Group Identity and Agency Frictions: Evidence using Big Data

By Jitendra Aswani*

Does group identity alleviate agency frictions? Using a novel database of 3.78 million Indian last names with multiple identities to perceive the managers' and board members' identities, I provide evidence that the manager with a similar social identity as the board members relatively earns higher compensation. A firm benefits from paying the cost of in-group favoritism as the manager's reciprocation increases the long-term value and reduces agency frictions. A 1% increase in compensation due to group identity increases the firm's value by 1.8%. These findings imply that learning agent's specialization through homophily can enhance efficiency and welfare in principal-agent settings.

***KEYWORDS**: Group Identity, In-group Favoritism, Managerial Compensation, Firm Value, Statistical Discrimination.*

***JEL Codes**: D21, D22, G30, G40, J71*

* Aswani: jaswani@mit.edu

I am grateful for feedback, guidance, and comments provided by the Oliver Hart, Peter KniDaniel Green, Samuel Hanson, Elisabeth Kempf, Sanjai Bhagat, Tracy Wang, Yihui Pan (Discussant), Sris Chatterjee, N.K Chidambaran, Gayane Howakimian, Ashok Banerjee, Kevin Donovan, Nadya Malenko, David Hirshleifer, Thomas Chemmanur,Thomas Philippon, Hersh Shefrin, Ray Fisman, Alex Butler, Ross Levin, Ganet Gao, Cian Ruane, Vishan Gandhi Nigam (Discussant), Ata Can Bertay (Discussant), Sitian Liu, Gaurav Sood, Tristan Fitzgerald, Yeejin Jang, Svenja Dube, Elizabeth Blankespoor, Darya Yuferova, Iftekhar Hasan, Hongyu Shen, Md Moshi Ul Alam, and seminar participate at Contract Theory Workshop (Harvard Econ Dept.), Ph.D. Finance Workshop (Harvard Business School), Ph.D. Colloquium of Financial Management Association (FMA, 2020), Australasian Finance & Banking Conference (2020) at UNSW, Royal Economic Society Annual Metting 2021, North American Summer Meeting of the Econometric Society (2021) at Montréal, Asian Meeting of the Econometric Society (2021) at Malaysia, North East Universities Development Consortium (NEUDC) Meeting 2021, Financial Management Association 2022, American Finance Association (AFA) 2024 (paper session), and SIOE 2024 (University of Chicago). I also thank Mahazarin Banaji (Harvard Behavioral Lab), Mahesh Srinivasan (UC Berkeley Behavioral Lab), and Jay Van Bavel (NYU Behavioral Lab) for providing guidance on social psychology and social neuroscience literature. The author acknowledges the help from Fordham University for cloud computing and Harvard University's Alexander and Diviya Magaro Peer Pre-Review program.

"The relationship of agency is one of the oldest and commonest codified modes of social interaction."

-Stephen Ross (1973)

The separation of firm ownership and control creates agency conflicts—moral hazard, information asymmetry, and related challenges (Ross, 1973; Jensen and Meckling, 1976; Fama, 1980; Fama and Jensen, 1983; Hermalin and Weisbach 1991; Jensen 1993; Shleifer and Vishny 1997). While corporate governance research advocates incentive alignment (Yermack 1995; Hall and Liebman 1998; Mehran and Tracy 2001), agency frictions persist.[1] This paper offers a behavioral solution rooted in identity economics and social psychology: shared group identity between managers and board members raises compensation costs but also boosts long-run firm value, reducing agency frictions through informational gains about managerial specialization.

Consider two scenarios. A German board choosing between equally qualified German and Vietnamese candidates favors the German for cultural ease and communication. Or picture an eight-member board with one woman: she hesitates more discussing firm progress with a male manager than a female one. These illustrate that group identity fosters trust and ease decision-making frictions for firm benefit. Therefore, the board diversity should be with respect to manager's identities.

Neoclassical models long ignored social identities in corporate decisions, but Akerlof and Kranton (2000) sparked identity economics showing similarity breeds favoritism, dissimilarity discrimination.[2] I integrated Akerlof and Kranton's (2005) identity utility into Holmstrom's (1979) model and found that the in-group bias raises managerial pay but the firm value rises more than the cost of incremental pay due to the extra effort by the manager because of this in-group bias.

Most of the empirical studies on group identity are lab based which uses fMRIs and implicit association tests (IATs)[3] to capture implicit biases or field based using randomize control trails (RCTs), but the generalization of the results from these studies proves challenging. This paper fills

[1]The agency problem spans finance (Fama, 1980), economics (Spence and Zeckhauser, 1971), accounting (Fan and Wong, 2002), management (Douma, George, and Kabir, 2006; Peng and Jiang, 2010), and organizational studies (Schulze, Lubatkin, Dino, and Buchholtz, 2001).

[2]Identity models explain workplace discrimination (Akerlof and Kranton, 2000; Bertrand and Mullainathan, 2004), poverty (Akerlof and Kranton, 2000), contracts (Akerlof and Kranton, 2005), development (Basu, 2006), public goods (Croson, Marks, and Snyder, 2008), teams (Eckel and Grossman, 2005), banking (Fisman, Paravisini, and Vig, 2017), and investments (Morse and Shive, 2011; Grinblatt and Kelohraju, 2001). Favoritism splits into taste-based (Becker, 1957) and statistical (Arrow, 1972; Phelps, 1972).

[3]IAT measures unconscious associations via response speed (Greenwald et al. 1998).

that gap. Following Fershtman and Gneezy (2001) empirical design, I use India's surname-based stereotypes where last names reveal family lineage, native language, place, and caste[4]—fueling kinship cooperation, not distrust (Alesina and Giuliano 2013; Moscona, Nunn, and Robinson 2017).[5]

By merging Socio-Economic Caste Census (SECC) data for 474 million Indians with Linguistic Survey of India (LSI) data using a machine learning algorithm, I develop a novel identity database of 3.78 million surnames encoding language, caste, and place. Mapping this database to 2,324 Bombay Stock Exchange (BSE)/National Stock Exchange (NSE) non-financial firms (2004–2018), I find that shared identity (homophily) between a manager and the (corporate) board yields 6–8% higher managerial compensation—the in-group favoritism cost. Yet firm value such as market-to-book (MB) ratio and similar other measures rises 11–12%.[6] Residual firm value, which represents the post-compensation firm value, increases by 5.6% to 9.5%. These results suggest that the 1% pay rise in managerial compensation due to group identity boosts firm value by 1.8%, reducing agency frictions via specialization signals.

I test the findings for external validity and concerns about simultaneity and self-selection biases. These results are robust across various definitions of homophily (such as majority board ties and board-scaled measures), different firm value proxies (including sales growth, return on assets (ROA), return on equity (ROE), and logarithm of market capitalization), as well as profitability metrics (Bhagat and Bolton, 2008). To mitigate concerns about simultaneity bias, I use the two-stage least squares (2SLS) method following Palia (2000) and obtain the same results. Given the potential for self-selection bias, the results are tested using exogenous shocks, such as the deaths or retirements of directors, based on the Fracassi and Tate (2012) identification strategy. These results are robust and align with behavioral science experiments, which suggest that individuals (managers) exert additional effort for groups (boards/firms) they identify with (Cikara and Bavel, 2014).

Next, I check whether these results differ for firms with different organizational structures. The business group firms[7] show muted pay premiums but higher residual value—stronger reciprocation. Complex firms (measured as the square root of the number of subsidiaries) prioritize merit (negative

[4]Caste divisions in India: Brahmins, Kshatriyas, Vaishyas (upper); Shudras (ST), Harijans (SC) per SECC. Caste predicts economic outcomes (Prakash, 2015; Deshpande, 2000; Kijima, 2006).

[5]Natural identities show stronger fMRI-detected bias than induced. Huettel and Kranton (2012) distinguish natural identities (gender, race, caste, language, place) from induced (alumni, clubs).

[6]MB captures value/performance (Ritter, 1981; McDonald et al., 2008).

[7]Family-controlled, outperforming via capital access (Khanna and Palepu, 2000).

identity-pay link) yet gain value. Public sector undertakings (PSUs) show null effects—bureaucratic neutrality. Independent directors' ties raise no pay but cut value, validating Hwang and Kim (2009) and Bebchuk and Fried (2003) with natural identities. These results highlight exceptions in which the expected relationship between group identity, managerial compensation, and firm value does not hold.

Finally, I perform a series of checks to further validate the robustness of the results. These additional tests include using an alternative dataset for measuring group identity based on Indianchildnames.com, which contains approximately 2,225 surnames. I also examine different compensation measures, such as compensation relative to peers and a combination of salary and bonus. Other considerations include migration exclusions, controlling for influential CEO characteristics (as noted in Adams et al., 2005), induced ties (as discussed by Fracassi and Tate, 2012), geographic splits, founder CEOs, a homophily index based on all identities, and a continuous measure of homophily. All these tests confirm that in-group favoritism not only increases managerial compensation but also enhances firm value by more than the cost of in-group favoritism, due to incremental managerial reciprocation, suggesting that group identity helps mitigate agency frictions.

Contribution: This paper contributes to agency theory (Jensen and Meckling, 1976; Fama, 1980), particularly behavioral agency theory (Wiseman and Gomez-Mejia, 1998; Sanders and Carpenter, 2003; Pepper and Gore, 2012). Decades of research underpin agency conflicts: Arrow (1971) and Wilson (1968) highlight risk-sharing mismatches; Ross (1973) highlights incentives; Mitnick (1975) highlights institutional structures; Grossman and Hart (1983) highlight effort-consumption links that demand optimal pay-risk trade-offs. Unlike incentive fixes, this study proposes a behavioral remedy: group identity between the agent (manager) and the principal (board) theoretically and empirically reduces frictions via identity economics.

Executive pay debates (Edmans, Gabaix, and Jenter, 2017) pit "rent extraction" against "shareholder value" views (Fama, 1980). Here, identity-driven rents (hidden pay from board similarity) prove non-costly if managers boost value exceeding pay.[8] Board diversity studies yield mixed results on performance/value/monitoring—genetic (Delis et al., 2017), ancestral (Giannetti and Zhou, 2019), gender (Adams and Ferreira, 2009 vs. Gregory-Smith et al., 2014).[9] This shifts focus from

[8]See Edmans, Gabaix, and Jenter (2017) for compensation views.

[9]Delis et al. (2017) link genetic diversity to performance gains; Giannetti and Zhou (2019) to volatility; Adams and Ferreira

diversity to manager-board homophily reducing frictions.

In identity economics/discrimination, it advances theory/evidence on homophily's effects on managerial pay, firm value, agency frictions; proves statistical discrimination's Pareto efficiency. The literature covers implicit biases (Bertrand et al., 2005), taste-based biases (Becker, 1971), and statistical biases (Phelps, 1992; Schwab, 1986)—the latter is efficient via group signals sans prejudice. Prior work stresses taste distortions: Fisman et al. (2018) 39% selection boost from hometown ties; Fisman, Paravisini, and Vig (2017) credit/repayment gains from religion/caste; Gompers et al. (2016) venture capitalists (VCs) homophily harming early investments. This confirms Norman's (2003) efficiency via specialization information. Controlling for demographics (gender, education, age), results hold—though merit-homophily trade-offs warrant further study (Moisson and Tirole, 2020).

Lastly, this paper also contributes to emerging literature on big data in finance especially the applications in corporate governance. Unlike, the big data application in selecting the directors (Erel et al., 2021) or limiting the insider trading (Zhu, 2021), this paper shows how it can be used to learn the additional characteristics of a manager and the directors which can assist in curbing the agency frictions.[10]

Closest to Cai et al. (2013), Fracassi and Tate (2012), Huber, Lindenthal, and Waldinger (2021). Cai et al. (2013) find family ties raise Chinese private firm pay/roles—but this shows non-familial in-group bias boosts pay *and* value/friction reduction. Fracassi and Tate (2012) tie experience-based CEO-director links to value loss absent oversight; here natural identities raise value under strong governance (high promoter holdings). Huber et al. (2021) quantify ethnic discrimination's stock/dividend/return costs; this nets favoritism costs/benefits: 6–8% pay rise, 11–12% value gain, 5.6–9.5% residual value—extra effort reciprocation proves net beneficial, reducing agency frictions for long-term value.

Section 2 discusses the theoretical model and its predictions. Section 3 elaborates on the descriptive data and variables. Section 4 explains the methodology and empirical specification. Section 5 discusses the results. Section 6 concludes the article.

(2009) find gender diversity hurts performance despite monitoring benefits; Gregory-Smith et al. (2014) find null effects.

[10]For review of papers on big data applications in finance, refer Goldstein, Spatt, and Ye (2021).

2. THEORETICAL MODEL

The model's motivation comes from Ross (1973), Holmstrom (1979), and Edmans, Gabaix, and Jenter (2017).[11] According to this model, the board of directors hires a manager to run the firm. Firm value is denoted by $V(a, S, \epsilon)$, which increases with the manager's action a and firm size S. The function b(S) measures the effect of effort on firm value for a firm of size S. Suppressing the dependence on S and ϵ for simplicity, firm value is treated as a function of managerial effort only.

$$V(a) = S + b(S)a + \epsilon$$

The manager receives compensation $c(V)$ that is contingent on firm value. The board's limited liability is assumed, which means $c(V) \leq V$; that is, manager's pay should not be more than firm value. Similarly, the limited liability of the manager is assumed, which means $c(V) \geq 0$ and reservation wage, $w \geq 0$.

The model is modified by introducing group identity into the preference function by assigning different weights to the managerial effort component. It defines the agent utility function using the identity. This type of agent utility function is suggested by Akerlof and Kranton (2005), and a similar method of incorporating group identity into preferences is adopted in Chen and Li (2009) and Chen and Chen (2011). This theoretical result is consistent with the literature on minimum-effort coordination games.[12] The complete model for this study is presented in **Appendix B**.

The following model setup is from Holmstrom (1979):For simplicity, I consider risk neutrality and additive preferences and therefore, $u(x) = x$ and $v(c) = c$. For details, refer to Edmans, Gabaix, and Jenter (2017, Pp-39).

$maxE(V - c)$ (Principal's Problem)

s.t. $E(u(c - g(a))) \geq E(u(w))$ (Participatory Constraint)

$c = \phi + \theta V$ (Managerial Compensation)

[11] In the first-order model of principal and agent used by Holmstrom (1979), the agent's utility maximization problem is replaced with the first order constraint and it is based on the assumption that an optimum exists and is differentiable. The principal and agent both are risk-neutral in this setting. See Mirrlees (1976) and Rogerson (1985) to know more about first order and other related models.

