1916-2008 period from Fair (2010); all data for this study are listed in Table 1.

[Table 1 here]

Extending the Fair Presidential Voting Equation with Charisma

Because we were interested in the relative charisma difference between the two nominees, we used the charisma difference score between the candidates as the independent variable. To control for effects due to speech length (i.e., longer speeches have more opportunities to use charismatic tactics) we included the difference in number of sentences between the Democrat and the Republican speeches as a control variable. We estimated the following model:

$$V_t = \beta_0 + \beta_1 G_t . I_t + \beta_2 P_t . I_t + \beta_3 Z_t . I_t + \beta_4 DPER + \beta_5 DUR + \beta_6 I_t + \beta_7 WAR_t + \beta_8 Charisma_t + \beta_9 Length + u_t \quad (2)$$

Where *Charisma* denotes the difference between the Democrat and Republican candidate charisma scores ($C^{dem} - C^{rep}$). Thus, a positive value of charisma indicates that the Democrat candidate employed more rhetorical items than did his Republican counterpart. *Length* is the difference in speech length between the Democrat and Republican candidate ($L^{dem} - L^{rep}$). We controlled for length instead of a relative measure (i.e., *Charisma* divided by *Length*) because the

latter does not capture the number of times an audience is exposed to rhetorical signals. For instance, if one speech is twice as long as another, we cannot assume that the two speeches will have the same impact, even if they use the same proportion of rhetorical signals.

To test the interaction hypothesis, we extend Equation (2) to the following:

$$V_t = \gamma_0 + \gamma_1 G_t . I_t + \gamma_2 P_t . I_t + \gamma_3 Z_t . I_t + \gamma_4 DPER + \gamma_5 DUR + \gamma_6 I_t + \gamma_7 WAR_t + \gamma_8 Charisma_t + \gamma_9 Length_t + \gamma_{10} Performance\ Signal_t + \gamma_{11} Performance\ Signal * Charisma_t + \varphi_t \quad (3)$$

Where *Performance Signal* is defined as the ex-ante (predicted) absolute victory margin using only attributional information, that is, the predictions derived from Eq. 1 (see section "Measuring performance signals"). *Performance Signal*Charisma* is the interaction of the later variable with charisma. Thus, the test for the hypotheses in the full specification is a test of the following coefficients (to the extent that the directions of the effects are as predicted):

1. Hypothesis 1: $\gamma_1, \gamma_2, \gamma_3 = 0$ (that the economic variables matter)
2. Hypothesis 2: $\gamma_8, \gamma_9 = 0$ (that charisma matters)
3. Hypothesis 3: $\gamma_{11} = 0$ (that charisma has a stronger effect when performance signals are ambiguous; that is, a positive simple slope for *Charisma* when the value of *Performance Signal* is small and a nonsignificant slope when the value of *Performance Signal* is large).

We also tested whether $\gamma_4, \gamma_5, \gamma_6 = 0$ to determine whether the incumbency variables matter.

Measuring charisma: We use an objective measure of charisma based on rhetorical tactics employed by leaders (Antonakis et al., 2011). We did not directly measure how charismatic the leader seems to voters because such a measure would be biased and unavailable for non-contemporary leaders. We thus measure the projection of charisma via objectively measured

rhetorical skills (e.g., see Emrich et al., 2001; House et al., 1991) and assume the projection engendered charisma in the eyes of observers via the psychological mechanisms we discussed.

The measure pertains to substantive statements, framing and vision creation. We list these tactics with a brief explanations (see Antonakis et al., 2011 for details). *Framing* and *creating a vision* is accomplished through (a) metaphors: simplify the message, trigger an image, aid recall; (b) rhetorical questions: create an intrigue and interest in knowing the answer; (c) stories and anecdotes: elicit an image, create identification with the protagonists, distill a message into a moral; (d) contrasts and comparison: define the vision in terms of what it should or should not be and focus attention on the message; (e) three-part lists: provide "proof" for the arguments, focus attention, and show completeness; *Substantive statements* include*:* (f) expressing moral conviction: highlights value systems and provides justification for the mission; (g) expressing the sentiments of the collective: showing similarity of the leader with the followers; (h) setting high and ambitious goals: shows ambition, aligns efforts towards goals; (i) creating confidence goals can be achieved: raises self-efficacy belief. In view of their complex semantic and syntactic structure, these items are coded by trained human coders and not by computer.

Because the sample period goes back to the early 1900s, we did not have video and/or audio recordings for all the candidates. Thus, we did not code nonverbal charisma; however, indicators of nonverbal charisma (e.g., facial expressions, body language, use of voice) correlate strongly with the nine verbal strategies, ($r = .48$, $p < .001$, uncorrected for unreliablity, Antonakis et al., 2011). Leaders who are very expressive with words are also expressive non-verbally; thus, the markers we have used should capture charisma, broadly defined. Also, because candidates use speechwriters, a concern is that speeches might not reflect the charisma of candidates but rather that of speechwriters. Even if the coded rhetoric might, to a certain extent, reflect the charisma of

the speechwriter, nominees have an important say in developing their speech. Ultimately the speech will affect charismatic inference on the candidate because the candidate and not the speech writer delivers the speech (see e.g., Bligh, Kohles, & Meindl, 2004; House et al., 1991).

Two graduate students coded all 48 acceptance speeches. We provided extensive training to the coders on practice speeches. To minimize biases from their knowledge of the candidates, we masked information associated candidates' identities. For example, we concealed individual names (unless not used to name someone, e.g., in "the Taft Act" or the "George Washington bridge"), years and any mentions of party affiliation or ideology (i.e., "Republican", "Democrat", "liberal", "conservative" and related terms). So as not to create a void in meanings and confusion for the coders, we left certain references that could allow the coders to identify a general period in time untouched (e.g., references to the "iron curtain", "prohibition" or the "Vietnam war").

Each coder independently coded each speech at the sentence level for the presence or absence of the charismatic items. If a candidate employed a charisma item over several sentences (e.g., using a metaphor across two consecutive sentences), only the first sentence was coded. We summed the frequency of use of each charisma item to arrive at a charisma score. The coders correlated very highly with each other for both the Democrat ($r = .90$, $p < .001$) and Republican candidate ($r = .89$, $p < 001$) speeches. To ensure a maximum reliability for the charisma score of each candidate, we used the mean score across the two coders for analyses. The charisma and speech length scores are included in Table 1. We checked whether using difference scores was defensible (Edwards & Parry, 1993) both here and elsewhere (see below).

Convergent validity of charisma measure: To ensure that our scores were valid indicators of charisma, we examined whether they converged with theoretically similar measures. We obtained data from six studies. Data from (a) Lichtman's (2012) "Keys to the Whitehouse", were

available for all contenders. We predicted the charisma scores of Lichtman's model (i.e., by computing the difference of charisma between the Democrat and Republican nominees) using our charisma difference score and controlling for the difference in speech length; we also used a robust variance estimator. Because Lichtman coded for the presence (1) or absence (0) of charisma for both candidates, we model the range of scores as an ordered probit for the values of -1, 0, and 1 (i.e., $C^{Dem} < C^{Rep} = -1$; $C^{Dem} = C^{Rep} = 0$; $C^{Dem} > C^{Rep} = 1$). Results indicated strong convergent validity with this measure ($n = 24$, standardized $\beta = .67$, $z = 2.93$, $p < 0.01$).

For the other five studies, we only found charisma data for the election winner (i.e., the president) and these data were not available full sample period of our study. Nonetheless, these data were still useful given that the measures were derived from different methods. For these specifications, we controlled for speech length and party affiliation, and we used cluster-robust standard error corrections (at the candidate level because we have different charisma scores for presidents who served more than one term, whereas the other researchers gave only one overall score). Our charisma measure was significantly related to: (b) Simonton's (1988) ratings of charisma of U.S. presidents ($n = 18$, standardized $\beta = .44$, $t = 2.20$, $p < 0.05$); (c) Emrich et al.'s (2001) measure of image based rhetoric in U.S. presidential inaugural addresses ($n = 18$, standardized $\beta = .56$, $t = 2.58$, $p = 0.01$) as well as to (d) House, Spangler, and Woyke's (1991) measure of *charismatic effects*, ($n = 9$, standardized $\beta = .59$, $t = 3.16$, $p < .01$) and (d) House, Spangler, and Woyke's (1991) measure of *charismatic behaviors*, ($n = 9$, standardized $\beta = .60$, $t = 3.02$, $p < .01$).[2] Note, for these first four studies (Emrich et al., 2001; House et al., 1991; from

[2] The zero-order correlation between Simonton's (1988) ratings of charisma of U.S. presidents and Emrich et al.'s (2001) measure of image based rhetoric in US presidential inaugural addresses is .54; thus, we estimated the models for these two variables simultaneously (and hence gain in estimation efficiency). Consequently, the sample size is reduced to 18 observations for both models—the sample for Simonton's (1988) ratings of charisma of U.S. presidents would have otherwise amounted to 22 observations (with a convergent validity estimate of .33, $p < .10$).

Lichtman, 2012; Simonton, 1988), the sample size weighted average standardized beta was quite high (i.e., $\bar{\beta}$ = .58); corrected for unreliability in our measure (Murphy & Davidshofer, 2005) and assuming a reliability of .80 for the other measures suggests that our measure correlates very highly (i.e., $\bar{\beta}_{corrected}$ = .71) with these other measures of charisma. Thus, there is strong evidence for the convergent validity in our measure of charisma.