[12] For more details on this topic, refer to Camerer (2003).

$a^* \epsilon$ argmax $E(u(c - g(a)))$ (Agent's Problem)

$V = S + b(S)a + \epsilon$ (Firm Value)

g(a): The cost function of the manager's efforts and it is increasing and weakly convex. ($\frac{1}{2}ga^2$)

ϕ : Fixed component of the manager's compensation (Salary)

θ: Variable component of the manager's compensation

w: Reservation wage

ϵ: Random variable with a normal distribution. $\epsilon \sim N(0, \sigma^2)$

Agent utility function from Akerlof and Kranton (2005)[13]

$U(c, a, m) = c\text{-}a + I_m - t_m|\hat{a} - a|$

$\hat{a}$: Ideal effort

I_m : Additional utility due to group identity (or homophily)

Complete Homophily (or in-group): When a manager and board have the similar social identity.

In-complete Homophily (or out-group): When a manager and board have different social identities.

m: Variable to categorize different social identities such as native language, native place, or caste.

Consider the first-best benchmark in which effort is observable. Let a^* denote the effort level that the principal wishes to implement. The principal can simply direct the CEO to exert effort a^*, and so we can ignore the agent's problem. It is easy to show that the CEO is given a constant wage $c(V) = \bar{c}$, as this leads to efficient risk-sharing. The participatory constraint becomes $\bar{c} \geq w + g(a)$. Therefore, the principal maximizes,

$$\max E(V(a) - c)$$
$$\text{s.t. } c \geq w + g(a)$$

This defines the first-best effort level as

[13] I_m is identity utility, which is considered as 0 for simplicity (see, Akerlof and Kranton (2008), pg-213), and $t_m|\hat{a} - a|$ is disutility for diverging from the ideal effort level, denoted by $\hat{a}$. For in-group member, $|\hat{a} - a| > 0$ and for out-group member, $|\hat{a} - a| < 0$.

$$g'(a^*_{FB}) = b(S)$$
$$a^* = \frac{b(S)}{g}$$

This shows that the principal trades off the marginal increase in firm value from effort, b(S), against the CEO's marginal cost, $g'(a^*_{FB})$. Nevertheless, the effort is usually unobservable, so the first-best outcome is not possible. This is the effort level, excluding the impact of identity, as shown in prior literature. Next, I use the agent utility function with identity to check the effort level for complete homophily (in-group) and for incomplete homophily (out-group), and compare them.

Proposition. *The managerial efforts are higher in the case of complete homophily than incomplete homophily.*

Proof:At equilibrium, the managerial effort under complete homophily is
$a^*_{in-group} = \frac{1}{1-t_m}[\frac{\theta b(s)}{k} - t_m\hat{a}]$.

Similarly, in the case of incomplete homophily, the managerial effort is
$a^*_{out-group} = \frac{1}{1+t_m}[\frac{\theta b(s)}{k} + t_m\hat{a}]$.

$a^*_{in-group}$ and $a^*_{out-group}$ depends on the managerial effort which depends on three factors (sensitivity of effort to identity (t_m), sensitivity of the firm value to effort ($b(S)$), and sensitivity of the managerial compensation (variable component) to firm value(θ)) and their dynamics. Next, I simulate the model while restricting the parameters to compare $a^*_{in-group}$ and $a^*_{out-group}$. The sensitivity parameters, i.e., (t_m), $b(S)$, and θ are restricted to values between 0 and 1. These restrictions are intuitive considering that firm value cannot exceed effort, compensation cannot exceed firm value, and effort is not completely driven by group identity. By combining the above two equations and rearranging,

$$(1 - t_m)a^*_{in-group} + (1 + t_m)a^*_{out-group} = \frac{2}{k}(b\theta)$$

As b represents sensitivity of the firm value to effort and θ represents sensitivity of compensation to firm value, applying the chain rule in derivatives, the above equation can be expressed as

$$(1 - t_m)a^*_{in-group} + (1 + t_m)a^*_{out-group} \cong \frac{dc}{da}$$

This shows that the impact of a manager's effort on compensation depends on the identity overlap between a manager and the board and the sensitivity of the manager's utility to identity, i.e., (t_m). In other words, the impact of a manager's effort on compensation is the weighted average of the manager's in-group and out-group efforts weighted by the sensitivity of manager's utility to identity.

The $\frac{dc}{da}$ has an upper bound of 1 (complete effort converts into compensation) and a lower bound of 0 (effort does not convert into compensation at all).

$$0 \leq (1 - t_m)a^*_{in-group} + (1 + t_m)a^*_{out-group} \leq 1$$

The lower bound provides the real solution as in complete homophily, and for incomplete homophily, both $a^*_{in-group}$ and $a^*_{out-group}$ would be 0 and cannot be the solution. Therefore, from the upper bound,

$$(1 - t_m)a^*_{in-group} + (1 + t_m)a^*_{out-group} = 1$$

For complete homophily, $a^*_{in-group}$ is $\frac{1}{1-t_m}$ and for incomplete homophily, $a^*_{out-group}$ is $\frac{1}{1+t_m}$. As t_m is positive and lies between 0 and 1, this shows,

$$a^*_{in-group} > a^*_{out-group}$$

This implies that, under complete homophily, the manager's effort at equilibrium exceeds the effort under incomplete homophily.

Lemma 1. *Higher managerial efforts lead to higher firm value at equilibrium for complete homophily compared to the case of incomplete homophily.*

Proof. As

$a^*_{in-group} > a^*_{out-group}$

Therefore, at equilibrium (for same b(S)),

$V^*_{in-group} > V^*_{out-group}$

This implies that a firm in which there is complete homophily has a higher value than a firm with incomplete homophily, this derives from the difference in managerial efforts in both cases.

Lemma 2. *Higher managerial efforts and firm value for complete homophily lead to higher managerial compensations at equilibrium compared to incomplete homophily.*

Proof. From the participation constraint,

$$E[c] = \begin{cases} E[c]_{in-group} = w + \frac{1}{2}k(a^*_{in-group})^2 - \frac{\nu}{2}\theta^2_{in-group}\sigma^2, \text{ For complete homophily} \\ E[c]_{out-group} = w + \frac{1}{2}k(a^*_{out-group})^2 - \frac{\nu}{2}\theta^2_{out-group}\sigma^2, \text{ For in-complete homophily} \end{cases} \quad (3)$$

$$\text{Principal's utility} = E[V - c] = \begin{cases} S + b(S)a^*_{in-group} - E[c]_{in-group} \\ S + b(S)a^*_{out-group} - E[c]_{out-group} \end{cases}$$

$\theta^*_{out-group} > \theta^*_{in-group}$

As, $a^*_{in-group} > a^*_{out-group}$

Therefore, from eq(3), $E[c]^*_{in-group} > E[c]^*_{out-group}$

This theoretical prediction implies that the manager's total compensation is higher for a firm with group identity. This incremental compensation is the cost of the in-group favoritism to the firm.[14]

In sum, both firm value and manager's total compensation (and salary) are higher when the manager and the board share the same identity (i.e., due to group identity or homophily). Although

[14]One limitation of this model is that the cost of managerial effort can be both pecuniary and non-pecuniary in nature. Following Holmstrom and Milgrom (1987), only the pecuniary cost of effort is considered.

additional managerial compensation is a cost to the firm due to in-group favoritism, this cost is offset by the manager's extra efforts to increase firm value. The complete model with proofs is available in Internet Appendix IA-M. In the next section, I verify whether these theoretical predictions hold in empirical analysis.

3. DATA AND DESCRIPTIVES

3.1. Data Sources

Surname data serve as the key to identifying managers and directors. My approach follows the literature in sociology, anthropology, and social neuroscience. The identity database, which links Indian surnames to native place, native language, and caste, is constructed using SECC (2011) data on 474 million Indians and the LSI . Internet Appendix Table IA2 Panel A and Panel B provide a preview of the SECC data with the procedure to extract surnames. The SECC data, after filtering, include information on Indian surnames, their native place (district (county) and state), and caste. The LSI data provide information on the most populous language in the district.[15]By merging these databases using machine learning algorithms, I develop a novel database of 3.78 million Indian surnames with multiple identities.[16] This database contains information on Indian surnames, native language, caste, and native place. Figure 1 shows the state-wise, language-wise, and caste-wise distribution of these 3.78 million Indian surnames.

[Insert Figure 1 Here]

To externally validate the surname database developed using the SECC and LSI data, I use an alternative surname dataset obtained from "Indianchildnames.com." This dataset contains 2,225 unique common Indian surnames with their native language and native place.

ProwessDx (CMIE), Global Compustat, and the BoardEx databases provide manager and board member names, executive compensation data, firm fundamentals, board characteristics, and managerial demographics. Extracting surnames from these names was not straightforward due to the various designations and their patterns, such as Mr./MR/Mr/(MR)/(Mr.)/(Mr), Dr./Dr/(Dr),

[15]The Appendix provides details on constructing the identity database using the SECC and LSI datasets.

[16]This is inline with Ancestory.com database which suggests there are almost 4 million unique last names in India. See, `https://blogs.ancestry.com/cm/indian-surnames-a-mix-of-culture-caste-and-religion/`. However, Ancestory.com lacks the identities of these last names.

Prof., Sh./Shri, I.A.S., in addition to the army designations (Maj. (Major), Gen. (General), Col. (Colonel), navy designations (Vice Admiral, Commodore) and many others. The algorithm employed here differs from those used to extract surnames from the SECC database. Table IA2 Panel C shows the 200 most common Indian surnames among managers with their ranking. For instance, 'Gupta', 'Jain', 'Agarwal', 'Patel', and 'Shah' are the top five surnames among Indian managers. I merge ProwessDx and identity database using unsupervised machine learning with the surname as an identifier.[17] Section 4.1. provides a detailed discussion of the methodology.

The sample comprises a panel of non-financial firms listed on two leading Indian stock exchanges - the Bombay Stock Exchange (BSE) and the National Stock Exchange (NSE), spanning from 2004-2018. There are 2,324 firms in the sample and 28,026 firm-year observations. Table 1 Panel A shows the filtering of observations from the ProwessDx database to form the sample. Since the concept of CEO is relatively new in India, the sample also includes CEO-equivalent positions such as president, managing director, and chairman.[18] The yearly distribution of observations is provided in Appendix Table IA1 and shows that the sample is balanced.

The main dependent variables are managerial compensation, firm value, and residual firm value (i.e., firm value conditional on the previous year's managerial compensation). The compensation data from ProwessDx include total compensation, salary, bonus, perks, contributions to provident fund (PF), and sitting fees. Figure 3 illustrates the distribution of total compensation and its components. [19]

The main managerial compensation variable is the logarithm of total compensation. However, for external validity, I also use an additional set of compensation variables such as the logarithm of salary, bonus, sitting fees, and total compensation relative to the industry average. Logarithm forms are used due to skewness in the distribution. The firm value is measured as the ratio of the market value of equity to the book value of equity (*Market to Book ratio*). I use the market-to-

[17]Unsupervised machine learning algorithm is used because of some unknown designations in the ProwessDx and the human error in the census data as some of the surnames in the census data are given as "Kuma r", "Sharm a", "Ra i", "S harma", and similar others.

[18]Bertrand and Hallock (2001) use the CEO, chairman, Managing Director positions for the top corporate jobs. One can argue that the chairman is also a board position; therefore, including it can make identification complicated. To resolve such a concern, an additional robustness test is performed by excluding the observations where the manager is chairman (see Table IA10).

[19]Total compensation paid to an executive can be either salary or a combination of salary and bonus or a combination of salary, bonus and sitting fees. The main component of the total compensation is salary, but salary with bonus and sitting fees cover 90 % of total compensation.

book ratio because it is a well-established measure for firm value (see Ritter (1981) and Downes and Heinkel (1981)) as well as firm performance (see McDonald, Khanna, & Westphal, 2008; Morck, Shleifer, & Vishny, 1989, and others). To show how firm value is affected by homophily even after paying the cost of in-group favoritism, the residual firm value is computed. The residual firm value (*Residual Firm Value*) is measured as the residual from regressing firm value on the previous year's managerial compensation. The underlying theory for using it as a proxy of managerial reciprocation is that the determinants of residual firm value should only be correlated with firm and governance characteristics. If homophily explains part of this residue, it is due to the extra effort the manager invests in the firm. For the robustness check, sales growth (*Sales Growth*) and return on equity (*ROE*) are used as additional proxies for firm value. *Sales Growth* is calculated as the growth in sales compared to the previous year. *ROE* is the ratio of the net income to the total book value of equity.[20] The main independent variables are three homophily (group identity) measures - homophily language (*Homophily Language*), homophily place (*Homophily Place*), and homophily caste (*Homophily Caste*). The homophily measures are dummy variables that take on the value of 1 if the manager and at least one board member share the same social identity. Otherwise, it is 0. For robustness, I also consider alternative measures of homophily, such as mapping managerial identity to the identity of the majority of board members, and the percentage of board members' with which managerial identity matches.

For controls, firm size (*Size*), firm age (*Firm age*), institutional ownership (*Institutional Ownership*), promoters' ownership (*Promoters Ownership*), leverage (*Leverage*), volatility (*Volatility*), Big 5 auditing firms (*BIG5*), board size (*Board Size*), number of board meetings (*Number of Meetings*), and return on assets (*ROA*) are used. The data used to construct these variables are sourced from ProwessDx (CMIE). *Size* is measured as the log of market capitalization. Market capitalization is expressed in USD. Nevertheless, the way firm size is measured can affect the association between compensation (or firm value) and firm size (Luo, Liu, and Tripathy (2021)); I also conduct a robustness check using the log of assets as a measure of firm size (refer to Table IA13 for results). *Firm Age* is the firm's age since the year of incorporation. *Promoters Ownership* is the percentage of promoters' equity holdings relative to the total equity issued by the firm. *Institutional Ownership*

[20]Results of external validity tests are provided in Internet Appendix Table IA 8.

is the percentage of institutional investors equity holdings relative to the total equity issued by the firm. *Leverage* is the ratio of total debt to total assets. *ROA* is measured as the ratio of net income to total assets. The total debt and total assets are in millions of USD. *Board Size* is the total number of directors on board. *Volatility* is the return volatility. *BIG5* is a dummy variable that takes the value 1 if the firm's auditor is one of the BIG 5 auditing firms - Deloitte & Touche, Ernst & Young, Pricewaterhouse Coopers, Grant Thornton, or KPMG, and 0 otherwise.[21] *Number of Meetings* is the number of board meetings in a year. The institutional ownership, promoter ownership, board size, number of board meetings, and the BIG 5 constitute the internal and external governance characteristics.