Finally, our measure of charisma was also related to two measures concerned with the use of metaphor rather than with charisma per se, which provides a useful approximation for the validity of our measure (uncorrected for unreliability): (e) Mio et al.'s (2005) measure of presidents' *overall metaphor density* (*n* = 26, standardized β = .81, *t* = 5.15, *p* < .01) and (f) Mio et al.'s (2005) presidents' *metaphor density score for inspiring passages* (*n* = 26, standardized β = .74, *t* = 4.36, *p* < .01). Thus, there is strong evidence for the validity of our charisma measure.

Stability of charisma measure: We examined if charisma is consistent over time and thus ensured that the acceptance speech is a good proxy of how charismatic a contender is in general. We obtained charisma scores for between two and four additional speeches from a randomly selected subsample of six presidential candidates. We examined whether the scores derived from several speeches (*n* = 27; including acceptance speeches) resembled each other across the subsample of candidates. Using a fixed-effects model with cluster robust standard errors, controlling for speech length, and including a dummy variable to control for the unique nature of acceptance speeches we find a high intraclass correlation (ρ = .33, bootstrapped *SE* =.15, *z* = 2.14, *p* < .05); note too that acceptance speeches had significantly higher charisma scores (β = 39.86, standardized β = .67, *SE* = 7.57, *t* = 5.26, *p* < .01), which shows the importance that candidates

---

We estimated the model with House et al.'s (2005) measures separately (in one seemingly unrelated model) because it would otherwise have constrained the sample size of the other models to 9 observations.

give to the acceptance speech. Thus, there is some evidence of consistency over time—particularly when considering our small sample size—and the fact that candidates place a large emphasis on the acceptance speech. This result provides additional evidence of construct validity.

Measuring performance signals: For the *Performance Signal* measure, we used the Fair model (i.e., Eq. 1) to estimate "out-of-sample" predictions of the Democratic vote share. An out-of-sample prediction is a forecast that does not use the observed values for this event in estimation. In our case, this means that to obtain an out-of-sample prediction for election year *t*, we do not use the data for election year *t* to fit the model; we only use the sample data for year *t* to generate predicted estimates from the fitted model. For example, the out-of-sample prediction for 2008 uses 1916-2004 data to estimate the model parameters. On the basis of the fitted model, we generate predictions for 2008 by multiplying the estimated coefficients of the fitted model with the observed values for the predictors for 2008 to obtain a forecast of $\hat{V}$ for 2008. Similarly, the out-of-sample prediction for 1916 uses data from 1920-2008 to fit the model. Using the out-of-sample prediction method makes for very objective model test (Meese & Rogoff, 1983). Model fitting and forecasting evaluation are separated and the modeler does not "cheat" by including the observed data of a particular year to fit the model and then predict from it.

To generate values for the variable *Performance Signal* we used the absolute difference of the out-of-sample predictions of the econometric model (i.e., from Eq. 1) subtracted from 50% (i.e., *Performance Signal* = $\left|\widehat{V_t} - 50\right|$). Thus, the value of *Performance Signal* is the predicted margin of victory when only considering attributional processes (the economy) and ignoring inferential processes (candidate charisma). Values close to 0% indicate a small margin of victory, showing a mitigated view of the economic fundamentals: Performance signals are ambiguous. As values move away from 0%, the economic signals indicate whether the economy is doing

well or not. Note that *Performance Signal* has sufficient variation, ranging from .24 to 14.91 (mean = 5.02, *SD* = 4.44) and that for about half (i.e., 13/24) of the elections *Performance Signal* takes on values of less than 5% which would qualify these elections as close-call elections (The Washington Post, 2012); these are the situations where charisma should matter most.

As a measure of convergent validity for *Performance Signal,* we correlated the out-of-sample prediction of the vote share with the actual vote share received ($r(22) = .85$, $p < .001$; note, we exploit this correlation later on when estimating a two-stage least squares regression).

Results

Hypothesis tests: We estimated all models using *Stata* version 13—see Table 2. We first show the results of the econometric model (Model 1), to which we add *Charisma* and *Length* (Model 2). The full model includes the performance signal-charisma interaction (Model 3).

[Table 2, Figure 1 here]

The results of Model 1 are the same as those presented in Fair (2010). The regression model predicted a large portion of the variance in the vote share (91.15%). Adding *Charisma* and *Length* (Model 2) improved the *r*-square (to 93.48%), as did the *Performance Signal* and the interaction of *Performance Signal * Charisma* variables (to 95.55%).

We tested the hypotheses, using the full specification (Model 3). Hypothesis 1 was supported; the set of econometric variables were simultaneously predictive of the vote share as hypothesized: The growth variables *G* and *Z* are positive predictors and the inflation variable *P* is a negative predictor ($F(3, 12) = 33.21$, $p < .001$). With respect to the incumbency variables, they were also simultaneously predictive too ($F(3, 12) = 18.07$, $p < .001$) in the expected direction (i.e., the incumbent variable was positive and the duration variable was negative). The result also support Hypothesis 2 regarding the addition of *Charisma* and *Length* ($F(2, 12) = 5.35$, $p < .05$) to the

model. The simple main effect of charisma alone (i.e., when *Performance Signal* = 0) was significant too (β = .13, *SE* = .05, $p < .05$). This simple main effect suggests that in a predicted dead-heat election charisma can make a very large difference to the election outcome (the standardized beta is .42). These results provide support for Hypothesis 2.

As regards Hypothesis 3, the interaction of *Performance Signal* * *Charisma* was significant (t =2.26, $p < .05$). The form of the interaction provided strong support that charisma matters more when performance signals are ambiguous. Holding all other predictors constant at the means, the simple slope of *Charisma* under clear performance signals (+1 SD from the mean *Performance Signal*) was not significant (β = -.05, *SE* =.04, t = 1.18, $p > .10$). However, the simple slope of *Charisma* under ambiguous performance signals (-1 SD from the mean of *Performance Signal*) is positive and significant (β = .12, *SE* = .05, t = 2.47, $p < .05$).

Forecasting accuracy: To show the model is not overfitted (Roberts & Pashler, 2000), we examined the forecasting accuracy by comparing the out-of-sample predicted values from the various specifications to the actual vote received and calculated the absolute error (i.e., predicted less actual value). Overall, our models (i.e., Models 2, 3 and 4, see next) perform substantially better than does the Fair model (i.e., Model 1). The mean absolute error (MAE) for Models 2, 3, and 4 is 2.70, 2.76, and 2.90 respectively, whereas it is 2.85 for the Fair model (see Table 1).

We can also compare if the election is "called" correctly (a practically-useful outcome). Model 3 correctly identifies the winner in 21 out of 24 elections; the hit rate for Model 1 is 17 of the 24. Interestingly, and confirming our theorizing about the importance of clear performance signals, the actual winning margin of the correctly called elections in the cases where Model 1 was right was $Actual_{\text{correct}}$ = 7.30% (*SD* = 3.60) with an average predicted winning margin for this model at $Predicted_{correct}$ = 6.58% (*SD* = 4.35); however, the actual winning margin of those

elections that were incorrectly called by this model was only $Actual_{incorrect}$ = 1.37% (*SD* = 1.27) with an average predicted winning margin for this model at $Predicted_{incorrect}$ = 1.22% (*SD* = 1.22). Thus, when Model 1 predicts the election will be close (i.e., the performance signal is ambiguous) it actually is close and this is when this model is likely to get it wrong because it ignores individual differences in times of attributional ambiguity (note $Actual_{\text{correct}} > Actual_{incorrect}$, $t = 4.21$, $p < .001$; $Predicted_{correct} > Predicted_{incorrect}$, $t = 3.17$, $p < .01$).

As an additional and more basic check for the interaction hypothesis we looked at the predictive accuracy of Model 2 (the model without the interaction but with *Charisma*) versus Model 1. Model 2 called 21/24 elections right, whereas Model 1 got 17/24 right (likelihood ratio $\chi^2(1) = 8.52$, $p < .01$). We then compared hit rates for the 13 elections predicted to be a close call (i..e, < 5% victory margin, The Washington Post, 2012). Model 2 got 10 correct, whereas Model 1 only got 6 correct. This difference was significant (likelihood ratio $\chi^2(1) = 4.49$, $p < .05$). Overall, these results provide converging evidence in support for Hypothesis 3.

Robustness checks for endogeneity: There are two concerns with endogeneity for the estimates of *Performance Signal* and *Charisma*. First, although *Performance Signal* is theoretically exogenous with respect to *V* (vote share), endogeneity could bias results because we use an ex-ante estimated regressor that has some degree of uncertainty in it. We thus reestimated Model 3 (referred to as Model 4 for this analysis) using two-stage least squares analysis (2SLS) (Antonakis, Bendahan, Jacquart, & Lalive, 2010). To do so, we need to have an instrument—an exogenous sources of variance that cannot possibly vary as a function of omitted causes in *Performance Signal* or *V*. We used the actual margin of victory as our instrument, because it is perfectly observed and an absolute value of victory for either one of the parties it is exogenous to *V*. We therefore estimate the model with robust standard errors "instrumenting" *Performance*

Signal and *Performance Signal * Charisma* with the actual margin of victory, its square and cube as well as the interaction of these three variables with *Charisma,* and included the rest of the variables as exogenous regressors (see Wooldridge, 2002 for how to instrument for endogenous interactions). The instruments were "strong," the overdidentification test nonsignificant ($\chi^2(4) = 2.53$, $p > .10$),and the endogeneity test of *Performance Signal* and *Performance Signal * Charisma* indicated that they were endogenous with respect to vote share (Hausman, 1978); still, because the 2SLS point estimates (reported in Model 4, Table 2) did not differ from the Model 3 OLS estimates, as indicated by a Wald test ($\chi^2(12) = 14.27$, $p > .10$) we retain the OLS estimates. All hypothesis tests were nonetheless still supported using 2SLS, and the interaction pattern was similar.