As managerial characteristics such as age (*Age*), gender (*Gender*), and education (*Education*) are not available in ProwessDx, this information is gathered from the BoardEx database. Due to data limitations, these characteristics are not available for all managers and this constraint limits the number of observations to 3,542 for the sub-sample analysis.

3.2. Descriptives

Figure 2 presents maps showing the distribution of managers and directors by their native place in India. These maps indicate that, although directors and managers come from all over India, most of them are originally from southern India.[22]

[Insert Figure 2 Here]

The pie chart in Figure 2 indicates that most managers and directors in India belong to upper castes, and only 19-21% come from a lower caste. This distribution is more skewed at the managerial level than at the director's level. Figure 2 also shows the distribution of managers and directors by their native language and indicates that most managers and directors speak Marathi, Tamil, or Western Hindi. This group comprises 14-18% of the sample. Other common native languages spoken by managers and directors are Malayalam, Gujarati, and Kannada. Native place, native language, and caste form the identities of managers and directors. [23]

[21] In India, the Big 5 accounting firms also work through their affiliated firms. The information on the affiliated firms has been gathered from the BIG5 auditing firms websites.

[22] Indian Media has also pointed out that half of the top ten Indian billionaires figuring in the list of Bloomberg Billionaires are Gujaratis (a state in northern India) and one in three CFOs of Nifty index companies is south Indian. For details check, https://economictimes.indiatimes.com/news/politics-and- nation/stereotypes-dominate-indias-rich-list/articleshow/46857906.cms?from=mdr

[23] Appendix Table IA3 shows the industry-wise distribution of the homophily variables. Certain industries such as metals and

[Insert Figure 3 Here]

Table 1 Panel B reports summary statistics for different components of managerial compensation in millions of USD. The average (maximum) total compensation of a CEO or CEO-equivalent position in India is 192,000 USD (2,461,000 USD).[24] The average (maximum) salary, bonus, sitting fees, and perks are 132,000 (1,116,000), 262,000 (3,663,000), 3,000 (23,000), and 17,000 (138,000) USD. Unlike in the US, the mean variable pay in India is 17% of total remuneration. Summary statistics and distribution of managerial compensation are comparable to Hill and Thomas (2014).

Table 1 Panel C shows the summary statistics of the variables used in the analysis. The mean (max) of log total compensation (*Log Compensation*) and industry-relative compensation (*Industry Relative Compensation*) are 10.611 (13.634) and 0.003 (3.752). The mean(max) of log salary (*Log Salary*) is 10.4 (13.493). Log compensation, log salary, industry-relative compensation, and logs of salary, bonus, and sitting fees are used to demonstrate how executive compensation is affected by in-group favoritism.

[Insert Table 1 Here]

To show how firm value is affected by group identity (homophily), market-to-book ratio (*Market to Book ratio*) is used. However, for external validation, sales growth (*Sales Growth*) and the sum of the market value of equity and book value of debt (*Log Market Value*) are used as proxies for firm value. The mean (max) values of *Market to Book ratio*, *Sales Growth*, *Log Market Value* are 2.238 (19.260), 0.145 (7.067), and 4.180 (9.610), respectively. The mean (max) of *Residual Firm Value* is -0.014 (8.470). Residual firm value captures the impact of managerial reciprocation of in-group favoritism.

The three main independent variables are homophily language, homophily place, and homophily caste. These variables are dummy variables and take on the value of 1 if the board and the manager share the same identity, i.e., native language, native place, or caste. In 40-42% of the sample's observations, the board and the manager speak the same native language or come from the same native place. This factor indicates that the sample is balanced regarding these identities.

mining, consumer durables, media and print, wine cables, wood products, fertilizers, and similar sectors have more homophily between the manager and the board than other industries. There can be a concern that homophily constructs might be time-invariant, but it is not valid, as these constructs change with time due to manager or director turnover.

[24]To provide a sense of these numbers compared to US executive compensation, for the period 1982 through 1988, the average salary and bonus for the CEOs of large public companies were 843,000 USD. (Jensen and Murphy, 1990).

However, the homophily caste variable shows skewness. In 89% of the observations, the board and the manager belong to the same caste. As upper-caste individuals dominate board and managerial positions, the skewness in homophily caste results from it. Table 1 Panel D shows the distribution of homophily variables in detail. The control variables include firm size (*Size*), firm age (*Firm age*), institutional ownership (*Institutional Ownership*), promoters ownership (*Promoters Ownership*), leverage(*Leverage*), volatility (*Volatility*), Big 5 auditing firms (*BIG5*), board size (*Board Size*), the number of board meetings (*Number of Meetings*), and return on assets (*ROA*). In the sample, the average firm size, age, and leverage are 16.9, 34 years, and 15.9%, respectively. In India, promoters (founders) own most shares, and the ownership distribution clearly reflects this. The average promoter ownership and institutional ownership are 54% and 9.5%, respectivelyThe definitions of the variables are provided in Appendix A.

For board characteristics, the average board size is 10 members. On average, 27.5% of observations involve BIG5 auditors, and Indian boards conduct 8 meetings per yea managerial demographics, the minimum (average) age (*Age*) of a manager is 40 (64) years old, and the average level of educationualification (*Education*) is a Master's degree. The average number of women on the board (*Gender Diversity*) is only 9-10%. Panel D shows that the sample is balanced for homophily language and homophily place. Homophily caste is skewed as most managers and directors in India belong to the upper caste. This implies that either lower-caste individuals do not get opportunities to reach such posts or are crowded out by lobbying by upper-caste managers and directors. Testing the implications is beyond the scope of this article, but they can be examined in future research. The univariate analysis of total compensation, firm value, and residual firm value based on homophily constructs is provided in the Online Appendix Table IA4. Univariate analysis supports the main argument that there is a significant difference in executive compensation, firm value, and residual firm value when a manager shares the same identity as the board compared to when they do not. The covariate balance between the control variables for the three homophily constructs is provided in the Online Appendix Table IA5. It confirms that there is not much difference in covariates between the comparison groups.

4. METHODOLOGY

4.1. Surname Matching Using Unsupervised Machine Learning

The lack of a common identifier between the identity database and ProwessDx necessitates the use of a text similarity algorithm from the unsupervised machine learning literature. Excluding less common designations from the names of managers and directors is difficult, as doing so could reduce the number of observations and generate a measurement error.[25] To address this issue, a text similarity algorithm is used for merging. The Jaccard algorithm, a well-known text similarity algorithm, is used to discern a manager's identity and the remaining board members.[26] For matching, surnames are first divided into bi-grams, which are then used to compute similarity scores. The Jaccard similarity score is the ratio of the number of common bi-grams between two surnames to the total number of bi-grams across the two surnames.

$$\text{Jaccard Similarity} = (A \cap B)/(A \cup B) \tag{1}$$

Here, A and B are two surnames.

It can be understood by computing the Jaccard similarity score of 'Smith' and 'Thomas'. For instance, the bi-grams of surnames such as 'Smith' and 'Thomas' are $\{sm, mi, it, th\}$ and $\{th, ho, om, ma, as\}$, respectively. As there is only one common bi-gram –'th', between these two surnames, '$A \cap B$' is 1. The total number of unique bi-grams between them –'$A \cup B$', is 8. This means that the Jaccard similarity score is 0.125, which is quite low because the surnames in this example are too discrete. For exact matching, the similarity score threshold is set at 0.9 and above. To confirm the matching, a manual check was also conducted.

4.2. Empirical Specification

To examine the role of group identity (or homophily) in the executive compensation setting, I ran a regression of the log of total compensation on homophily constructs, while controlling for

[25] For instance, 'Sharma' is a well-known Indian surname. However, the data has various other forms of this surname such as'S harma', 'sharm a', 'Sharmams' etc. These erroneous surnames have formed due to man-made errors when doing the entry in the census report.

[26] Another common text similarity algorithm is the Cosine algorithm. These two algorithms give different results if the two documents or sentences in question are too different in length. In this case, both give similar results as the comparison is between the two surnames rather than between the two documents.

firm characteristics, governance characteristics, and managerial demographics. [27] The empirical specification for this test is:

(2)

$$LogCompensation_{i,t} = \alpha_{i,t} + \beta_1 Homophily_{i,t} + \beta_2 X_{i,t} + \beta_3 L_{i,t} + \beta_4 Z_{i,t} + \gamma_{firm} + \gamma_{year} + \gamma_{designation} + \epsilon_{i,t}$$

For external validation, in place of a log of total compensation, I also use log of salary, total compensation relative to the industry peers, and log of salary plus bonus plus sitting fees. Homophily (or group identity) is a dummy variable that takes a value of 1 if a manager and at least one board member share the identity; otherwise, 0. Separately, I tested three natural identity traits: native language, same caste, and native place. For external validation, I also measure homophily differently. This paper uses a simple baseline homophily measure to estimate in-group favoritism.

The detailed literature surveys on executive compensation by Core, Guay, and Larcker (2002), Frydman and Jenter (2010), and Edmans, Gabaix, and Jenter (2017) suggest that the main determinants of executive pay are firm fundamentals, governance characteristics, and manager demographics. Following that suggestion, X is a vector of firm characteristics such as firm size, firm age, return on assets (ROA), leverage, and volatility. L is the vector of governance characteristics such as board size, number of board meetings, one of the BIG 5 auditors, institutional ownership, and promoter ownership. Z is the vector of the manager's demographics (such as education, age, gender, and others). To account for unobservables, firm fixed effects γ_{firm} and year fixed effects γ_{year} are used. As results can be affected due to the managerial position, managerial designation fixed effects ($\gamma_{designation}$) are also used. Because managerial demographics are unavailable for the full sample, the subsample results control for managers' age, gender, and education are provided in Internet Appendix Table IA6. The following empirical specification is implemented to check the impact of homophily on firm value:

(3)
$$FirmValue_{i,t} = \alpha_i + \beta_1 Homophily_{i,t} + \beta_2 X_{i,t} + \beta_3 L_{i,t} + \gamma_{firm} + \gamma_{year} + \nu_{i,t}$$

The coefficient of interest is β_1. As homophily is based on language, place, and caste, β_1 can be

[27] As the number of observations available for the managerial demographics is limited; a sub-sample analysis is provided controlling for the vectors of the managerial demographics.

seen as $\beta_{Language}$, β_{Place}, and β_{Caste}. Firm value is measured by the market-to-book ratio. X and L are the firm characteristics and governance characteristics listed earlier. Based on prior literature, all βs of firm fundamentals and governance are expected to be positive except for volatility and leverage (Hovakimian, Hovakimian, and Tehranian (2004)). Next, I decompose the firm value explaining last year's managerial compensation and the residual firm value. The impact of homophily on the residual firm value is examined to check the reciprocation of the manager toward in-group favor. The residual firm value is an error term from the regression of firm value on the previous year's managerial compensation.

$$FirmValue_{i,t} = \omega_i + \phi_i LogCompensation_{i,t-1} + \mu_{i,t} \tag{4}$$

The residual firm value (*Residual Firm Value*) is μ. It captures firm value conditioned on managerial compensation. An increase in residual firm value proxies a greater managerial effort to increase firm value. To test whether residual firm value (*Residual Firm Value*) is positively associated with the homophily variables, the following regression specification is used,

$$ResidualFirmValue_{i,t} = \alpha_i + \beta_1 Homophily_{i,t} + \beta_2 X_{i,t} + \beta_3 L_{i,t} + \gamma_{firm} + \gamma_{year} + \nu_{i,t} \tag{5}$$

The interpretation and expectations of βs are the same as for the firm value regression. If firm characteristics and governance characteristics alone explain residual firm value, we expect β_1 to be statistically insignificant. However, if β_1 is positive and significant, the manager exerts an additional effort to increase the firm value more than what is paid for. This change occurs due to shared group identity with the board.

4.3. Endogeneity Concerns and Identification Strategy

Two econometric issues with the above described empirical specifications must be addressed. First, the executive compensation and firm value are endogenously determined, and therefore, the described empirical specifications may not capture it completely. Second, as the firm hires a particular manager for an unobservable reason, there is a self-selection problem. Following Palia's (2000) example, I use the two-stage least square method to mitigate the simultaneity issue. Similarly, to

minimize self-selection bias, following the example of Fracassi and Tate (2012), I use the death or retirement of a director as an exogenous shock to board identity and the homophily constructs. To implement the two-stage least square method, the fitted value of $LogCompensation_{i,t}$ from eq (6) is used in eq (7).

$$LogCompensation_{i,t} = \alpha_{i,t} + \beta_1 Homophily_{i,t} + \beta_2 FirmValue_{i,t} + \beta_3 X_{i,t} + \beta_4 L_{i,t} + \beta_5 Promoter_Dummy + \gamma_{firm} + \gamma_{year} + \gamma_{designation} + \epsilon_{i,t} \quad (6)$$

$$FirmValue_{i,t} = \alpha_i + \beta_1 Homophily_{i,t} + \beta_2 LogCompensation_{i,t} + \beta_3 X_{i,t} + \beta_4 L_{i,t} + \gamma_{firm} + \gamma_{year} + \epsilon_{i,t} \quad (7)$$

The coefficient of interest is β_1. The promoter dummy (*Promoter Dummy*) is used as an instrument; it equals 1 if the manager is also a firm's promoter and 0 otherwise. In the first stage, the estimated regression is the logarithm of the total managerial compensation on the promoter dummy as an instrument and all the firm and governance level controls. In the second stage, the estimated regression looks at firm value concerning three group identity measures, controlling for the fitted value from the first stage regression with other controls.