Second, it is possible that the charisma of political candidates is determined by—rather than caused by—expected or previous electoral outcomes rending estimates inconsistent (Antonakis et al., 2010). Perhaps the political party that has lost the previous election, or which is more likely to lose the current election as a result of economic and incumbency factors, selects a more charismatic candidate to reverse the electoral outcome. In order to test for this potential reverse causality, we regressed *Charisma* on the predicted values from Model 1; we also regressed *Charisma* on the actual vote share from the previous election (while controlling for *Length*). Both results were non-significant, indicating that charisma is exogenous to such selection effects.

Robustness check for *r*-square: Because of the small sample size, we checked if Model 3 does better than chance in terms of variance prediction. We used two Monte Carlo simulations, to see how a model having the same *n*-size and *k*-predictors would do under different conditions. The first simulation used normally distributed random variables. For the second simulation we used the observed correlation matrix of the data for Model 3 from which we calculated the mean

absolute correlation between the 12 variables. The mean correlation was .28. To err on the side of caution, we created a correlation matrix manipulating the mean correlation between the variables at .30, .40 and .50. Then, we introduced random "shocks" in each predictor by adding $x*e$ to each variable (where x was manipulated from .1 to 1, and where e is a normal distributed random variable with a mean of 0 and a SD of 1). We crossed the two manipulations and ran both simulations 1,000 times (for $n = 24$). The highest upper boundary of the 95% confidence interval we obtained was .58; the lower 95% bootstrapped boundary of Model 3' r-square was .93, which was significantly higher that .58 ($\chi^2(1) = 819.16$, $p < .001$). Given the out-of-sample forecasts and the simulation evidence, it is unlikely that Model 3 is "overfitted."

Predicting the 2012 election (*ex-ante*): To forecast the 2012 election (done prior to the election), the two authors independently coded the nomination acceptance speeches of Barack Obama and Mitt Romney. We coded each sentence for the presence or absence of the nine charismatic tactics. Obama's speech had 1,908 coding events (212 sentences * 9 coding categories); Romney's had 2,430 coding events (270 sentences * 9 coding categories). Agreement statistics (Landis & Koch, 1977) on the two codings indicated strong agreement ($\kappa_{Obama} = .69$, $SE = .02$, $z = 30.28$, $p < .001$, agreement = 94.81%; $\kappa_{Romney} = .67$, $SE = .02$, $z = 33.28$, $p < .001$ agreement = 96.63%). We obtained very similar findings when combining both speeches or when only examining agreement for each of the 9 coding categories. We therefore averaged the final scores of the coders. Obama received a score of 177.5; Romney's was 131 (charisma difference = 46.5).

We used forecasts of economic data from approximately two months and two weeks before the election (Fair, 2012); we report predictions for Models 1 and 3 from two weeks before the election (substantive projections for two weeks and two months prior were the same). Model 1 predicted that Obama would lose the popular vote with 49.05% ($SE = 1.76$, $z = 27.81$, $p < .001$,

90% CIs 45.97 to 52.13). Evidently, the election was too close to call on the basis of this model as some commentators noted (New York Times, 2012), reflecting the then mitigated view of the economy. We used the Model 1 prediction to compute the value of the *Performance Signal* (i.e., |49.05097% - 50%|= .94903%) for Model 3. This model showed that Obama would win 55.50 % of the two-party vote (SE = 2.88, t = 19.25, $p < .001$, 90% CI's 50.37 to 60.64). We tested whether 55.50% differed from 50%; this result suggested that it was unlikely that Obama would lose this election (F (1, 12) = 3.65, p = .08). Model 4 (2SLS) predicted 57.79% for Obama (F(1, 12) = 13.59, $p < .001$ for Δ from 50%), who actually received 51.96% of the two-party vote.

Brief discussion

Controlling for the idiosyncrasy of the context of the U.S. presidential election, results of this study suggest that when performance signals are ambiguous, leaders are selected on the basis of how charismatic they are. However, when performance clearly positive (or negative), incumbents or their parties are rewarded (or punished) at the next elections. Although the effects of charisma were quite strong, they should be lower-bound estimates to the extent that politicians face intense competition before being nominated by their party to run for the presidential race.

STUDY 2

Although Study 1 made a relatively strong case for external validity, we could only assume that inferential and attributional processes drove these results. Here we used an experimental design to directly test the effect of these two psychological processes. Using realistic stimulus materials, we tested if our theory explains retention decisions about an incumbent CEO.

Materials and procedure

We designed a TV business report, *Business Line*, portrayed by actors and professionally produced to provide an ecologically valid context to manipulate information about a company's

performance and the charisma of its CEO. We asked participants to watch the newscast, which profiled a fictitious U.K.-based company called *BlueTech*. The news anchor introduced the report stating that there were recent increased trade volumes in the firm's stock and that investors were wondering whether to buy, hold, or sell their stocks. He then spoke to a reporter from the London Stock Exchange about *BlueTech's* stock price who in turn conveyed analysts' recent recommendations. The anchor then read a statement from the CEO of the firm regarding recent happenings, and then interviewed a former employee (a financial analyst) of the CEO.

Given the focus of the newscast, we manipulated CEO charisma (high vs. low) and firm performance signals (either clearly good, ambiguous, or clearly poor), resulting in 2 x 3 between subjects experimental design. We thus created six versions of the stimulus video—with each version of the video lasting approximately four minutes (e.g., for the charismatic CEO condition with good company performance see http://www.youtube.com/watch?v=zN1RPthANFA). We hosted the videos on a private YouTube channel and embedded them in Qualtrics. Participants only accessed one condition. After watching the newscast we asked participants to suppose they were on the board of directors of the company and then to vote on whether the current CEO should be reappointed or replaced; it is this decision that we model as the selection decision.

Participants

Our usable sample was 717 (47.84% women) U.S. participants recruited via Amazon's Mechanical Turk. These participants were thoroughly screened on appropriate control questions and memory checks regarding the content of the videos (cf. Mason & Suri, 2012) and were on average 32.65 years old (SD = 11.45). They represented all 20 industries listed in the North American Industry Classification System (United States Census Bureau, 2013) and all 50 states.

Most (58.17%) had a received a college degree; 31.94% occupied at least a managerial position and 42% of participants owned stocks or mutual funds.

Manipulated variables

CEO charisma: We manipulated CEO charisma in three ways: (a) directly through the CEO statement read by the anchor—in terms, of how many charismatic rhetorical tactics were used (i.e., six vs. none; note, both versions of the CEO statement had the same number of sentences and words); (b) indirectly via how others described the leader in this case by the news anchor, and (c) indirectly through descriptions about the CEO in an interview given by a former employee of the CEO (cf. Bligh, Kohles, & Pillai, 2011; Meindl, 1990).

Firm performance signals: We manipulated firm performance through (a) presenting stock price over the last quarter on a graph where it was either increasing by 8%, decreasing by 8% or being almost constant (ambiguous signal), (b) analysts' recommendations (buy, hold, sell), and (c) the recommendations of the finance specialist interviewed on the show (buy, hold, sell).

Manipulation checks

To verify participants' perceptions of CEO charisma, we used the MLQ's idealized influence (attributes), idealized influence (behaviors) and inspirational motivation scales (Avolio, Bass, & Jung, 1995). We measured leader effectiveness with the MLQ's effectiveness scale. Cronbach alpha reliabilities of the four scales were .91, .89, .94 and .89. We then regressed each of the scales on the respective manipulated factor (note, all *F*-tests below are heteroscedastic robust).

Results showed that for idealized influence (attributes), the mean of the charismatic CEO condition was higher than that of the non-charismatic one (mean = 2.88, SD = .83 vs mean = 1.06, SD = .77, model $F(1, 715) = 928.10$, $p < .001$, *r*-square = .57). The results for idealized influence (behaviors) (mean = 2.70, SD = .83 vs mean = 1.23, SD = .82, model $F(1, 715) =$

572.24, $p < .001$, r-square = .44), and inspirational motivation were similar (mean = 3.43, SD = .72 vs mean = 1.21, SD = .90, model $F(1, 715) = 1345.96$, $p < .001$, r-square = .65). For effectiveness, the means of the three conditions differed as a function of performance cues ($mean_{negative}$ = 1.51, SD = .90, $mean_{ambiguous}$ = 1.79, SD = 1.08, $mean_{positive}$ = 2.53, SD = .94, model $F(2, 714) = 73.65$, r-square = .16; contrast $F(2, 714) = 55.29$, $p < .001$, all p-levels Bonferroni adjusted). These results suggest that the manipulations had their intended effects.

Results

Using a linear probability model (LPM) with a heteroscedastic-robust estimate of the variance, we regressed the binary variable—reflecting whether the CEO should be reappointed—on the manipulated variables, *Charisma* and *Performance Cues* and the interaction (the omitted category was *Ambiguous Performance cues*). We used a LPM instead of a probit regression because the LPM is a consistent estimator of binary outcome models having marginal effects isomorphic to the observed coefficients (cf. Moffitt, 1999); predicted probabilities and differences in marginal effects using the probit regression, nonetheless gave the same results. We probed the interaction effects using post-estimation tests (e.g., testing differences in predicted probabilities), which we undertook using the delta method. Refer to Table 3 for the results.