As the board can select a manager of the same identity to reduce frictions in corporate decision-making (Damaraju and Makhija, 2018), this can affect the causal inference of results. I use a similar identification strategy to mitigate this concern in Fracassi and Tate (2012). The director's death or retirement is used as an exogenous shock as these events are unrelated to the firm's conditions. Voluntary retirement is not included, as it may arise from firm-related issues.[28] As ProwessDx does not have information on the director's date of death, this information is collected from the BoardEx database. The empirical specification for this test is

$$Y_{i,t} = \alpha_i + \beta_1 Shock + \beta_2 Homophily_{i,t} + \beta_3 Shock\text{x}Homophily_{i,t} + \beta_4 X + \beta_5 L + \gamma_{firm} + \epsilon_{i,t} \quad (8)$$

[28]There is a possibility that even after the director's death or retirement, the board identity remains the same. Therefore, I ran the additional test as diff-in-diff between two groups where, in one group, board identity (and hence, the homophily) changed after the exogenous shock compared to the group where board identity didn't change. The results for this test are available in Internet Appendix.

Here, Y is log compensation (*Log Compensation*), firm value (*Firm Value*), or firm value conditional on managerial compensation (*Residual Firm Value*). *Shock* is a dummy variable that equals 1 if any director has died or retired in that firm year; otherwise, 0. X is a vector of firm fundamentals and L is a vector of governance characteristics. The coefficient of interest is β_3. The firm (γ_{firm}) fixed effects are used to control for firm-level unobservables. [29]

5. RESULTS AND DISCUSSION

5.1. Group Identity, Executive Compensation, and Firm Value

Table 2 Columns (1) - (3) show the baseline results of how group identity (or homophily) affects managerial compensation. The results show that when a manager and the board share the same native language or come from the same native place, the manager earns higher compensation. The impact of group identity is stronger for native language than native place as $\beta_{Language} > \beta_{Place}$. The $\beta_{Language}$ is 0.066, and β_{Place} is 0.071, but the difference is not statistically significant. When the board and the manager are of the same caste, homophily negatively affects total compensation. However, the latter result is insignificant after the endogeneity test. Nevertheless, this result is similar to Freshtman and Gneezy (2001), whose behavioral experiment's predictions found that Eastern Jews discriminate against their people. However, this result can also be due to the skewness in the caste variable toward upper caste. Managerial designation, firm, and year fixed effects are used to control for the impact of unobservables.

[Insert Table 2 Here]

Table 2 Columns (4)-(6) show the impact of group identity on executive compensation with controls. These results confirm that if the manager and the board speak the same native language or belong to the same native place, the manager earns higher compensation than peers. $\beta_{Language}$ and β_{Place} are 0.078 and 0.077, and both coefficients are significant at the 1% level. This implies that in a firm where the board is more homophilic to the manager due to native language or native place, managers earn 7- 8% higher compensation. For external validation, I replace total compensation with salary, total current compensation (salary, bonus, and sitting fees), and compensation relative

[29] As time-fixed effects can affect the coefficient of the *Shock* dummy variable, they are not included in this specification.

to industry peers, and confirm the main results. Detailed results for external validation are provided in Table IA7 Panel (A)-(C). The coefficient of the homophily caste shows that the variable is negative and significant, but interpretation is difficult as the main variable is highly skewed. However, in untabulated results, I examine upper caste and lower caste groups separately and find that coefficients for homophily language and place are positive and significant. Among the controls, *Size*, *ROA*, *Institutional Ownership*, *Promoters Ownership*, *Number of Meetings*, *Firm Age*, and *BIG5* are significant and positively associated with managerial compensation. As expected, the coefficients of *Leverage* and *Volatility* are significant and negatively associated with compensation. The direction and magnitude of the controls' coefficients are in line with the findings in the executive compensation literature. I confirm these findings using a different database (based on "Indianchildnames.com") to measure homophily constructs and also using different measures of homophily. Detailed results for the same are provided in Table IA7 Panel D and Table IA8 Panels (A)-(B).

Table 3 Columns (1)-(3) exhibit the impact of homophily on firm value. The results show that homophily in the native language is positively associated with firm value and the results are significant at the 5% level. Similar to the executive compensation results, homophily based on native language is stronger than that based on native place, i.e., $\beta_{Language} > \beta_{Place}$. $\beta_{Language}$ is 0.125 and is significant at the 1% level. β_{Place} is 0.058, but it is significant at 12%. This result implies that firms in which there is more homophily between a manager and the board have a 12.5% higher value. Collectively, the 1% increase in executive compensation due to group identity leads to 1.8% increase in firm value. Among the controls - Size, Institutional Ownership, Promoters Ownership, Number of Meetings, and BIG5, are significant and positively associated with firm value. The coefficient of Leverage and Volatility is significant at the 5% level and negatively associated with firm value. To avoid the time trend and firm-level unobservables, firm and year fixed effects are used. For external validation, I replace market-to=book value of equity with two other firm value measures - sales growth and market value of equity plus book value of debt, and confirm the results. Detailed results for external validation are provided in Table IA7 Panel (A)-(C).

[Insert Table 3 Here]

Table 3 Columns (4)-(6) show the results of group identity on residual firm value. The results

indicate that homophily due to language is positively associated with residual firm value. Similar to the previous results, the impact of homophily based on native language is stronger than that based on native place, i.e., $\beta_{Language} > \beta_{Place}$. $\beta_{Language}$ is 0.095, which is significant at the 1% level. β_{Place} is 0.056, which is significant at the 15% level. This result implies that firms with more homophily between a manager and board have a 5.6-9.5% higher residual firm value, i.e., higher firm value excluding the cost of in-group favoritism. As managerial demographics are unavailable for the full sample, I show the results (Table IA6) using sub-sample analysis. I find that a male manager earns higher compensation than a same-level female manager, which is significant at the 1% level. Manager's age and education, however, do not significantly affect the total compensation.

There is a possibility that the results appear due to the specific dependent variables or the way homophily variables are constructed. To mitigate such concerns, I replace log total compensation with log salary, compensation relative to the industry average, and log salary plus bonus plus sitting fees. The market-to-book value is replaced by sales growth, ROE, log of market value of equity plus book value of debt. I also estimated new homophily constructs based on 'indianchildnames.com' database. The external validity test confirms the main findings of the study. The results for the external validity tests are provided in Internet Appendix Table IA7.

5.2. Endogeneity Tests Results

The endogenous relationship between executive compensation and corporate governance is always an issue (Hermalin and Weisbach, 1998; Wintoki et al., 2012) for this research question. As executive compensation and firm value are simultaneously determined (Palia (2001)), I use two-stage least squares (2SLS) regression to show the relation between firm value conditional on compensation and the homophily constructs. The results of the second stage regression are provided in Table 4 Panel A. The fitted value of log compensation from the first stage is negatively associated with firm value, supporting the notion that managerial compensation is a cost to the firm. $\beta_{Language}$ is 0.124 and is significant at the 1% level. β_{Place} is 0.096 and is significant at the 10% level. The coefficient of the fitted value of the total compensation is 0.570 and is significant at the 1% level. These results confirm the main findings that homophily increases firm value even when the cost of managerial compensation is considered. Internet Appendix Table IA12 provides the test statistics and p-values

of the Durbin (score) and Wu-Hausman tests. These statistics confirm that 2SLS is more efficient than the ordinary least squares (OLS) method in addressing the self-selection problem. [30]

[Insert Table 4 Here]

It is possible that the board chooses a manager who matches their identity to avoid conflicts in corporate decision-making (Damaraju and Makhija, 2018). Following Fracassi and Tate (2012), to mitigate self-selection bias, the director's death or retirement is used as a shock to the board identity and hence to the homophily variables.

Table 4 Panel B reports that the $\beta_{Language}$ for log compensation and residual firm value are 0.045 (significant at 5%) and 0.065 (significant at 1%), respectively. and the same for β_{Place} are 0.062 (significant at 5%) and 0.042 (significant at 10.6%), respectively. These results confirm the main finding that group identity between a manager and the board based on native language or native place leads to higher managerial compensation. Although the increment in managerial compensation is a cost to the firm, group identity also incentivizes the manager to exert more effort toward the firm's growth, and thus increase its value.

5.3. Additional Hypotheses

Powerful CEO

A powerful CEO influences compensation level and increases firm performance variability (Adams et al. 2005). To mitigate the possibility that results are driven by the position of the CEO rather than homophily. I also control for the manager's influence within the organization. Following Adams et al. (2005), I measure a powerful manager (Powerful CEO) as a dummy variable equal to 1 if the manager is both the founder (promoters) and chairman of the board. Results are provided in Table 5 and show that, even after controlling for manager's influence in the organization, homophily based on language and place remains positively associated with executive compensation and residual firm value. These findings support the main results of the article. The $\beta_{Language}$ for log compensation and residual firm value are 0.078 (significant at 1%) and 0.095 (significant

[30]The difference between the Durbin (1954) and Wu–Hausman ('T-2' statistics of Wu(1974)) tests of endogeneity is that the former uses an estimate of the error term's variance based on the model assuming the variables being tested are exogenous, while the latter uses an estimate of the error variance based on the model assuming the variables being tested are endogenous. Under the null hypothesis that the variables being tested are exogenous, both estimates of the error variance are consistent.

at 1%) and the same for β_{Place} are 0.077 (significant at 1%) and 0.056 (significant at 9.95%). I also find a positive and statistically significant relation between a powerful manager and executive compensation, but no relation between a powerful manager and firm value. This result implies rent extraction due to the manager's influence on the organization's decision-making process.

[Insert Table 5 Here]

College Ties and Past Experience

Farcassi and Tate (2012) show that powerful CEOs are more likely to appoint directors with ties to the CEO and that CEO-director ties reduce firm value. There is a possibility that an increase in compensation, firm value, and residual firm value is not due to homophily based on native language or native place, but due to induced identities such as college ties or past employment between a manager and the directors. I control for college ties and past employment connections between a manager and the directors to address this concern. Using BoardEx's education database and network database, I created variables based on college ties (*College Ties*) and past experiences (*Experience Ties*). *College Ties* is a dummy variable that takes the value 1 if there is at least one director on the board from the same college as the manager; otherwise, 0. Similarly, *Experience Ties* is a dummy variable that equals 1 if the manager has experience working with at least one board director; otherwise, 0.

[Insert Table 6 Here]

The results of this test are shown in Table 6. Panel A shows the results when controlling for college ties. $\beta_{Language}$ is 0.131, and β_{Place} is 0.126 for compensation, and both are statistically significant. $\beta_{Language}$ is also positive and significant for residual firm value. It implies that homophily based on native language between a manager and the board increases executive compensation and residual firm value, even controlling for college ties. The college ties between a manager and the board negatively impact compensation and residual firm value. Panel B shows the results controlling for ties from previous employment. $\beta_{Language}$ and β_{Place} are positive and statistically significant for executive compensation and residual firm value. This confirms the main results that homophily based on native language and native place increases executive compensation and reduces agency frictions by increasing residual firm value. The results also show that ties due to past employment

reduce total compensation, firm value, and residual firm value. These results are in line with Farcassi and Tate (2012).

Promoter as a Manager

In India, promoters (or founders) have a stronger influence on firms, as they are the major shareholders. Therefore, it is crucial to check whether the results are different for firms where one of the promoters is a manager. As this information is not available for the full sample, a sub-sample test is undertaken. The results are provided in Table 7. Columns (1) and (2) show that there is no impact due to homophily on compensation. These results are intuitive as promoters have no incentive to earn extra compensation based on their identity as they hold the majority of shares in the firm. Columns (4)-(5) show that homophily based on native language and native place increases firm value by 18.6% and 15.6%, respectively. This result is high compared to the full sample results, implying that the promoter being a manager, governs the firm better and undertakes constructive investment decisions. Columns (7)-(8) show that homophily due to language and place increases residual firm value by 14.6% and 11.2%. This implies that even after paying the cost of in-group favoritism, firm value of these homophily variables increases. This result is prominent in when the manager is also one of the promoters. See Table IA-P1.

Business Group vs. Non-Business Group Firms

India's business groups are collections of publicly traded firms with a presence in different industries and a significant amount of common ownership and control, usually by a family. Tarun and Palepu (2000) compare the performance of business group firms with the non-business group firms in India. They show that business group firms outperform non-business group firms. They argue that the group-affiliated firms have disproportionately good access to international sources of capital due to the business group's reputation.

The current study proposes an alternative channel for the better performance of group affiliated firms', i.e., a manager's reciprocation toward a firm due to homophily. The results are provided in Table 8. I do not find that homophily affects managerial compensation differently in business group firms than in others. However, its impact on firm value and residual firm value is more prominent in business group firms. The coefficients of interests, $\beta_{BGxHomophilyLanguage}$ and $\beta_{BGxHomophilyPlace}$,

for residual firm value are 7.3% (significant at 10%) and 4.9% (significant at 25%). In business group firms, managers primarily hire individuals from the same family or those known to the promoter (founder). Therefore, the manager's reciprocation towards the firm increases its value. Still, the increment in compensation due to homophily is not different from non-business group firms. See Table IA-P2.[31]

Social Independence of Board

Hwang and Kim (2009) show that social ties between a manager and independent board members affect how directors monitor and discipline the CEO. Using US data, they show that conventionally independent boards that are socially connected with the manager grant higher compensation. However, unlike this paper, they solely focus on independent board members. I replicate their test for Indian corporates and find that social ties between the manager and independent board members do not significantly affect managerial compensation, but they negatively affect future firm value and residual firm value. The coefficients of $\beta_{Language}$ and β_{place} for firm value are 10.5% and 11.2%, both significant at the 5% level. This supports Hwang and Kim's (2009) conclusion that independent board members should be both conventional and socially independent to monitor and discipline the manager. The results are provided in Table IA-P3.

Group Identity in Complex Firms

Complex firms are difficult to govern, and the need for managerial skillset may dominate group identity. In that case, group identity would not lead to favoritism. I test this hypothesis using Craswell, Stokes, and Laughton (2002) measurement of firm complexity - the square root of the number of subsidiaries. The regression results for managerial compensation, firm value, and residual firm value on firm complexity and its interaction with group identity variables are provided in Table 7.

[Insert Table 7 Here]

[31]The empirical specification for these regressions does not include firm fixed effects as BG dummy variable is used as a firm fixed effect.

The coefficient of *Complexity* is positive for log compensation, implying that managers are overpaid to handle complex firms. However, the coefficient of *Firm Complexity X Homophily Language* is -0.129 at the 1% level, indicating that group identity in complex firms reduces managerial compensation. However, firm value and residual firm value increase by 12.8% and 8.4%, suggesting that it is beneficial for these firms to hire a manager who shares a group identity with board members, as it increases future firm value and reduces managers' compensation compared to their counterparts.

Group Identity in Government Firms

A manager's recruitment process and working culture in government firms differ from those in other firms.[32] Therefore, it would be interesting to know whether group identity matters in such organizations. For simplicity, I focus on public sector undertakings (PSUs), which are government commercial enterprises with autonomous power to take decisions up to the non-management level. Results are provided in Table 11. *PSU* is a dummy variable that takes a value of 1 if a firm is a central or state government's commercial enterprise; otherwise, 0. I find that PSUs pay managers less and their firm value is also less than that of other firms. It is intuitive as PSUs provide many non-pecuniary benefits to managers, and it is difficult to fire them from their positions. A stable job and less incentive to take risk reduce future firm value. I find no impact of group identity on managerial compensation and firm value in PSUs. These results align with intuition as government involvement in managerial recruitment is unbiased from any implicit stereotypes. The results are provided in Table IA-P4.

Robustness Checks

Additional tests are conducted to explore the robustness of the results. First, I examine whether cultural differences between the southern and northern regions of India influence homophily between a manager and directors. I find no evidence that results differ when the manager and board members are from the southern region compared to when they are from the northern region. Second, I examine whether migration of communities (group associated with a particular surname) from one region of India to another affects the results. Only one community, 'Kashmiri Pandits', migrated

[32]In India, the manager and the directors are hired by the government and the recruitment process is considered to be unbiased.

from the Kashmir region to other parts of India in 1991 due to riots. I exclude the surnames belonging to this community and run the same regressions. Results are robust to this test (refer to Table IA14). Third, as the principal and agent have multiple identities and bias may arise from any of them. Therefore, rather than using a particular homophily construct, I use a homophily dummy variable which takes the value of 1 if all three identities match between a manager and at least one board member (refer to Table IA8 Panel A). Results suggest that managerial compensation increases by 2.9% due to the homophily index, which implies the cost of in-group favoritism. However, net firm value also increases by 2.7%. Fourth, I examine whether homophily yields a higher return to equity holders or not by using return on equity ROE as the dependent variable. $\beta_{Language}$ and β_{Place} are positive and significant for ROE and Residual ROE consistent with the main findings (refer to Table IA9). Lastly, following Bhagat and Bolton (2008), I use current and future return on assets (ROA) as dependent variables to check the impact of homophily on profitability. The results (Internet Appendix Table IA11) suggest that homophily language and homophily place are positively associated with current and future ROA. These findings are in contrast with Cai et al. (2013), who show that profitability reduces if the manager and owner are from the same family. I presume that these differences arise from differences in the governance system. High promoter ownership maintains a higher level of scrutiny over managers, which leads to higher profitability when there is homophily. The results table for the robustness tests is provided in the Internet Appendix (IA).

6. HOMOPHILY CONSTRUCTS AND FIRM VALUE: A CHANNEL

Managers and directors have multiple identities, such as age, sex, religion, race (or caste), affiliation with a native place, native language, and many others. Then, why does homophily based on some identities affect firm decision and firm performance (or value) and not all of them? The reason lies in 'trust' and supervision (governance). As the underlying element of homophily is trust, homophily based on different identities reflects different levels of trust. If a firm hires a trustworthy manager under adequate supervision (or governance), the manager exerts additional effort to increase firm value. The firm will recruit this manager even if the firm has to pay extra compensation to do so.

An alternative view of trust is the reduction in information asymmetry. Individuals who are

similar in their personal or social characteristics and share a common identity (e.g., race, gender, culture) tend to have common interests and worldviews, which explains the formation of ties based on interpersonal attraction (McPherson, Smith-Lovin, and Cook, 2001). For this reason, board members understand the manager well, leading to better monitoring and a higher likelihood that the manager refrains from shirking because of possible board sanctions. In other words, trust and reduced information asymmetry imply lower agency and monitoring costs. Using a survey of 600 Chinese private family firms, Cai et al. (2013) show that family managers earn 18-22% higher salaries and receive more bonuses, but face weaker incentives. Considering demographic similarity between the CEO and the directors of US firms, Westphal and Zajac (1995) provide evidence that similarity results in more generous CEO compensation contracts. Fracassi and Tate (2012) show that CEO-director ties based on shared experiences (such as past employment) reduce firm value, particularly in the absence of other governance mechanisms that can substitute for board oversight. The main results show that the manager gets a 6-8% higher compensation based on homophily, such as native language and native place, and even after paying a higher level of compensation, these firms exhibit an 8-9% higher value. Managerial efforts can drive the latter result due to promoters' intense supervision, as they are the major equity holders in Indian firms (Khanna and Palepu (2000)).

In sum, I argue that as homophily based on different identities reflects different levels of trust relationships, therefore the level of association of homophily with executive compensation and firm value is different for different identities. Most forms of homophily lead to high executive compensation due to favoritism. However, various shared identities are differently associated with firm outcomes due to weaker or stronger governance.

7. CONCLUSION

This paper examines how group identity arising from implicit bias affects executive compensation, firm value, and agency frictions. Integrating the identity-based utility function of Akerlof and Kranton (2005) into Holmstrom's (1979) first-order principal-agent model reveals that in-group bias elevates managerial compensation. The same mechanism induces greater managerial effort, thereby enhancing future firm value.

To empirically prove these theoretical predictions, I construct a novel database of 3.78 million Indian surnames from 474 million individuals in the Socio-Economic Caste Census (SECC) and Linguistic Survey of India (LSI). Surnames encode multiple identities—native language, place, and caste—enabling identification of shared identities between managers and directors. Managers sharing native language or place with board members receive 6–8% higher compensation, reflecting the cost of in-group favoritism. Such firms exhibit 11–12% higher value; after adjusting for favoritism costs, residual firm value rises by 5–8%, indicating heightened managerial effort. Results withstand external validity tests, endogeneity controls, geographic homophily variation, induced identities (college ties, past employment), managerial influence, and alternative homophily measures.

Across firm types, in-group favoritism yields marginally lower compensation premiums in business group firms, yet higher residual value, signaling stronger reciprocation. In complex firms, shared identity reduces compensation while elevating firm and residual value, underscoring merit's primacy. Public sector undertakings (PSUs) show no identity effects on compensation or value, consistent with impartial government recruitment.

In promoter-managed firms, group identity leaves promoter compensation unchanged but boosts firm and residual value, as promoters prioritize performance. Extending Hwang and Kim (2009), managerial ties with independent directors do not affect compensation but impair future firm value, validating their social independence hypothesis using natural identity measures.

In sum, in-group favoritism increases managerial compensation yet elevates firm value via reciprocated effort, suggesting group identity mitigates agency frictions. These findings imply that learning agent's specialization through homophily can enhance efficiency and welfare in principal-agent settings.

References

Adams, Renée B., Heitor Almeida, and Daniel Ferreira. "Powerful CEOs and their impact on corporate performance." The Review of Financial Studies 18.4 (2005): 1403-1432.

Akerlof, George A., and Rachel E. Kranton. "Economics and identity." The Quarterly Journal of Economics 115.3 (2000): 715-753.

Akerlof, George A., and Rachel E. Kranton. "Identity and the Economics of Organizations." Journal of Economic Perspectives 19.1 (2005): 9-32.

Akerlof, George A., and Rachel E. Kranton. 2008. "Identity, Supervision, and Work Groups." American Economic Review, 98 (2): 212-17.

Akerlof, George A., and Rachel Kranton. "Identity economics" The Economists' Voice 7.2 (2010).

Allport. "The nature of prejudice." Addison-Wesley Press, (1954).

Arrow, Kenneth. "The Theory of Discrimination." Discrimination in Labor Markets (1973): 3-33.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. "Social connections and incentives in the workplace: Evidence from personnel data." Econometrica 77. 4 (2009): 1047-1094.

Becker, Gary S. "The economics of discrimination."The University of Chicago Press, (1957).

Bertrand, Marianne, and Sendhil Mullainathan. "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination." American Economic Review 94.4 (2004): 991-1013.

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. "Implicit discrimination." American Economic Review 95.2 (2005): 94-98.

Bertrand, Marianne, and Kevin F. Hallock. "The gender gap in top corporate jobs." ILR Review 55.1 (2001): 3-21.

Bhagat, Sanjai, and Brian Bolton. "Corporate governance and firm performance." Journal of Corporate Finance 14. 3 (2008): 257-273.

Cai, Hongbin et al. "Family ties and organizational design: Evidence from Chinese private firms." Review of Economics and Statistics 95.3 (2013): 850-867.

Cikara, Mina, and Jay J. Van Bavel. "The neuroscience of intergroup relations: An integrative review." Perspectives on Psychological Science 9.3 (2014): 245-274.

Chen, Yan, and Sherry Xin Li. "Group identity and social preferences." American Economic Review 99.1 (2009): 431-57.

Chen, Roy, and Yan Chen. "The potential of social identity for equilibrium selection." American Economic Review 101.6 (2011): 2562-89.

Craswell, Allen, Donald J. Stokes, and Janet Laughton. "Auditor independence and fee dependence." Journal of Accounting and Economics 33. 2 (2002): 253-275.

Downes, David H., and Robert Heinkel. "Signaling and the valuation of unseasoned new issues." The Journal of Finance 37. 1 (1982): 1-10.

Edmans, Alex, Xavier Gabaix, and Dirk Jenter. "Executive compensation: A survey of theory and evidence." The Handbook of the Economics of Corporate Governance. Vol. 1. North-Holland, 2017. 383-539.

Fama, Eugene F. "Agency problems and the theory of the firm." Journal of political economy 88. 2 (1980): 288-307.

Fershtman, Chaim, and Uri Gneezy. "Discrimination in a segmented society: An experimental approach." The Quarterly Journal of Economics 116. 1 (2001): 351-377.

Fisman, Raymond, et al. "Social ties and favoritism in Chinese science." Journal of Political Economy 126.3 (2018): 1134-1171.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig. "Cultural proximity and loan outcomes." American Economic Review 107.2 (2017): 457-92.

Fracassi, Cesare, and Geoffrey Tate. "External networking and internal firm governance." The Journal of Finance 67. 1 (2012): 153-194.

Giannetti, Mariassunta, and Mengxin Zhao. "Board ancestral diversity and firm-performance volatility." Journal of Financial and Quantitative Analysis 54. 3 (2019): 1117-1155.

Goldstein, Itay, Chester S. Spatt, and Mao Ye. "Big data in finance." The Review of Financial Studies 34. 7 (2021): 3213-3225.

Gompers, Paul A., Vladimir Mukharlyamov, and Yuhai Xuan. "The cost of friendship." Journal of Financial Economics 119.3 (2016): 626-644.

Grinblatt, Mark, and Matti Keloharju. "How distance, language, and culture influence stockholdings and trades." The Journal of Finance 56.3 (2001): 1053-1073.

Hill, Jennifer G., and Randall S. Thomas, eds. Research Handbook on Executive Pay. Edward Elgar Publishing, 2012.

Holmstrom, Bengt. "Moral hazard and observability." The Bell Journal of Economics. (1979): 74. Hovakimian, Armen, Gayane Hovakimian, and Hassan Tehranian. "Determinants of target capital structure: The case of dual debt and equity issues." Journal of Financial Economics 71. 3 (2004): 517-540.

Hwang, Byoung-Hyoun, and Seoyoung Kim. "It pays to have friends." Journal of Financial Economics 93. 1 (2009): 138-158.

Jensen, Michael C., and William H. Meckling. "Theory of the firm: Managerial behavior, agency costs and ownership structure." Journal of Financial Economics 3. 4 (1976): 305-360.

Khanna, Tarun, and Krishna Palepu. "Is group affiliation profitable in emerging markets? An analysis of diversified Indian business groups." The Journal of Finance 55.2 (2000): 867-891.

Kumar, Alok, Alexandra Niessen-Ruenzi, and Oliver G. Spalt. "What's in a name? Mutual fund flows when managers have foreign-sounding names." The Review of Financial Studies 28.8 (2015): 2281-2321.

McDonald, Michael L., Poonam Khanna, and James D. Westphal. "Getting them to think outside the circle: Corporate governance, CEOs' external advice networks, and firm performance." Academy of Management Journal 51. 3 (2008): 453-475.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. "Birds of a feather: Homophily in social networks." Annual Review of Sociology 27. 1 (2001): 415-444.

Moscona, Jacob, Nathan Nunn, and James A. Robinson. "Keeping it in the family: Lineage organization and the scope of trust in sub-Saharan Africa." American Economic Review 107.5 (2017): 565-71.

Norman, Peter. "Statistical discrimination and efficiency." The Review of Economic Studies 70. 3 (2003): 615-627.

Palia, Darius. "The endogeneity of managerial compensation in firm valuation: A solution." The Review of Financial Studies 14. 3 (2001): 735-764.

Ritter, Jay R. Insider holdings and the pricing of initial public offerings. Rodney L. White Center for Financial Research, Wharton School, University of Pennsylvania, 1981.

Ross, Stephen A. "The economic theory of agency: The principal's problem." The American economic review 63. 2 (1973): 134-139.

Schwab, Stewart. "Is statistical discrimination efficient?." The American Economic Review 76. 1 (1986): 228-234.

Westphal, James D., and Edward J. Zajac. "Who shall govern? CEO/board power, demographic similarity, and new director selection." Administrative Science Quarterly (1995): 60-83.

Wiseman, Robert M., and Luis R. Gomez-Mejia. "A behavioral agency model of managerial risk taking." Academy of Management Review 23. 1 (1998): 133-153.