[Table 3]

The regression model predicted CEO reappointment. The effects of *Performance Cues* were significant in the full model ($F(2, 711) = 53.30$, $p < .001$), and main effects contrasts showed that negative cues significantly decreased the likelihood of CEO reappointment as compared to ambiguous cues ($F(1, 711) = 12.39$, $p < .001$); likewise, positive cues significantly increased the likelihood of leader reappointment with respect to ambiguous cues ($F(1, 711) = 91.99$, $p < .001$). These results provide strong support for Hypothesis 1. In addition, the main

effect of *Charisma* was significant too in the full model ($F(1, 711) = 82.29$, $p < .001$) as was the main effect contrast ($F(1, 711) = 152.95$, $p < .001$), providing strong support for Hypothesis 2.

The *Charisma*Performance cues* interactions were jointly significant ($F(2, 711) = 5.46$, $p < .01$). Compared to the omitted category, the coefficient of *Charisma*Negative Performance* cue was lower though not significantly so; the *Charisma*Positive Performance* cue coefficient was significantly lower. A more powerful test using a linear combination shows that the average coefficient of the two clear performance signals conditions each interacted with charisma was significantly lower than the omitted category ($\beta = -.18$, $SE = .07$, $t = 2.76$, $p < .01$). This finding provides strong support for Hypothesis 3a: The effect of charisma on leader reappointment is significantly positive and higher than when performance signals are ambiguous.

We probed the interaction by generating predicted values (Figure 2). Unexpectedly, and as also shown by the main effect result, *Charisma* had a significant effect on leader selection in both clear performance signal conditions: The difference in predicted probabilities between Conditions 2 (.53) and 1 (.16) was significant ($\Delta_{Cond.2\text{-}1} = .37$, $SE = .06$, $t = 6.56$, $p < .001$) as it was too between Conditions 6 (.96) and 5 (.70) ($\Delta_{Cond.6\text{-}5} = .26$, $SE = .05$, $t = 5.65$, $p < .001$). Of course, the difference in Conditions 4 (.74) and 3 (.24) was significant also ($\Delta_{Cond.4\text{-}3} = .50$, $SE = .06$, $t = 9.07$, $p < .001$). As indicated previously, the effect of *Charisma* on selection in the ambiguous performance signal condition was significantly higher than that in the other two conditions (i.e., $\Delta_{Cond.4\text{-}3} > (\Delta_{Cond.2\text{-}1} + \Delta_{Cond.6\text{-}5})/2$, $F(1, 711) = 7.63$, $p < .01$; note, $\sqrt{7.63}$ = the *t*-statistic of 2.76 in the above paragraph).

[Figure 2 here]

Using a more fine-grained analysis, we tested if the predicted probabilities differed from .50 (i.e., a random choice). In all conditions, save one (Condition 2), the difference of the

predicted probability from .50 was significant (Bonferroni adjusted tests). Thus, even though charisma did help the leader in Conditions 2 vs. 1, it still did not help the leader enough. These results provide some support for Hypothesis 3b, namely that charisma has an insufficient (albeit a positive effect) on leader selection when performance signals are negative; however, contrary to what we expected charisma does have a positive effect on selection when performance signals are positive, though the effect is comparatively smaller than in the ambiguous condition.

Finally, our "litmus test" argument suggested that in Conditions 1 and 2, the leader should not be reappointed; the predicted probability across these conditions was .32 (< .50 $F(1, 711) = 28.79$, $p < .001$). Similarly, in Conditions 5 and 6, the leader should be reappointed; predicted probability across these two conditions was .83 (> .50 ($F(1, 711) = 251.61$, $p < .001$). In addition, our theory also suggests that Conditions 1, 2, and 3, characterized by negative performance signals, or no charisma in the presence of ambiguous performance signals, should have a detrimental effect on leader selection as compared to Conditions 4, 5 and 6. These latter conditions have positive performance signals, or a charismatic leader along with ambiguous performance signals. The linear combination of estimators showed that the mean predicted probability across conditions 1, 2, and 3 (i.e., .31) was significantly below .50 ($F(1, 711) = 69.91$, $p < .001$). The mean predicted probability of Conditions 4, 5, and 6 (i.e., .80) was significantly above .50 ($F(1, 711) = 213.79$, $p < .001$). Moreover $Probability_{(\text{Cond. } 1, 2, 3)} < Probability_{(\text{Cond. } 4, 5, 6)}$ ($F(1, 711) = 56.73$, $p < .001$). These results provide further support for Hypothesis 3[3].

Robustness checks

[3]We also asked participants how they would apportion a $1,000 investment between *BlueTech* stock and low-risk government treasuries. Charisma had the greatest effect under conditions of attributional ambiguity. With a non-charismatic leader, subjects allocated 20.27% of their monies to company stock; with a charismatic leader the allocation was 30.58% (i.e., 50.89% higher). There was no effect of charisma in the positive performance signal condition; however, charisma had a strong effect in the negative performance signal condition, increasing the allocation by 42.10% (from 13.17% to 22.75%). Participant job-level did not affect findings.

We checked for the effect of job level because employees at lower levels may not understand the dynamics of top-level leadership, market indicators, or investment issues. Thus, we interacted all the job level dummy variables (i.e., $k - 1$) with performance cues, charisma, and the performance cue-charisma interactions; the main effect of job level as well as the interaction coefficients of job-level remained insignificant (F (20, 691) = 1.35, $p > .10$).

Brief discussion

Experimental evidence, using high fidelity video material provided further evidence that charisma matters much more in conditions of attributional ambiguity than in conditions where performance signals of the leader's organization are clearly positive or negative. This finding suggests that similar psychological mechanisms underlie the evaluation and selection of top-level leaders, whether in the political or business arena.

GENERAL DISCUSSION

Our results demonstrate that charisma's effect is most evident when performance signals are ambiguous; that is, inferential processes matter most in conditions of attributional ambiguity. Applied to the first context in which we tested our model, the U.S. presidency, our results show that the charisma of political candidates matters and significantly improves the prediction of the Fair model particularly when performance signals are unclear (i.e., in predicted close call elections). We replicated these results using an experimental design in a business context, showing that reappointment of a CEO for participants acting as board members depended both on attributional and inferential process; the latter, in particular mattered most when performance signals are ambiguous. Participants in the role of investors reacted similarly, which suggests that decision mindsets for selection and investment choices are similar.

The experimental results showed that charisma increased the likelihood of CEO

reappointment across all performance conditions. Although charisma matters most in ambiguous performance conditions, our expectation that it would not matter under clear performance signals was not supported; charisma mattered though to a lesser degree. Moreover, following our "litmus test" argument we found that clear negative performance signals significantly reduced the CEO's chances of reappointment. When performance signals were clearly positive, the CEO's chances of reappointment were significantly increased. Although our model offers a refined view allowing us to make predictions about leader selection when inferential and attributional signals are not pointing in the same direction it seems that a leader positively (or negatively) assessed through both inferential and attributional routes will be more (or less) desirable than one who is only positively (or negatively) assessed through only one of these two routes.

In the case of high-level leaders who are distant from observers, our results underscore the importance of observer category-based information processing using markers of charisma. Indeed, it is through effective image-building and discourse that leaders communicate their vision, but also build an aura about themselves that is inextricably bound to the vision. Our results also partly reconcile theories claiming leadership is an attribution versus those claiming that leaders matter. Leaders do have a role in affecting organizational performance, as field experiments manipulating leadership have shown (Antonakis et al., 2010). The halo effect from performance cues is not just illusionary. Our study cannot speak to this part of the argument because we cannot discern if top-level leaders matter for country (or firm) performance; still we know there is strong causal evidence indicating that leaders can even affect country economic performance (Jones & Olken, 2005). Interestingly, and based on findings from Study 2, the coefficients of the dummy variables of cue on ratings of three MLQ charisma scales was significant ($\chi^2(6) = 30.36$, $p < .001$) taking the mean *r*-square across the equations from .55,

when only the charisma dummy was included, to .59. Yet, the simultaneous linear combination of cues was insignificant (β = .20, *SE* =.15, z = 1.33, $p > .10$) and that of manipulated charisma was higher (β = 1.76, *SE* =.08, z = 21.40, $p < .001$) and significantly so. Thus, the "romance bias" on *ratings of leadership* in our data was minimal. Coupled with the results of the CEO decision choice, our findings show that something "in" the leader affects evaluations beyond performance cues, which in turn matter much for leader selection when performance signals are clear. If selectors are unsure about how well the organization is doing, they will infer how competent the leader is based on how charismatic the leader appears.

Our findings also shed light on charisma per se in situations that may or may not be characterized as crisis. As we argued, attributional ambiguity is not necessarily reminiscent of a crisis and such situations do not necessarily give rise to charismatic leaders. The phenomenon we have identified is something unique and could contribute to advancing our understanding of leadership, and other phenomena where performance is evaluated: For example, work or interview performance is a situation wherein a target sends signals to an evaluator via attribution and inferential channels. How these signals are pondered should largely follow our theorizing—inferential signals will matter much when attributional signals are ambiguous. Consider the case of hiring recently minted Ph.D.'s for a faculty position: Because many applicants may not have a well-developed publication record (i.e., performance signal) an important part of the selection decision will be based on how they sell themselves in the job talk and interview, which will rely on factors like their charisma and communication skills, declarative knowledge of the field and so forth. Of course, those that have a demonstrated publication record (or a bad record) will have an advantage (or a disadvantage) if they are charismatic too (or regardless of their charisma).