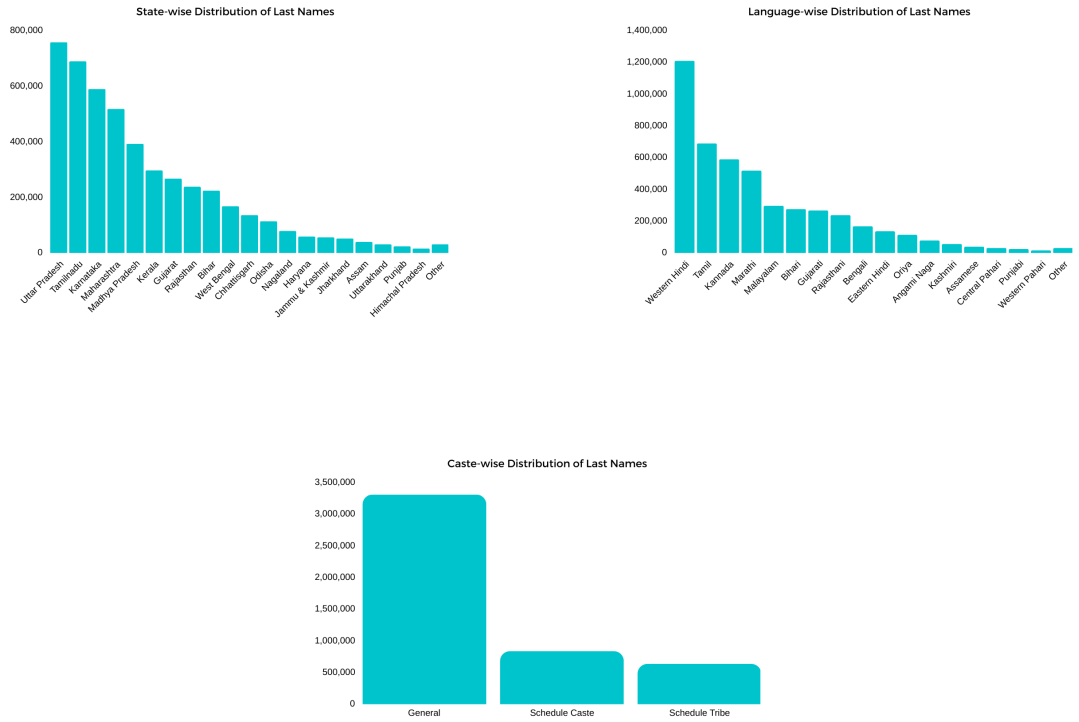
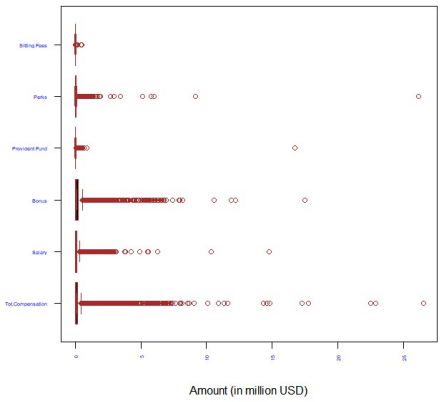


Figure 1: These figures show the state-wise, language-wise, and caste-wise distribution of 3.78 million Indian last names. Caste classification is based on administrative data, which classifies caste as scheduled caste (SC), scheduled tribe (ST), and other. The native language is from the 22 scheduled languages registered under the Indian constitution.



Managerial Compensation and Its Components

Figure 3: This figure shows the distribution of total managerial compensation and its components - salary, bonus, sitting fees, and contribution to provident fund. All values are in million USD.

Table 1

Panel A: Sample Selection

This table describes the sample selection procedure.

	Firms	Observations (Manager and Board Members)
Merging with surname and social identities data	24,326	1,473,836
Total remuneration is given	13,226	285,159
BSE and NSE listed non- financial firms	3,668	208,275
BSE and NSE listed non- financial firms (Manager)	3,668	41,507
Balanced Sample	2,324	28,026

Panel B: Managerial Compensation - Distribution and Summary Statistics

This table shows the components of managerial compensation (not in log form) and their summary statistics. Total Compensation (*Total Compensation*) is total managerial compensation. Salary (*Salary*) is salary of a manager. The other components of total compensation can be bonus (*Bonus*), contribution to provident fund (*Contribution to PF*), sittings fees (*Sitting Fees*), and perks (*Perks*). Components of total managerial compensation can vary depending on the position of a manager such as if manager is also a promoter or manager is also a chairman of the board. Total compensation and its components are in million USD. All variables are winsorized at 1%.

Variable	Obs	Mean	Std. Dev.	Min	Max
Total Compensation	28,026	0.192	0.373	0.000	2.461
Salary	21,328	0.132	0.183	0.002	1.116
Bonus	7,831	0.262	0.545	0.000	3.663
Perks	9,471	0.042	0.079	0.000	0.530
Sitting Fees	6,705	0.003	0.004	0.000	0.023
Contribution to PF	5,047	0.017	0.023	0.000	0.138

Panel C: Homophily Distribution (For Managers)

This table provides the distribution of observations based on homophily variables. The homophily variables are based on native language (*Homophily Language*), native place (*Homophily Place*), and caste (*Homophily Caste*). The variable definitions are provided in the Appendix.

	Language		Place		Caste	
Homophily	Freq.	Percent	Freq.	Percent	Freq.	Percent
0	16,240	57.95	16,897	60.29	3,018	10.77
1	11,786	42.05	11,129	39.71	25,008	89.23
Total	28,026	100	28,026	100	28,026	100

Panel D: Summary Statistics

This table provides summary statistics of variables used in the analyses. *Log Compensation* is log of total compensation. *Industry Relative Compensation* is compensation relative to industry average. *Firm value* is the ratio of market value to the book value of equity. *Residual Firm Value* is residual of regression of firm value on manager's total compensation of last year. *Homophily Language* is a dummy variable whose value is equal to 1 if the board (at least one director) and the manager speak the same Indian language, otherwise 0. *Homophily Place* is dummy variable whose value is equal to 1 if the board (at least one director) and the manager belong to same Indian state, otherwise 0. *Homophily Caste* is dummy variable whose value is equal to 1 if the board (at least one director) and the manager belong to same caste, otherwise 0. All other variables are defined in a variable definition sheet in the Appendix. All variables are winsorized at 1%.

Variable	Obs	Mean	Std. Dev.	Min	Max
Identity Variables					
Homophily Caste	28,026	0.892	0.310	0	1
Homophily Place	28,026	0.397	0.489	0	1
Homophily Language	28,026	0.421	0.494	0	1
HIL	28,026	0.224	0.379	0	1
HIP	28,026	0.182	0.351	0	1
HIC	28,026	0.874	0.302	0	1
Experience Ties	2,921	0.755	0.430	0	1
College Ties	3,299	0.267	0.443	0	1
Firm Characteristics					
Firm Value (MB Ratio)	28,026	2.238	3.073	-3.07	19.26
Residual Firm Value	27,908	-0.014	1.754	-5.143	8.470
Market Value (Debt + Market Cap)	28,026	627.36	2018.18	0.70	14899.99
Size	28,026	17.326	1.521	14.103	21.463
Firm Age	28,026	33.872	20.413	4	102
Institutional Ownership	28,026	9.540	12.006	0	48.365
Promoters Ownership	28,026	54.471	16.282	5.960	88.580
ROA	28,026	0.038	0.099	-1.435	0.416
Leverage	28,026	0.159	0.179	0	2.372
BG	28,026	0.489	0.500	0	1
PSU	28,026	0.024	0.154	0	1
Complexity	28,026	0.419	0.493	0	1
BIG 5	28,026	0.273	0.445	0	1
Board Characteristics					
Number of Board Meetings	28,026	7.855	3.155	1	29
Board Size	28,026	10.476	3.252	3	23
Shock	28,026	0.307	0.461	0	1
CEO-Chair Dummy	28,026	0.256	0.437	0	1
Manager Characteristics					
Log Compensation	28,026	10.611	2.158	3.412	13.634
Industry Relative Compensation	28,026	0.003	2.070	-5.699	3.752
Age	3,542	64.566	11.282	33	98
Gender	3,598	.982	0.132	0	1
Qualification	3,430	2	1.103	1	9
Powerful CEO	28,026	0.363	0.481	0	1
Promoter_Manager	10,802	0.933	0.249	0	1

Table 2: Group Identity and Executive Compensation

This table provides regressions results of executive compensation on three measures of group identity - homophily language, homophily place, and homophily caste. *Log Compensation* is log of total compensation. The total compensation is in million USD. *Homophily Language* is a dummy variable whose value equal to 1 if the board (at least one director) and the manager speaks same Indian language otherwise 0. *Homophily Place* is dummy variable whose value equal to 1 if the board (at least one director) and the manager belong to same Indian state otherwise 0. *Homophily Caste* is dummy variable whose value equal to 1 if the board (at least one director) and the manager belong to same caste otherwise 0. All other variables are defined in a variable definition sheet in the Appendix. The standard errors are in parentheses. All variables are winsorized at 1%. In all panels, *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

VARIABLES	(1) Log Compensation	(2) Log Compensation	(3) Log Compensation	(4) Log Compensation	(5) Log Compensation	(6) Log Compensation
Homophily Language	0.066***			0.078***		
	(0.021)			(0.021)		
Homophily Place		0.071***			0.077***	
		(0.022)			(0.022)	
Homophily Caste			-0.141***			-0.138***
			(0.041)			(0.041)
Size				0.094***	0.095***	0.093***
				(0.017)	(0.017)	(0.017)
Firm Age				0.032*	0.032*	0.032*
				(0.017)	(0.017)	(0.017)
ROA				1.133***	1.130***	1.126***
				(0.099)	(0.099)	(0.099)
Institutional Ownership				0.015***	0.015***	0.015***
				(0.001)	(0.001)	(0.001)
Promoters Ownership				0.007***	0.007***	0.007***
				(0.001)	(0.001)	(0.001)
Leverage				-0.148**	-0.146**	-0.149**
				(0.068)	(0.068)	(0.068)
Volatility				-5.924***	-5.915***	-5.883***
				(0.961)	(0.961)	(0.961)
Big 5				0.096**	0.094**	0.096**
				(0.038)	(0.038)	(0.038)
Number of Meetings				0.015***	0.015***	0.015***
				(0.004)	(0.004)	(0.004)
Board Size				-0.005	-0.005	-0.005
				(0.004)	(0.004)	(0.004)
Constant	10.583***	10.583***	10.737***	7.431***	7.414***	7.593***
	(0.011)	(0.011)	(0.037)	(0.668)	(0.668)	(0.668)
Observations	28,026	28,026	28,026	28,026	28,026	28,026
R-squared	0.758	0.758	0.758	0.762	0.762	0.762
Managerial Designation	Yes	Yes	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes

Table 3 : Group Identity, Firm Value, and Residual Firm Value

This table provides regression results of firm value and residual firm value on three measures of group identity - homophily language, homophily place, and homophily caste. *Firm value* is ratio of market value of equity to the book value of equity. Market value and book value of equity are in million USD. *Residual Firm Value* is residual of regression of firm value on manager's total compensation of last year. Total compensation is in million USD. *Homophily Language* is a dummy variable whose value is equal to 1 if the board (at least one director) and the manager speak the same Indian language, otherwise 0. *Homophily Place* is dummy variable whose value is equal to 1 if the board (at least one director) and the manager belong to same Indian state, otherwise 0. *Homophily Caste* is dummy variable whose value is equal to 1 if the board (at least one director) and manager belong to same caste, otherwise 0. All other variables are defined in a variable definition sheet in the Appendix. The dependent variable in Columns (1)-(3) is firm value and in Columns (4)-(6) is residual firm value in a year. The standard errors are in parentheses. All variables are winsorized at 1%. In all panels, *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

VARIABLES	(1) Firm Value	(2) Firm Value	(3) Firm Value	(4) Residual Firm Value	(5) Residual Firm Value	(6) Residual Firm Value
Homophily Language	0.124***			0.095***		
	(0.040)			(0.034)		
Homophily Place		0.058			0.056	
		(0.041)			(0.034)	
Homophily Caste			-0.122			-0.095
			(0.076)			(0.065)
Size	0.240***	0.240***	0.239***	0.225***	0.226***	0.224***
	(0.033)	(0.033)	(0.033)	(0.028)	(0.028)	(0.028)
Firm Age	0.053*	0.053*	0.053*	0.044	0.044	0.044
	(0.032)	(0.032)	(0.032)	(0.027)	(0.027)	(0.027)
ROA	4.243***	4.237***	4.234***	3.953***	3.948***	3.946***
	(0.185)	(0.185)	(0.185)	(0.157)	(0.157)	(0.157)
Institutional Ownership	0.048***	0.047***	0.047***	0.045***	0.045***	0.045***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Promoters Ownership	0.042***	0.042***	0.042***	0.037***	0.037***	0.037***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Leverage	0.136	0.138	0.135	0.153	0.154	0.152
	(0.127)	(0.128)	(0.128)	(0.108)	(0.108)	(0.108)
Volatility	-9.150***	-9.116***	-9.089***	-9.159***	-9.136***	-9.112***
	(1.803)	(1.803)	(1.803)	(1.526)	(1.527)	(1.526)
Big 5	0.317***	0.315***	0.316***	0.263***	0.261***	0.262***
	(0.071)	(0.071)	(0.071)	(0.060)	(0.060)	(0.060)
Number of Meetings	0.044***	0.044***	0.044***	0.043***	0.043***	0.043***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Board Size	0.013*	0.013*	0.013*	0.011*	0.011*	0.011*
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Constant	-7.021***	-6.985***	-6.835***	-8.351***	-8.335***	-8.208***
	(1.253)	(1.253)	(1.254)	(1.060)	(1.061)	(1.061)
Observations	28,026	28,026	28,026	28,026	28,026	28,026
R-squared	0.586	0.585	0.585	0.084	0.084	0.084
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes

Table 4: Resolving Simultaneity and Self-Selection Biases

This table provides results on resolving the simultaneity between firm value and executive compensation, and self-selection issues in executive hiring. *Log Compensation* is log of total compensation. *Firm value* is ratio of market value of equity to the book value of equity. *Residual Firm Value* is residual of regression of firm value on manager's total compensation of last year. Total compensation, market value of equity, and book value of equity are in million USD. *Homophily Language* is a dummy variable whose value is equal to 1 if the board (at least one director) and the manager speak the same Indian language, otherwise 0. *Homophily Place* is dummy variable whose value is equal to 1 if the board (at least one director) and manager belong to same Indian state, otherwise 0. *Homophily Caste* is dummy variable whose value is equal to 1 if the board (at least one director) and manager belong to same caste, otherwise 0. Panel A reports the regression results of firm value on group identity measures considering the simultaneity between firm value and executive compensation. Panel B reports the regression results of group identity and self-selection bias in hiring of a manager. All other variables are defined in variable definition sheet in the Appendix. The standard errors are in parentheses. All variables are winsorized at 1%. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Panel A: Group Identity and Simultaneity of Firm Value and Managerial Compensation

Panel A provides results of two-stage least squares (2SLS) regression used to resolve the simultaneity of firm value and managerial compensation . In the first stage, the estimated regression is logarithm of total managerial compensation on promoter dummy as a instrument and all firm and governance level controls. In the second stage, estimated regression is firm value on three group identity measures controlling for fitted value from first stage regression with other controls. The three measures of group identity are homophily language, homophily place, and homophily caste.