These findings, applied to leadership, also have important practical implications for leaders—who should project strong charisma in times of attributional ambiguity and also associate or dissociate themselves from performance signals depending on whether the signals are positive or negative—but also for evaluators who should carefully consider whether they are correctly pondering inferential and attributional information. Evaluators could be unduly affected by charismatic targets in situations where performance-based information is ambiguous. Thus, it behooves evaluators to try and obtain concrete information about targets' true competence.

Our results should apply to situations where top leaders are evaluated and information on prior performance and leader prototypicality is available. For example, when boards of directors select CEOs they usually will have a short-list of candidates along with information on (a) how the firm they managed performed, as well as (b) on how charismatic they are (from personal observation and from the reports of others who have seen, heard, or read about the CEO). Thus, although there is no direct tournament, several candidates may be considered and the inferential and attributional dynamics we identified will surely play a role in the selection decision.

There are other situations to which the model would apply, for example in how top-level leaders are selected by a tournament for associations or professional bodies (here performance signals might not be economic but outcomes like the evolution in the number of members). The model could also be extended to cases where more than two candidates are vying for office.

Reflections on U.S. presidential leadership

Given the unique context of Study 1, our theory not only benefitted from studying political leaders, but it can also help explain how political leaders are judged. For example, our results challenge the prevailing wisdom regarding Obama's victory in the 2008 U.S. presidential election—apparently Obama won because of his charisma as suggested too by scholars (e.g.,

Bligh & Kohles, 2009). We disagree: If we estimate Model 3 for the case of a 2008 race in which a Democrat President had been in office for the two previous terms *ceteris paribus* (i.e., we reverse the incumbency record for the two past terms), we would predict a Republican victory (35.08% Democrat vote share). Similarly, for the case of a 2008 election for which the economic conditions had been very good *ceteris paribus* (i.e., inputting values of our model economic variables one standard deviation above the sample mean), our model again predicts a Republican victory (42.01% Democrat vote-share); Obama's charisma would not have shifted the balance. In any case, the econometric model (Model 1) predicted a Democratic victor in 2008 while ignoring the charisma difference of the candidates. The economic signals did suggest—in a relatively strong way—a Democratic margin of victory of 5.78% (which is slightly above the mean of the predicted margins of victory for Model1). So, despite Obama's charisma surplus, these models suggest that Obama did not win because of his charisma but partly because of the macroeconomic performance conditions in which the election was contested.

Still, charisma makes a large difference in close-call elections (e.g., 2012). Interestingly, most of the polls had Romney ahead for most of October 2012. The economic model suggested that it would be close election (and that Obama would lose). The economic conditions were mitigated and sent ambiguous signals: The economy was growing, but not strongly, and inflation was low. It appears that the reason why Obama defied political gravity and won the election was because he had an incumbent advantage, and because he was more charismatic than Romney was.

Limitations and suggestions for future research

Our model can be extended for any general case where evaluators evaluate (or select) using attributional and inferential decision processes. Thus, we hope to see studies that can obtain data on inferential and attributional mechanisms to predict CEO or political successions

using field data, particularly given the limited sample size we had in Study 1. Experimental studies should also consider modeling business situations that include more than one candidate vying for a position of CEO—for example, an incumbent CEO and an apparent "challenger" (i.e., one who is being considered for the incumbent's position). Other interesting designs could include manipulating objective leader expertise and responsibility in performance outcomes.

With regard to boundary conditions, national culture, which affects the aforementioned psychological process (Morris & Peng, 1994) should moderate the model as could other contextual factors (Liden & Antonakis, 2009). For instance, "close" observers might not use heuristics much. Yet it is hard to get close to top-level leaders. Although boards of directors, who select CEOs, might decide in a more individuating fashion, they still do not work on a day-to-day basis with CEOs and part of the information they get on the CEO is staged (in board of directors meetings, shareholders addresses, etc.). They also receive information about the CEO's performance indirectly (accounting reports, share prices, etc.). In addition, boards may also anticipate what distant outsiders like shareholders and analysts—who are affected by CEO charisma (Fanelli, Misangyi, & Tosi, 2009; Flynn & Staw, 2004)—will expect of a CEO. Director decisions could, therefore, be partly explained by our model.

CONCLUSION

We extended established leader evaluation and selection paradigms by fusing two distinct theories. We tested the model in a political as well as a standard business context. Through the lenses of attributional and inferential processes, we proposed that selectors evaluate leaders based on prior organizational performance for which these leaders are thought to be causally responsible. However, we showed too that selectors use information on leader prototypicality,

charisma, in conditions of attributional ambiguity. Such mixed-models should provide better explanations as to why leaders are selected and might prove useful in other performance domains.

References

Antonakis, J. & Atwater, L. E. 2002. Leader distance: A review and a proposed theory. ***The Leadership Quarterly***, 13: 673-704.

Antonakis, J. & House, R. J. 2002. An analysis of the full-range leadership theory: The way forward. In B. J. Avolio & F. J. Yammarino (Eds.), ***Transformational and Charismatic Leadership: The Road Ahead***: 3-34. Amsterdam: JAI Press.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. 2010. On Making Causal Claims: A Review and Recommendations. ***The Leadership Quarterly***, 21(6): 1086-1120.

Antonakis, J., Fenley, M., & Liechti, S. 2011. Can Charisma Be Taught? Tests of Two Interventions. ***The Academy of Management Learning and Education***, 10(3): 374-396.

Avolio, B. J., Bass, B. M., & Jung, D. I. 1995. ***MLQ Multifactor leadership questionnaire: Technical Report***. Redwood City, CA: Mindgarden.

Awamleh, R. & Gardner, W. L. 1999. Perceptions of leader charisma and effectiveness: The effects of vision content, delivery, and prganizational performance. ***Leadership Quarterly***, 10(3): 345-373.

Bartels, L. M. 2000. Partisanship and Voting Behavior, 1952-1996. ***American Journal of Political Science***, 44(1): 35-50.

Bertrand, M. & Mullainathan, S. 2001. Are CEOs rewarded for luck? The ones without principals are. ***Quarterly Journal of Economics***, 116(3): 901-932.

Binning, J. F., Zaba, A. J., & Whattam, J. C. 1986. Explaining the biasing effects of performance cues in terms of cognitive categorization. ***Academy of Management Journal***: 521-535.

Bligh, M. C., Kohles, J. C., & Meindl, J. R. 2004. Charisma under crisis: Presidential leadership, rhetoric, and media responses before and after the september 11th terrorist attacks. ***The Leadership Quarterly***, 2(15): 211-239.

Bligh, M. C. & Kohles, J. C. 2009. The enduring allure of charisma: How Barack Obama won the historic 2008 presidential election. ***The Leadership Quarterly***, 20(3): 483-492.

Bligh, M. C., Kohles, J. C., & Pillai, R. 2011. Romancing leadership: Past, present, and future. ***The Leadership Quarterly***, 22(6): 1058-1077.

Boeker, W. & Goodstein, J. 1993. Performance and successor choice: the moderating effects of governance and ownership. ***Academy of Management Journal***, 36(1): 172-186.

Calder, B. J. 1977. An Attribution Theory of Leadership. In B. M. Straw & G. R. Salancik (Eds.), ***New directions in Organizational Behavior***: 179-204. Chicago: St Clair.

Cantor, N. & Mischel, W. 1977. Traits as prototypes: Effects on recognition memory. ***Journal of Personality and Social Psychology***, 35(1): 38-48.

Chen, C. C. & Meindl, J. R. 1991. The Construction of Leadership Images in the Popular Press: The Case of Donald Burr and People Express. ***Administrative Science Quarterly***, 36(4): 521-551.

Conger, J. A. & Kanungo, R. N. 1987. Toward a behavioral theory of charismatic leadership in organizational settings. ***Academy of Management Review***, 12(4): 637-647.

Connelly, B. L., Tihanyi, L., Crook, T. R., & Gangloff, K. A. 2014. Tournament Theory: Thirty Years of Contests and Competitions. ***Journal of Management***, 40(1): 16-47.

Datta, D. K. & Guthrie, J. P. 1994. Executive succession: Organizational antecedents of CEO characteristics. ***Strategic Management Journal***, 15(7): 569-577.

Datta, D. K. & Guthrie, J. P. 1997. Contextual influences on executive selection: Firm characteristics and CEO experience. ***Journal of Management Studies***, 34(4): 537-560.

Day, D. V. & Lord, R. G. 1988. Executive Leadership and Organizational Performance: Suggestions for a New Theory and Methodology. ***Journal of Management***, 14(3): 453-464.

Den Hartog, D. N. & Verburg, R. M. 1997. Charisma and rhetoric: communicative techniques of international business leaders. ***Leadership Quarterly***, 8(4): 355-391.

Den Hartog, D. N., House, R. J., Hanges, P. J., & Ruiz-Quintanilla, S. A. 1999. Culture specific and cross-culturally generalizable implicit leadership theories: Are attributes of charismatic/transformational leadership universally endorsed? ***The Leadership Quarterly***, 10(2): 219-256.

Edwards, J. R. & Parry, M. E. 1993. On the use of Polynomial Regression Equations as an alternative to Difference Scores in Organizational Research. ***Academy of Management Journal***, 36(6): 1577-1613.

Emrich, C. G., Brower, H. H., Feldman, J. M., & Garland, H. 2001. Images in words: Presidential rhetoric, charisma, and greatness. ***Administrative Science Quarterly***, 46(3): 527-557.

Erickson, D. J. & Krull, D. S. 1999. Distinguishing Judgments About What From Judgments About Why: Effects of Behavior Extremity on Correspondent Inferences and Causal Attributions. ***Basic and Applied Social Psychology***, 21(1): 1-11.

Fair, R. C. 1978. The Effect of Economic Events on Votes for President. ***The Review of Economics and Statistics***, 60(2): 159-173.

Fair, R. C. 2009. Presidential and Congressional Vote-Share Equations. ***American Journal of Political Science***, 53(1): 55-72.

Fair, R. C. 2010. Presidential and Congressional Vote-Share Equations: November 2010 Update: Yale University.

Fair, R. C.; http://fairmodel.econ.yale.edu/vote2012/index2.htm [accessed: 9 September 2012 & 26 October 2012].

Fanelli, A., Misangyi, V. F., & Tosi, H. L. 2009. In Charisma We Trust: The Effects of CEO Charismatic Visions on Securities Analysts. ***Organization Science***, 20(6): 1011-1033.

Fiske, S. T. 1995. Social Cognition. In A. Tesser (Ed.), ***Advanced Social Psychology***: 149-193. Boston: McGraw-Hill.

Flynn, F. J. & Staw, B. M. 2004. Lend me your wallets: The effect of charismatic leadership on external support for an organization. ***Strategic Management Journal***, 25: 309-330.

Frese, M., Beimel, S., & Schoenborn, S. 2003. Action training for charismatic leadership: Two evaluations of studies of a commercial training module on inspirational communication of a vision. ***Personnel Psychology***, 56(3): 671-697.

Gaissmaier, W. & Marewski, J. N. 2011. Forecasting elections with mere recognition from small, lousy samples: A comparison of collective recognition, wisdom of crowds, and representative polls. ***Judgment and Decision Making***, 6(1): 73-88.

Gardner, W. L. & Avolio, B. J. 1998. The charismatic relationship: A dramaturgical perspective. ***Academy of Management Review***, 23(1): 32-58.

Hausman, J. A. 1978. Specification Tests in Econometrics. ***Econometrica***, 46(6): 1251-1271.

Hilger, S., Mankel, S., & Richter, A. 2013. The use and effectiveness of top executive dismissal. ***The Leadership Quarterly***, 24(1): 9-28.

Hodgson, G. M. 2006. What Are Institutions? ***Journal of Economic Issues***, 40(1): 1-25.

Hogg, M. A. 2001. A social Identity theory of Leadership. ***Personality and Social Psychology Review***, 5(3): 184-200.

House, R. J. 1977. A 1976 Theory of Charismatic Leadership. In J. G. Hunt & L. L. Larson (Eds.), ***The Cutting Edge***: 189-207. Carbondale: Southern Illinois: University Press.

House, R. J., Spangler, W. D., & Woycke, J. 1991. Personality and charisma and the U.S. presidency: A psychological theory of leader effectiveness. ***Administrative Science Quarterly***, 36: 364-396.

House, R. J. & Shamir, B. 1993. Toward the integration of transformational, charismatic, and visionary thories. In M. M. Chemers & R. Ayman (Eds.), ***Leadership theory and research: Perspectives and directions*** 167-188. San Diego: Academic Press.

Howell, J. M. & Frost, P. J. 1989. A laboratory study of charismatic leadership. ***Organizational Behavior and Human Decision Processes***, 43(2): 243-269.

Hutzschenreuter, T., Kleindienst, I., & Greger, C. 2012. How new leaders affect strategic change following a succession event: A critical review of the literature. ***The Leadership Quarterly***, 23(5): 729-755.

James, M.; Democratic convention's final night draws 35.7 million viewers. http://www.latimes.com/entertainment/envelope/cotown/la-et-ct-obama-speech-dnc-ratings-20120907,0,4580270.story [accessed 15 Sept. 2012].

Jones, B. F. & Olken, B. A. 2005. Do leaders matter? National leadership and growth since World War II. ***The Quarterly Journal of Economics***: 835-864.

Jones, E. E. & Davis, K. E. 1965. From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), ***Advances in experimental social psychology***, Vol. 2: 219-266. New York: Academic Press.

Kelley, S. & Mirer, T. W. 1974. The simple act of voting. ***The American Political Science Review***, 68(2): 572-591.

Kets de Vries, M. 1988. Origins of charisma: Ties that bind the leader and the led. In J. A. Conger & R. N. Kanungo (Eds.), ***Charismatic leadership: The elusive factor in organizational effectiveness***: 237-252. San Francisco: Jossey-Bass Publishers.

Khurana, R. 2002. ***Searching for a corporate savior : the irrational quest for charismatic CEOs***. Princeton, N.J.: Princeton University Press.

Landis, J. R. & Koch, G. G. 1977. The measurement of observer agreement for categorical data. ***Biometrics***, 33(1): 159-174.

Lewis-Beck, M. S. & Stegmaier, M. 2000. Economic Determinants of Electoral Outcomes. ***Annual Review of Political Science***, 3(1): 183-219.

Lichtman, A. J. 2012. ***Predicting the next president: The keys to the White House 2012*** (2012 ed.). Lanham, Md.: Rowman & Littlefield.

Liden, R. C. & Antonakis, J. 2009. Considering context in psychological leadership research. ***Human Relations***, 62(11): 1587-1605.

Lord, R. G., Binning, J. F., Rush, M. C., & Thomas, J. C. 1978. The effect of performance cues and leader behavior on questionnaire ratings of leadership behaviorr. ***Organizational Behavior and Human Performance***, 21(1): 27-39.

Lord, R. G., Foti, R. J., & De Vader, C. L. 1984. A Test of Leadership Categorization Theory: Internal Structure, Information Processing, and Leadership Perceptions. ***Organizational Behavior and Human Performance***, 34(3): 343-378.

Lord, R. G. & Maher, K. J. 1994. ***Leadership and information processing: Linking perceptions and performance***: Psychology Press.

Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. 1996. Effectiveness correlates of transformational and transactional leadership: A meta-analytic review of the MLQ literature. ***The Leadership Quarterly***, 7(3): 385-425.

Martinko, M. J., Harvey, P., & Douglas, S. C. 2007. The role, function, and contribution of attribution theory to leadership: A review. ***The Leadership Quarterly***, 18(6): 561-585.

Mason, W. & Suri, S. 2012. Conducting behavioral research on Amazon's Mechanical Turk. ***Behavior Research Methods***, 44(1): 1-23.

Mayer, W. G. 2008. What Exactly Is a Swing Voter? Definition and Measurement. In W. G. Mayer (Ed.), ***The Swing Voter in American Politics***: 1-31. Washington, D.C.: Brookings Institution Press.

Meese, R. A. & Rogoff, K. 1983. Empirical exchange-rate models of the seventies: Do they fit out of sample? ***Journal of International Economics***, 14(1-2): 3-24.

Meindl, J. R. & Ehrlich, S. B. 1987. The romance of leadership and the evaluation of organizational performance. ***Academy of Management Journal***, 30(1): 91-109.

Meindl, J. R. 1990. On leadership: An alternative to the conventional wisdom. ***Research in Organizational Behavior***, 12: 159-203.

Mio, J. S., Riggio, R. E., Levin, S., & Reese, R. 2005. Presidential leadership and charisma: The effects of metaphor. ***Leadership Quarterly***, 16(2): 287-294.

Moffitt, R. A. 1999. New developments in econometric methods for labor market analysis. In O. Asherzfelter & D. Card (Eds.), ***Handbook of Labor Economics***: 1367-1397: Elsevier.

Morris, M. W. & Peng, K. 1994. Culture and Cause - American and Chinese attributions for social and physical events. ***Journal of Personality and Social Psychology***, 67(6): 949-971.

Murphy, K. R. & Davidshofer, C. O. 2005. ***Psychological Testing. Principles and Applications*** (5th ed.). New Jersey: Pearson Education.

New York Times; Through an Economic Lens, an Election Too Close to Call; http://www.nytimes.com/2012/01/08/your-money/an-election-too-close-to-call-as-seen-in-an-economic-lens.html?_r=0; 27 March 2013, 2012.

Roberts, S. & Pashler, H. 2000. How persuasive is a good fit? A comment on theory testing. ***Psychological Review***, 107(2): 258-267.

Rush, M. C., Thomas, J. C., & Lord, R. G. 1977. Implicit Leadership Theory: A Potential Threat to the Internal Validity of Leader Behavior Questionnaires. ***Organizational Behavior and Human Performance***, 20: 93-110.

Shamir, B., House, R. J., & Arthur, M. B. 1993. The motivational effects of charismatic leadership: A self-concept based theory. ***Organization Science***, 4(4): 577-594.

Shamir, B., Arthur, M. B., & House, R. J. 1994. The rhetoric of charismatic leadership: a theoretical extenson, a case study, and implications for research. ***The Leadership Quarterly***, 5(1): 25-42.

Shamir, B. 1995. Social Distance and Charisma - Theoretical Notes and an Exploratory-Study. ***The Leadership Quarterly***, 6(1): 19-47.

Shamir, B. & Howell, J. M. 1999. Organizational and contextual influences on the emergence and effectiveness of charismatic leadership. ***The Leadership Quarterly***, 10(2): 257-283.