VARIABLES	(1) Firm Value	(2) Firm Value	(3) Firm Value
Log Compensation	-0.570***	-0.562***	-0.540***
	(0.125)	(0.125)	(0.118)
Homophily Language	0.124***		
	(0.047)		
Homophily Place		0.096*	
		(0.049)	
Homophily Caste			-0.050
			(0.057)
Size	0.276***	0.274***	0.272***
	(0.019)	(0.018)	(0.018)
Firm Age	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)
ROA	7.226***	7.206***	7.158***
	(0.361)	(0.361)	(0.346)
Institutional Ownership	0.045***	0.045***	0.044***
	(0.004)	(0.004)	(0.004)
Promoters Ownership	-34.334***	-34.219***	-33.786***
	(2.653)	(2.658)	(2.559)
Leverage	0.025***	0.024***	0.024***
	(0.002)	(0.002)	(0.001)
Volatility	-0.832***	-0.828***	-0.834***
	(0.111)	(0.111)	(0.110)
Big 5	0.772***	0.769***	0.755***
	(0.049)	(0.049)	(0.047)
Number of Meetings	0.015	0.015	0.013
	(0.011)	(0.011)	(0.010)
Board Size	-0.029***	-0.029***	-0.030***
	(0.007)	(0.007)	(0.007)
Constant	2.435**	2.397**	2.334**
	(0.947)	(0.947)	(0.913)
Observations	28,026	28,026	28,026
R-squared	0.057	0.061	0.071
Industry	Yes	Yes	Yes
Year	Yes	Yes	Yes

Panel B: Group Identity and Self-Selection in Hiring of a Manager

This panel provides regression results of managerial compensation, firm value, and residual firm value on three measures of group identity considering the self-selection bias in hiring of manager. The selection bias is mitigated using the death or the non-voluntary retirement of a director as a exogenous shock to group identity measures. In Columns (1) - (3), the dependent variable is logarithm of total compensation (*Log Compensation*) of a manager in that firm-year. In Columns (4)-(6), the dependent variable is firm value (*Firm Value*). In Columns (7)-(9), the dependent variable is residual firm value (*Residual Firm Value*). All other variables are defined in variable definition sheet in Appendix. The standard errors are in parentheses.All variables are winsorized at 1%. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

VARIABLES	(1) Log Compensation	(2) Log Compensation	(3) Log Compensation	(4) FirmValue	(5) FirmValue	(6) FirmValue	(7) Residual Firm Value	(8) Residual Firm Value	(9) Residual FirmValue
Shock	-0.094***	-0.099***	-0.064	-0.116***	-0.108**	-0.109	-0.106***	-0.093**	-0.074
	(0.019)	(0.020)	(0.064)	(0.034)	(0.040)	(0.069)	(0.034)	(0.039)	(0.059)
Homophily Language	0.063			0.105**			0.069*		
	(0.050)			(0.049)			(0.038)		
Shock x Homophily Language	0.045**			0.059***			0.065***		
	(0.022)			(0.019)			(0.016)		
Homophily Place		0.057			0.043			0.041	
		(0.039)			(0.055)			(0.047)	
Shock x Homophily Place		0.062**			0.043*			0.042	
		(0.023)			(0.023)			(0.026)	
Homophily Caste			-0.134***						-0.095
			(0.034)						(0.092)
Shock x Homophily Caste			-0.014						-0.005
			(0.066)						(0.050)
Size	0.093***	0.094***	0.093***	0.238***	0.239***	0.238***	0.225***	0.225***	0.223***
	(0.023)	(0.022)	(0.022)	(0.022)	(0.021)	(0.022)	(0.020)	(0.019)	(0.019)
Firm Age	0.032	0.032	0.032	0.054	0.054	0.054	0.042	0.045	0.045
	(0.062)	(0.061)	(0.060)	(0.054)	(0.054)	(0.055)	(0.049)	(0.049)	(0.050)
ROA	1.131***	1.127***	1.123***	4.241***	4.234***	4.230***	3.956***	3.946***	3.943***
	(0.089)	(0.086)	(0.084)	(0.315)	(0.315)	(0.314)	(0.284)	(0.282)	(0.281)
Institutional Ownership	0.015***	0.015***	0.015***	0.047***	0.047***	0.047***	0.045***	0.045***	0.045***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Promoters Ownership	0.007***	0.007***	0.007***	0.042***	0.042***	0.042***	0.037***	0.037***	0.037***
	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Leverage	-0.144	-0.141	-0.144*	0.141	0.143	0.141	0.155	0.158	0.156
	(0.085)	(0.086)	(0.085)	(0.245)	(0.244)	(0.246)	(0.180)	(0.176)	(0.178)
Volatility	-5.882***	-5.871***	-5.830***	-9.103***	-9.059***	-9.027***	-9.090***	-9.089***	-9.059***
	(1.119)	(1.103)	(1.094)	(2.500)	(2.504)	(2.515)	(1.721)	(1.717)	(1.722)
Big 5	0.098***	0.096***	0.097***	0.318***	0.316***	0.318***	0.263***	0.263***	0.264***
	(0.015)	(0.016)	(0.015)	(0.082)	(0.080)	(0.080)	(0.066)	(0.063)	(0.063)
Number of Meetings	0.016***	0.016***	0.016***	0.044***	0.044***	0.044***	0.043***	0.044***	0.044***
	(0.002)	(0.002)	(0.002)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
Board Size	-0.001	-0.001	-0.000	0.018	0.018	0.018	0.018*	0.015	0.015
	(0.004)	(0.004)	(0.004)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Constant	7.407***	7.398***	7.554***	-7.047***	-7.015***	-6.877***	-8.324***	-8.358***	-8.246***
	(2.000)	(1.988)	(1.896)	(2.003)	(2.001)	(2.118)	(1.710)	(1.756)	(1.848)
Observations	28,026	28,026	28,026	28,026	28,026	28,026	28,026	28,026	28,026
R-squared	0.762	0.762	0.762	0.586	0.586	0.586	0.087	0.084	0.084
Designation	Yes	Yes	Yes	No	No	No	No	No	No
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 5: Powerful CEO

This table provides regression results of managerial compensation, firm value, and residual firm value on group identity when a manager is powerful. *Log Compensation* is log of total compensation. *Firm value* is ratio of market value of equity to the book value of equity. *Residual Firm Value* is residual of regression of firm value on manager's total compensation of last year. Total compensation, market value of equity, and book value of equity are in million USD. *Homophily Language* is a dummy variable whose value is equal to 1 if the board (at least one director) and manager speak the same Indian language, otherwise 0. *Homophily Place* is dummy variable whose value is equal to 1 if the board (at least one director) and manager belong to the same Indian state, otherwise 0. *Homophily Caste* is dummy variable whose value is equal to 1 if the board (at least one director) and manager belong to same caste, otherwise 0. *Powerful CEO* is a dummy variable which takes value of 1 if a manager is chairperson, CEO, and promoter; otherwise 0. In Columns (1) - (3), the dependent variable is logarithm of total compensation (*Log Compensation*) of a manager in that firm-year. In Columns (4)-(6), the dependent variable is firm value (*Firm Value*). In Columns (7)-(9), the dependent variable is residual firm value (*Residual Firm Value*). All other variables are defined in a variable definition sheet in the Appendix. The standard errors are in parentheses. All variables are winsorized at 1%. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

VARIABLES	(1) Log Compensation	(2) Log Compensation	(3) Log Compensation	(4) FirmValue	(5) FirmValue	(6) FirmValue	(7) Residual Firm Value	(8) Residual Firm Value	(9) Residual FirmValue
Homophily Language	0.078***			0.124***			0.095***		
	(0.021)			(0.040)			(0.034)		
Homophily Place		0.077***			0.059			0.056	
		(0.022)			(0.041)			(0.034)	
Homophily Caste			-0.137***			-0.122			-0.097
			(0.041)			(0.076)			(0.065)
Powerful CEO	0.032	0.031	0.031	-0.035	-0.036	-0.032	-0.027	-0.027	-0.028
	(0.021)	(0.021)	(0.021)	(0.039)	(0.039)	(0.040)	(0.033)	(0.033)	(0.033)
Size	0.093***	0.095***	0.092***	0.241***	0.241***	0.243***	0.226***	0.226***	0.225***
	(0.017)	(0.017)	(0.017)	(0.033)	(0.033)	(0.033)	(0.028)	(0.028)	(0.028)
Firm Age	0.032*	0.032*	0.032*	0.053*	0.053	0.051	0.044	0.044	0.044
	(0.017)	(0.017)	(0.017)	(0.032)	(0.032)	(0.032)	(0.027)	(0.027)	(0.027)
ROA	1.132***	1.129***	1.125***	4.244***	4.239***	4.233***	3.954***	3.949***	3.947***
	(0.099)	(0.099)	(0.099)	(0.185)	(0.185)	(0.185)	(0.157)	(0.157)	(0.157)
Institutional Ownership	0.015***	0.015***	0.015***	0.048***	0.047***	0.048***	0.045***	0.045***	0.045***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Promoters Ownership	0.007***	0.007***	0.007***	0.042***	0.042***	0.042***	0.037***	0.037***	0.037***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Leverage	-0.146**	-0.145**	-0.148**	0.135	0.136	0.136	0.152	0.153	0.151
	(0.068)	(0.068)	(0.068)	(0.128)	(0.128)	(0.128)	(0.108)	(0.108)	(0.108)
Volatility	-5.923***	-5.914***	-5.882***	-9.152***	-9.118***	-9.076***	-9.160***	-9.137***	-9.113***
	(0.960)	(0.960)	(0.961)	(1.803)	(1.803)	(1.803)	(1.526)	(1.527)	(1.527)
Big 5	0.097**	0.095**	0.097**	0.316***	0.314***	0.314***	0.262***	0.261***	0.262***
	(0.038)	(0.038)	(0.038)	(0.071)	(0.071)	(0.071)	(0.060)	(0.060)	(0.060)
Number of Meetings	0.015***	0.015***	0.015***	0.044***	0.044***	0.043***	0.043***	0.043***	0.043***
	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Board Size	-0.005	-0.005	-0.005	0.013*	0.013*	0.016**	0.011*	0.011*	0.011*
	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Constant	7.433***	7.416***	7.594***	-7.021***	-6.986***	-6.838***	-8.352***	-8.336***	-8.207***
	(0.667)	(0.668)	(0.668)	(1.253)	(1.253)	(1.254)	(1.060)	(1.061)	(1.061)
Observations	28,026	28,026	28,026	28,026	28,026	28,026	28,026	28,026	28,026
R-squared	0.762	0.762	0.762	0.586	0.585	0.586	0.084	0.084	0.084
Manager Designation	Yes	Yes	Yes	No	No	No	No	No	No
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6: Group Identity and Ties Based on College and Past Experience

This table reports the main results controlling the induced identity variables such as college-level ties and ties based on shared experiences. *Log Compensation* is log of total compensation. *Firm value* is ratio of market value of equity to the book value of equity. *Residual Firm Value* is residual of regression of firm value on manager's total compensation of last year. Total compensation, market value of equity, and book value of equity are in million USD. *Homophily Language* is a dummy variable whose value is equal to 1 if the board (at least one director) and manager speak the same Indian language. otherwise 0. *Homophily Place* is dummy variable whose value is equal to 1 if the board (at least one director) and manager belong to same Indian state, otherwise 0. *Homophily Caste* is dummy variable whose value is equal to 1 if the board (at least one director) and manager belong to same caste, otherwise 0. College ties (*College Ties*) is a dummy variable, which takes value of 1 if a manager and at least one board member went to the same college, otherwise 0. Past experience ties (*Past Experience Ties*) is a dummy variable, which takes value of 1 if a manager and at least one board member worked for the same firm in the past, otherwise 0. Panel A panel provides regression results of managerial compensation, firm value, and residual firm value on group identity controlling for college level ties between a manager and board members. Panel B provides regression results of managerial compensation, firm value, and residual firm value on group identity controlling for past employment level ties between a manager and board members. All other variables are defined in the Variable Definition Sheet in the Appendix. The standard errors are in parentheses. All variables are winsorized at 1%. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Panel A: College Ties

This panel provides regression results of managerial compensation, firm value, and residual firm value on group identity, controlling for college-level ties between a manager and board members. In Columns (1) - (3), the dependent variable is the logarithm of total compensation (*Log Compensation*) of a manager in that firm-year. In Columns (4)-(6), the dependent variable is firm value (*Firm Value*). In Columns (7)-(9), the dependent variable is residual firm value (*Residual Firm Value*).