Simonton, D. K. 1988. Presidential style: Personality, biography, and performance. ***Journal of Personality and Social Psychology***, 55(6): 928-936.

The Washington Post; 2012 Election Map: The race for the presidency.; http://www.washingtonpost.com/wp-srv/special/politics/election-map-2012/president/; 2 September, 2012.

Tosi, H. L., Misangyi, V. F., Fanelli, A., Waldman, D. A., & Yammarino, F. J. 2004. CEO charisma, compensation, and firm performance. ***The Leadership Quarterly***, 15(3): 405-420.

Tversky, A. & Kahneman, D. 1974. Judgment under Uncertainty: Heuristics and Biases. ***Science***, 185(4157): 1124-1131.

United States Census Bureau. 2013. North American Industry Classification System. ***http://www.census.gov/eos/www/naics/***, [accesed 10 July 2013].

Waldman, D. A., Ramirez, G. G., House, R. J., & Puranam, P. 2001. Does leadership matter? CEO leadership attributes and profitability under conditions of perceived environmental uncertainty. ***Academy of Management Journal***, 1(44): 134-143.

Weber, M. 1947. ***The theory of social and economic organization*** (T. Parsons, Trans.). New York: The Free Press.

Weber, R., Camerer, C., Rottenstreich, Y., & Knez, M. 2001. The illusion of leadership: Misattribution of cause in coordination games. ***Organization Science***, 12(5): 582-598.

Wooldridge, J. M. 2002. ***Econometric analysis of cross section and panel data***. Cambridge, MA: MIT Press.

Zajac, E. J. & Westphal, J. D. 1996. Who shall succeed? How CEO/board preferences and power affect the choice of new ceos. ***Academy of Management Journal***, 39(1): 64-90.

TABLE 1

Data and predictions from the Fair and the extended Presidential Equation (Study 1)

Yr.	Democrat candidate	Republican candidate	*V*	*I*	*DPER*	*DUR*	*WAR*	*G*	*P*	*Z*	C^{dem}	C^{rep}	L^{dem}	L^{rep}	$\hat{V}_a$	$\hat{V}_b$
1916	**W Wilson**	CE Hughes	51.68	1	1	0	0	2.23	4.25	3	40	46	182	354	49.46	50.78
1920	J M Cox	**WG Harding**	36.15	1	0	1	1	-11.46	0	0	1	61	6	266	47.24	39.34
1924	J W Davis	**C Coolidge**	41.74	-1	-1	0	0	-3.87	5.16	10	58.5	40.5	241	362	43.01	40.03
1928	A Smith	**H Hoover**	41.24	-1	0	-1	0	4.62	0.18	7	45.5	41.5	328	200	43.41	44.11
1932	**FD Roosevelt**	H Hoover	59.15	-1	-1	-1.25	0	-14.59	7.16	4	46	68	189	339	64.91	64.19
1936	**FD Roosevelt**	A Landon	62.23	1	1	0	0	11.84	2.48	9	42	56.5	102	212	64.18	61.74
1940	**FD Roosevelt**	W Willkie	54.98	1	1	1	0	3.90	0.08	8	50.5	58	137	338	55.75	52.64
1944	**FD Roosevelt**	T Dewey	53.78	1	1	1.25	1	4.23	0	0	20.5	30.5	67	131	50.82	53.14
1948	**HS Truman**	T Dewey	52.32	1	1	1.50	1	3.64	0	0	26	36.5	146	78	49.56	52.28
1952	A Stevenson	**D Eisenhower**	44.71	1	0	1.75	0	0.73	2.35	7	21.5	19.5	72	49	45.72	46.18
1956	A Stevenson	**D Eisenhower**	42.91	-1	-1	0	0	-1.45	1.90	5	43	43	143	183	43.88	44.00
1960	**JF Kennedy**	R Nixon	50.09	-1	0	-1	0	0.46	1.94	5	50.5	85	122	179	48.44	49.89
1964	**L Johnson**	B Goldwater	61.20	1	1	0	0	5.09	1.27	10	43	72	127	122	60.78	60.23
1968	H Humphrey	**R Nixon**	49.43	1	0	1	0	5.05	3.13	7	60	80.5	138	254	50.60	50.77
1972	G McGovern	**R Nixon**	38.21	-1	-1	0	0	5.95	4.80	4	51.5	57.5	105	209	42.42	40.85
1976	**J Carter**	G Ford	51.05	-1	0	-1	0	3.81	7.63	5	61.5	47.5	140	171	49.41	56.21
1980	J Carter	**R Reagan**	44.84	1	1	0	0	-3.66	7.86	5	68	80	247	203	46.41	48.05
1984	W Mondale	**R Reagan**	40.88	-1	-1	0	0	5.42	5.25	8	46.5	88.5	190	287	37.24	30.60
1988	M Dukakis	**GHW Bush**	46.17	-1	0	-1	0	2.21	2.96	4	84.5	118.5	119	295	49.76	47.55
1992	**B Clinton**	GHW Bush	53.62	-1	-1	-1.25	0	2.95	3.31	2	102	83.5	280	319	46.15	48.88
1996	**B Clinton**	B Dole	54.74	1	1	0	0	3.26	2.03	4	70.5	87	354	320	52.74	51.38
2000	**A Gore**	GW Bush	50.26	1	0	1	0	2.01	1.64	7	86	88.5	311	303	49.03	51.28
2004	J Kerry	**GW Bush**	48.77	-1	-1	0	0	1.99	2.25	1	123	74.5	317	275	43.95	44.12
2008	**B Obama**	J McCain	53.69	-1	0	-1	0	-2.26	3.05	1	93.5	80	210	284	55.78	56.16

Note: The values of *P* for 1920, 1944, and 1948 before multiplication by zero are 16.535, 5.690, and 8.480, respectively, and the values of *Z* are 5, 14, and 5 (Fair 2010). Winner of popular vote in boldface. C^{dem} = charisma score of Democrat candidate; C^{rep} = charisma score of Republican candidate; L^{dem} = number of sentences of the Democratic candidate's speech; L^{rep} = number of sentences of the Republican candidate's speech; $\hat{V}_a$ is the out-of-sample prediction from the fair model (see Eq. 1); $\hat{V}_b$ is the out-of-sample prediction from our full model (see Eq. 3). Underlined values in $\hat{V}_a$ and $\hat{V}_b$ refer to wrong out-of-sample predictions (i.e., 7 for the Fair model and only 3 for Model 3; Model 4, not included above, had 3 wrong predictions).

TABLE 2

Regression Estimates (Study 1): Predicting vote-share

Variables	(1)	(2)	(3)	(4)
G·I	.67**	.64**	.71**	.74**
	(6.22)	(6.33)	(7.03)	(9.90)
P·I	-.65*	-.86**	-.47	-.21
	(2.31)	(3.08)	(1.50)	(.68)
Z·I	.99**	.90**	1.03**	1.10**
	(4.30)	(4.14)	(5.10)	(6.23)
DPER	2.92*	2.31†	2.17†	2.17**
	(2.18)	(1.83)	(1.86)	(2.66)
DUR	-3.41*	-4.32**	-3.96**	-3.71**
	(2.87)	(3.70)	(3.74)	(4.70)
I	-1.91	.39	-2.09	-3.64
	(.85)	(.17)	(.89)	(1.57)
War	5.06†	5.11*	7.12**	8.45**
	(1.99)	(2.19)	(3.10)	(4.34)
Charisma		.03	.13*	.19**
		(1.20)	(2.49)	(4.88)
Length		.01	.00	-.00
		(1.21)	(.27)	(.52)
Performance Signal			-.44*	-.68**
			(2.19)	(4.99)
Performance Signal*Charisma			-.02*	-.03**
			(2.26)	(6.13)
Constant	47.38***	48.05**	49.79**	50.71**
	(77.55)	(74.92)	(49.33)	(65.55)
Root MSE	2.50	2.29	2.04	1.57
R^2	.91**	.93**	.96**	.95**
ΔR^2 F-test		2.50	2.78	19.15**

**$p < .01$, *$p < .05$, †$p < .10$; $n = 24$; sample period: 1916-2008; *t*-statistics in parentheses; Columns 1-3 OLS regression models reporting unstandardized estimates for Models 1, 2, and 3 respectively. Column 4 reports results from the 2SLS model. The *ΔR^2 F-test* compares the change in *r*-square to the previous (constrained) model (for model 3, *ΔR^2 F-test* for the interaction alone is $F(1, 12) = 5.13$, $p < .05$). The significance of the parameters and joint tests remain unchanged with robust standard errors for Model 3. Note, because our dependent variable *V* is bounded we reestimated the models using a fractional logit model with robust standard errors; we also estimated the models using quantile (median) regression, which is robust to outliers, to approximate the conditional median. The overall pattern of results using these two modeling approaches produced the same substantive results (with higher levels of significance).