VARIABLES	(1) Log Compensation	(2) Log Compensation	(3) Log Compensation	(4) Firm Value	(5) Firm Value	(6) Firm Value	(7) Residual Firm Value	(8) Residual Firm Value	(9) Residual Firm Value
Homophily Language	0.131**			0.305**			0.313**		
	(0.058)			(0.136)			(0.123)		
Homophily Place		0.126**			0.094			0.118	
		(0.060)			(0.148)			(0.134)	
Homophily Caste			-0.125			-0.162			-0.137
			(0.087)			(0.212)			(0.191)
College Ties	0.071	0.068	0.080	-0.055	-0.039	-0.037	-0.044	-0.028	-0.025
	(0.061)	(0.061)	(0.061)	(0.129)	(0.129)	(0.129)	(0.116)	(0.116)	(0.116)
Size	0.040*	0.040*	0.037	0.019	0.012	0.009	0.074	0.067	0.064
	(0.023)	(0.023)	(0.023)	(0.129)	(0.129)	(0.129)	(0.116)	(0.117)	(0.117)
Firm Age	0.005***	0.006***	0.005***	-0.052	-0.056	-0.066	-0.060	-0.063	-0.072
	(0.001)	(0.001)	(0.001)	(0.085)	(0.085)	(0.086)	(0.077)	(0.077)	(0.078)
ROA	1.940***	1.960***	1.972***	6.082***	6.148***	6.164***	5.801***	5.869***	5.882***
	(0.329)	(0.329)	(0.329)	(0.635)	(0.635)	(0.635)	(0.574)	(0.574)	(0.574)
Institutional Ownership	0.013***	0.013***	0.013***	0.053***	0.051***	0.051***	0.052***	0.050***	0.050***
	(0.003)	(0.003)	(0.003)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Promoters Ownership	-0.004*	-0.004*	-0.004*	0.127***	0.125***	0.125***	0.115***	0.113***	0.113***
	(0.002)	(0.002)	(0.002)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)
Leverage	0.798***	0.830***	0.846***	2.221***	2.269***	2.278***	2.004***	2.052***	2.061***
	(0.203)	(0.202)	(0.203)	(0.548)	(0.548)	(0.548)	(0.495)	(0.495)	(0.495)
Volatility	-27.132***	-27.014***	-26.525***	15.765	15.476	15.421	17.476*	17.167*	17.144*
	(5.352)	(5.351)	(5.345)	(11.419)	(11.428)	(11.428)	(10.321)	(10.331)	(10.332)
Big 5	0.482***	0.485***	0.484***	-0.234	-0.231	-0.228	-0.180	-0.178	-0.174
	(0.058)	(0.058)	(0.058)	(0.194)	(0.194)	(0.194)	(0.175)	(0.175)	(0.175)
Number of Meetings	-0.011	-0.011	-0.010	-0.001	-0.002	-0.001	-0.010	-0.010	-0.009
	(0.014)	(0.014)	(0.014)	(0.024)	(0.024)	(0.024)	(0.022)	(0.022)	(0.022)
Board Size	-0.010	-0.010	-0.009	-0.059***	-0.060***	-0.060***	-0.051***	-0.052***	-0.052***
	(0.008)	(0.008)	(0.008)	(0.020)	(0.020)	(0.020)	(0.018)	(0.018)	(0.018)
Constant	10.976***	10.981***	11.190***	-2.470	-1.975	-1.346	-6.659*	-6.186	-5.594
	(0.435)	(0.436)	(0.435)	(4.273)	(4.273)	(4.312)	(3.861)	(3.862)	(3.898)
Observations	3,299	3,299	3,299	3,299	3,299	3,299	3,299	3,299	3,299
R-squared	0.412	0.412	0.411	0.789	0.789	0.789	0.377	0.375	0.375
Manager Designation	Yes	Yes	Yes	No	No	No	No	No	No
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Past Employment Ties

This panel provides regression results of managerial compensation, firm value, and residual firm value on group identity controlling for past employment level ties between a manager and board members. In Columns (1) - (3), the dependent variable is logarithm of total compensation (*Log Compensation*) of a manager in that firm-year. In Columns (4)-(6), the dependent variable is firm value (*Firm Value*). In Columns (7)-(9), the dependent variable is residual firm value (*Residual Firm Value*).

VARIABLES	(1) Log Compensation	(2) Log Compensation	(3) Log Compensation	(4) Firm Value	(5) Firm Value	(6) Firm Value	(7) Residual Firm Value	(8) Residual Firm Value	(9) Residual Firm Value
Homophily Language	0.145**			0.403**			0.413***		
	(0.064)			(0.164)			(0.139)		
Homophily Place		0.183***			0.445**			0.323**	
		(0.067)			(0.182)			(0.155)	
Homophily Caste			-0.229**			0.191			0.215
			(0.096)			(0.281)			(0.238)
Past Employment Ties	-0.065	-0.066	-0.070	-0.144	-0.151	-0.156	-0.143	-0.151	-0.155
	(0.068)	(0.068)	(0.068)	(0.123)	(0.123)	(0.123)	(0.104)	(0.104)	(0.105)
Size	0.009	0.011	0.008	0.081	0.080	0.070	0.142	0.137	0.131
	(0.024)	(0.024)	(0.024)	(0.146)	(0.146)	(0.146)	(0.124)	(0.124)	(0.124)
Firm Age	0.006***	0.006***	0.005***	-0.015	-0.001	0.001	-0.027	-0.014	-0.009
	(0.001)	(0.001)	(0.001)	(0.138)	(0.138)	(0.139)	(0.117)	(0.117)	(0.117)
ROA	3.043***	3.029***	3.092***	8.293***	8.232***	8.337***	7.042***	7.017***	7.085***
	(0.387)	(0.387)	(0.386)	(0.857)	(0.858)	(0.858)	(0.726)	(0.727)	(0.727)
Institutional Ownership	0.018***	0.018***	0.017***	0.058***	0.058***	0.055***	0.054***	0.053***	0.052***
	(0.003)	(0.003)	(0.003)	(0.011)	(0.011)	(0.011)	(0.009)	(0.009)	(0.009)
Promoters Ownership	-0.003	-0.003	-0.004	0.135***	0.134***	0.134***	0.120***	0.119***	0.119***
	(0.002)	(0.002)	(0.002)	(0.013)	(0.013)	(0.013)	(0.011)	(0.011)	(0.011)
Leverage	0.800***	0.828***	0.864***	-3.298***	-3.243***	-3.202***	-2.292***	-2.219***	-2.196***
	(0.216)	(0.215)	(0.215)	(0.709)	(0.708)	(0.709)	(0.601)	(0.600)	(0.601)
Volatility	-20.705***	-20.855***	-20.719***	-2.473	-1.970	-1.776	5.176	5.712	5.901
	(5.557)	(5.556)	(5.557)	(13.066)	(13.065)	(13.080)	(11.069)	(11.077)	(11.086)
Big 5	0.586***	0.588***	0.585***	-0.356	-0.394	-0.369	-0.410*	-0.441**	-0.424**
	(0.063)	(0.063)	(0.063)	(0.253)	(0.253)	(0.253)	(0.214)	(0.214)	(0.214)
Number of Meetings	0.007	0.008	0.009	-0.015	-0.016	-0.016	-0.018	-0.018	-0.019
	(0.014)	(0.014)	(0.014)	(0.028)	(0.028)	(0.028)	(0.024)	(0.024)	(0.024)
Board Size	0.012	0.012	0.013	0.007	0.005	0.005	-0.001	-0.002	-0.003
	(0.009)	(0.009)	(0.009)	(0.025)	(0.025)	(0.025)	(0.021)	(0.021)	(0.021)
Constant	10.663***	10.604***	10.938***	-5.299	-5.770	-5.647	-9.094*	-9.365*	-9.509*
	(0.468)	(0.469)	(0.467)	(6.147)	(6.153)	(6.213)	(5.207)	(5.217)	(5.265)
Observations	2,921	2,921	2,921	2,921	2,921	2,921	2,921	2,921	2,921
R-squared	0.425	0.425	0.425	0.750	0.750	0.749	0.413	0.412	0.411
Manager Designation	Yes	Yes	Yes	No	No	No	No	No	No
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7: Group Identity in Complex Firms

This table reports the regression results of executive compensation, firm value, and residual firm value on group identity in complex firms. Firm Complexity (*Firm Complexity*) is a dummy variable which takes value of 1 if complexity of a firm is above median; otherwise 0. The complexity of a firm is measured as square root of number of subsidiaries. Log Compensation (*Log Compensation*) is logarithm of total managerial compensation which is in million USD. *Homophily Language* is a dummy variable whose value is equal to 1 if the board (at least one director) and manager speak the same Indian language, otherwise 0. *Homophily Place* is dummy variable whose value is equal to 1 if the board (at least one director) and manager belong to same Indian state, otherwise 0. *Homophily Caste* is dummy variable whose value is equal to 1 if the board (at least one director) and manager belong to same caste, otherwise 0. All other variables are defined in a variable definition sheet in the Appendix. The standard errors are in parentheses. All variables are winsorized at 1%. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

VARIABLES	(1) Log Compensation	(2) Log Compensation	(3) Log Compensation	(4) Firm Value	(5) Firm Value	(6) Firm Value	(7) Residual Firm Value	(8) Residual Firm Value	(9) Residual Firm Value
Firm Complexity	0.465***	0.432***	0.388***	-0.273***	-0.231***	-0.166	-0.127***	-0.124***	-0.003
	(0.025)	(0.025)	(0.055)	(0.048)	(0.047)	(0.104)	(0.030)	(0.030)	(0.065)
Homophily Language	0.111***			-0.066			0.011		
	(0.023)			(0.044)			(0.028)		
Firm Complexity x Homophily Language	-0.129***			0.128*			0.084*		
	(0.036)			(0.069)			(0.043)		
Homophily Place		0.088***			-0.060			-0.003	
		(0.023)			(0.044)			(0.028)	
Firm Complexity x Homophily Place		-0.053			0.022			0.079*	
		(0.036)			(0.069)			(0.043)	
Homophily Caste			-0.053			-0.011			0.026
			(0.037)			(0.070)			(0.044)
Firm Complexity x Homophily Caste			0.026			-0.062			-0.103
			(0.057)			(0.108)			(0.068)
Size	0.053***	0.054***	0.053***	0.242***	0.241***	0.242***	-0.000	-0.001	-0.001
	(0.007)	(0.007)	(0.007)	(0.014)	(0.014)	(0.014)	(0.009)	(0.009)	(0.009)
Firm Age	0.006***	0.006***	0.006***	-0.005***	-0.005***	-0.005***	-0.001	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Institutional Ownership	0.027***	0.027***	0.026***	0.031***	0.031***	0.031***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Promoter Ownership	0.008***	0.008***	0.008***	0.020***	0.020***	0.020***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ROA	2.319***	2.316***	2.320***	5.867***	5.869***	5.869***	2.306***	2.307***	2.309***
	(0.100)	(0.100)	(0.100)	(0.190)	(0.190)	(0.190)	(0.119)	(0.119)	(0.119)
Leverage	0.141***	0.144***	0.141***	-0.873***	-0.874***	-0.872***	0.134**	0.136**	0.137**
	(0.054)	(0.054)	(0.054)	(0.102)	(0.102)	(0.102)	(0.064)	(0.064)	(0.064)
Volatility	-16.744***	-16.712***	-16.586***	-27.011***	-27.024***	-27.118***	-1.927	-1.924	-1.934
	(1.023)	(1.024)	(1.023)	(1.945)	(1.945)	(1.944)	(1.220)	(1.220)	(1.219)
Big5	0.432***	0.432***	0.427***	0.707***	0.704***	0.708***	-0.032	-0.033	-0.036
	(0.022)	(0.022)	(0.022)	(0.042)	(0.042)	(0.042)	(0.026)	(0.026)	(0.026)
Number of Meetings	0.051***	0.050***	0.050***	-0.018***	-0.018***	-0.018***	0.015***	0.015***	0.015***
	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
Board Size	0.031***	0.031***	0.031***	-0.042***	-0.041***	-0.041***	-0.007*	-0.007*	-0.006*
	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
Constant	8.125***	8.136***	8.232***	-2.002***	-1.996***	-2.024***	-0.261*	-0.251	-0.270*
	(0.129)	(0.129)	(0.133)	(0.245)	(0.244)	(0.252)	(0.154)	(0.153)	(0.158)
Observations	28,026	28,026	28,026	28,026	28,026	28,026	28,026	28,026	28,026
R-squared	0.548	0.548	0.548	0.193	0.193	0.193	0.020	0.020	0.020
Manager Designation	Yes	Yes	Yes	No	No	No	No	No	No
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Appendix A

Variable Definition Sheet

Variable	Description	Source
TotalCompensation	Log of manager's total compensation. Total compensation is in USD.	ProwessDx
Salary	Log of manager's salary. Salary is in USD.	ProwessDx
Firm Value (Market to Book Ratio)	Ratio of market value of equity to book value of equity. The market value of equity and book value of equity are in USD.	ProwessDx
Market Value	Sum of book value of debt and market value of equity. The market value of equity and book value of debt are in USD.	ProwessDx
Sales Growth	Growth in sales compare to last year	ProwessDx
Residual Firm Value	Residual from regression of market to Book ratio on last year's managerial compensation	ProwessDx
Homophily Language	Dummy variable which takes value of 1 if a manager and majority of board members have the same native language, otherwise 0.	Linguistic Survey of India (LSI)
Homophily Place	Dummy variable which takes value of 1 if a manager and majority of board members have the same native place, otherwise 0.	Socio-Economic Caste Census (SECC)
Homophily Caste	Dummy variable which takes value 1 if a manager and majority of board members have the same caste; otherwise 0.	Socio-Economic Caste Census (SECC)
Size	Log of market capitalization. Market capitalization is in USD.	ProwessDx
ROA	Ratio of net income to total assets.	ProwessDx
ROE	Ratio of net income to book value of equity.	ProwessDx
Promoters Ownership	Perentage of equity ownership by promoters.	ProwessDx
Institutional Ownership	Perentage of equity ownership by institutional investers.	ProwessDx
Leverage	Ratio of total debt to total assets.	ProwessDx
South	Dummy variable which takes value of 1 if the state is in southern part of India, otherwise 0.	Socio-Economic Caste Census (SECC)
Firm Age	Age of the firm from incorporation	ProwessDx
Board Size	Number of board members	BoardEx
Number of Meetings	Number of board meetings in a year.	BoardEx
Board Diversity	Percentage of women in board	BoardEx
BG	Dummy variable which takes value of 1 if a firm belongs to a business group, otherwise 0.	ProwessDx
Firm Complexity	Dummy variable, which takes value of 1 if complexity of a firm is above median, otherwise 0. The complexity of a firm is measured as square root of number of subsidiaries.	ProwessDx
PSU	Dummy variable which takes value of 1 if a firm is central or state government's commercial enterprise (also known as public sector undertakings (PSUs), otherwise 0.	ProwessDx
Volatility	Volatility of stock return.	Global Compustat
HIL	Dummy variable whose value is equal to 1 if the independent directors (at least one) and the manager speaks same Indian language, otherwise 0.	Linguistic Survey of India (LSI)
HIP	Dummy variable whose value is equal to 1 if the independent directors (at least one) and manager belong to same Indian state, otherwise 0.	Socio-Economic Caste Census (SECC)
HIC	Dummy variable whose value is equal to 1 if the independent directors (at least one) and the manager have the same caste, otherwise 0.	Socio-Economic Caste Census (SECC)
Shock	Dummy variable which takes value 1 if a director died or non-voluntary retired in that year; otherwise 0.	ProwessDx
Age	Managerial age in years.	BoardEx
Gender	Dummy variable which takes value 1 if a manager is male and 0 if a manger is female.	BoardEx
Qualification(Education)	Number of educational degrees by a manager.	BoardEx