TABLE 3

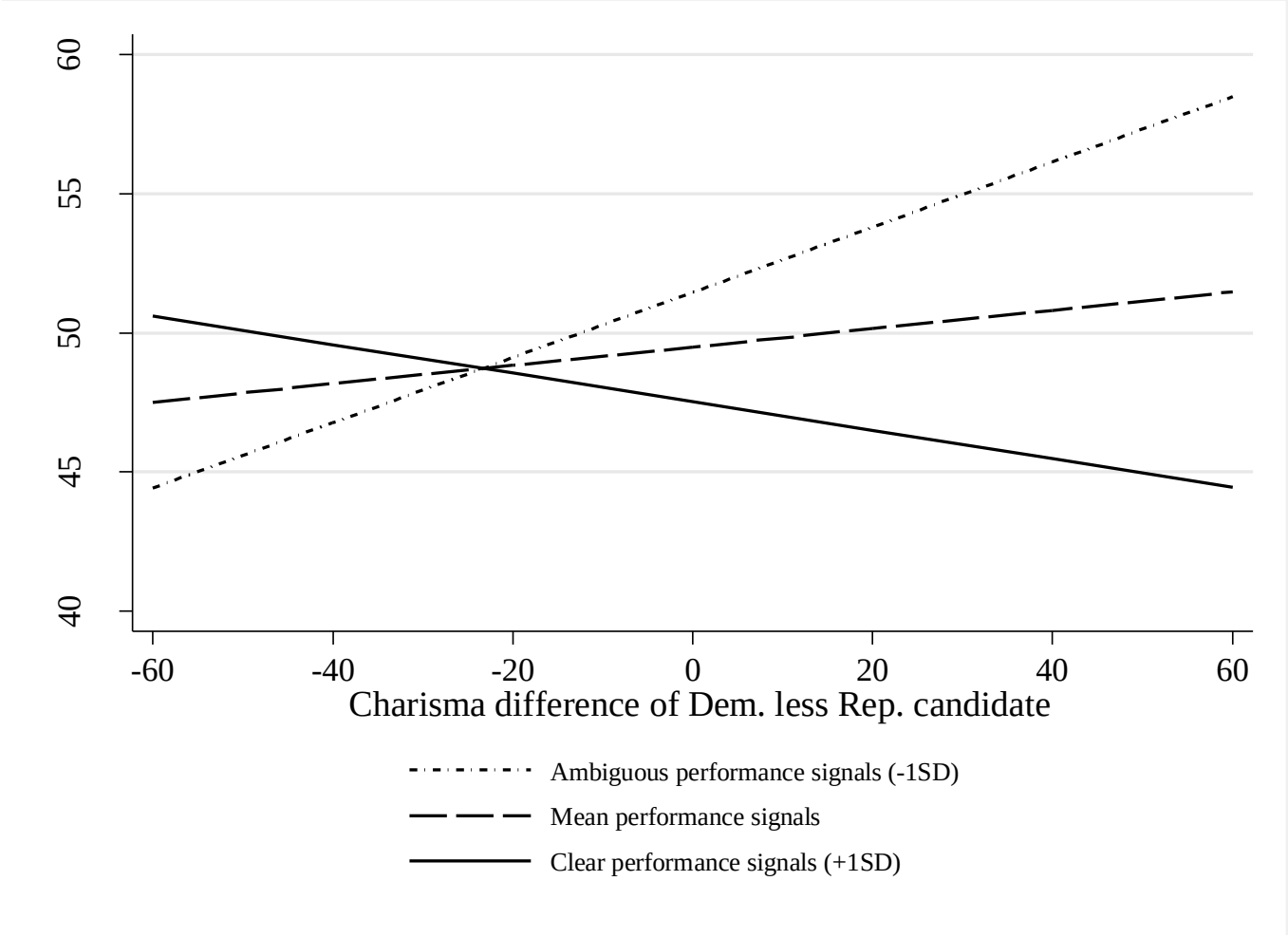
Regression Estimates (Study 2): Predicting CEO retention

Variables	(1)	(2)	(3)	(4)
Negative cue	-.13**		-.14***	-.08
	(2.97)		(3.45)	(1.51)
Positive cue	.36***		.35***	.47***
	(8.87)		(9.54)	(8.17)
Charisma		.39***	.38***	.50***
		(11.27)	(12.34)	(9.07)
Charisma*Neg. cue				-.13
				(1.62)
Charisma*Pos. cue				-.24**
				(3.30)
Constant	.47***	.36***	.29***	.24***
	(14.94)	(14.23)	(9.42)	(6.35)
R-squared	.17***	.15***	.31***	.32***
ΔR^2 F-test				5.46**

*** $p < .001$, ** $p < .01$, * $p < .05$; $n = 717$; robust t-statistics in parentheses; Columns 1-4 report unstandardized estimates for linear probability model for models predicting the selection decision. Change in r-square is for comparing models in Columns 4 and 3.

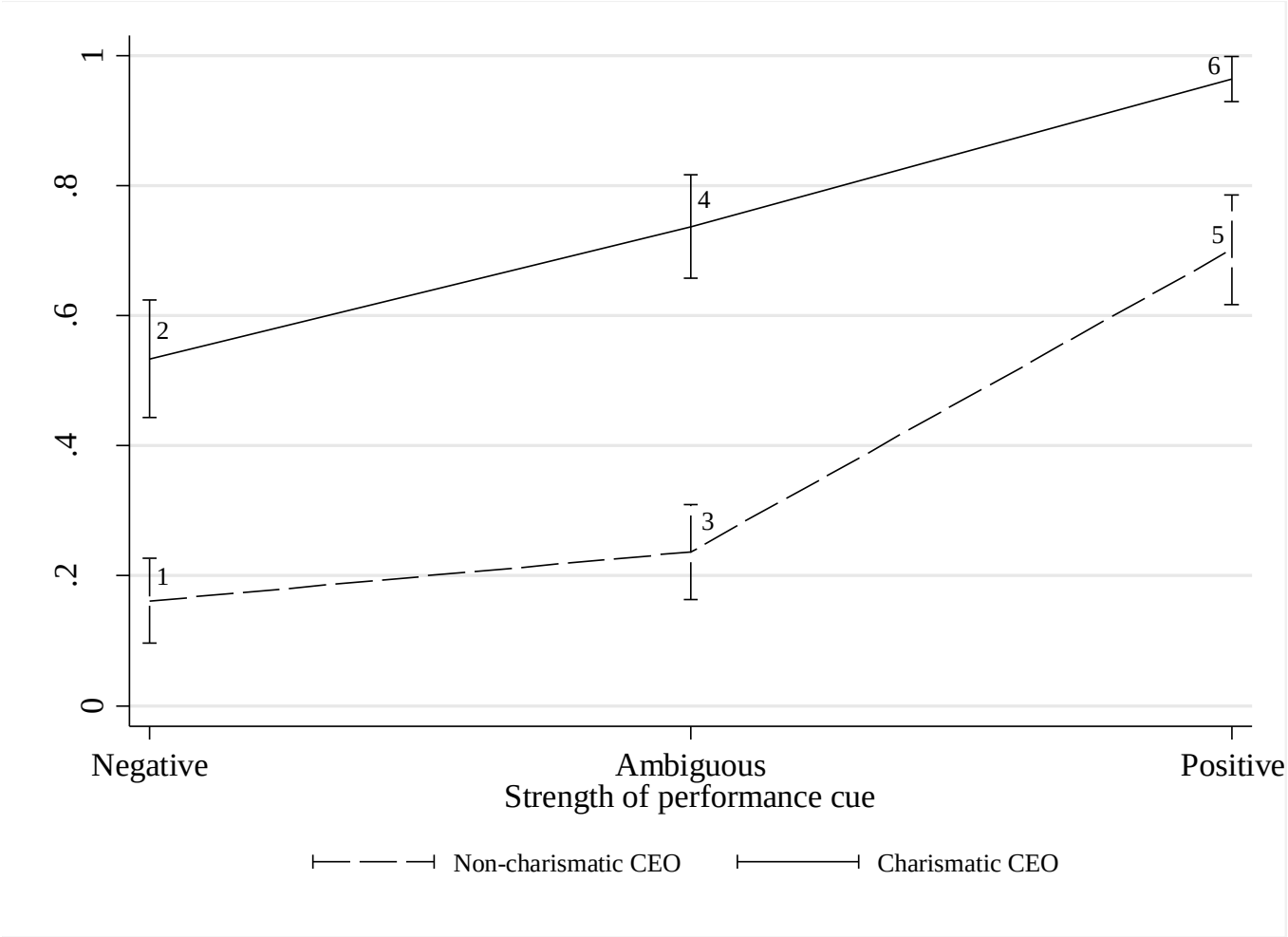
FIGURE 1

Interaction Between Strength Of Performance Signals And Charisma Difference Using Model 3 (Study 1)



Note: Prediction from the fitted values of Model 3; "Performance signals" refers to the predicted margin of victory using only economic and incumbency data (Model 1). The simple slope for "clear performance signals" is not significant; that of "ambiguous performance signals" positive and significant. The simple slope for "mean performance signals" is not significant.

FIGURE 2
Interaction Between Performance Signals And Charisma (Study 2)



Note: estimates are from a linear-probability model with heteroscedastic-robust estimate of the variance. Probit model predicted probability and marginal difference were the same as above. Error bars are 95% confidence intervals for the prediction. Numbers at prediction points are labels for the six experimental conditions.

Philippe Jacquart (jacquart@em-lyon.com) is an Assistant Professor of Leadership at EMLYON Business School. He received his doctorate in management from the Faculty of Business and Economics of the University of Lausanne and was postdoctoral researcher at The Wharton School at the University of Pennsylvania. His research focuses on leadership at the upper echelons of organizations.

John Antonakis (john.antonakis@unil.ch) is Professor of Organizational Behavior at the Faculty of Business and Economics at the University of Lausanne. He obtained his doctorate in Applied Management and Decision Sciences from Walden University and was postdoctoral researcher in cognitive psychology at Yale University. He currently does research on leadership, individual differences and research methods.