INSEAD The Business School for the World®

Working Paper

2022/24/EFE

Diversity in Corporate Boards and Firm Outcomes

Suresh Bhagavatula
IIM Bangalore, sureshbh@iimb.ernet.in

Manaswini Bhalla
IIM Bangalore, manaswinib@iimb.ernet.in
Corresponding author

Manisha Goel
Pomona College, manisha.goel@pomona.edu

Balagopal Vissa
INSEAD, balagopal.vissa@insead.edu

We examine the relationship between firm performance and mindset diversity among corporate directors, proxied by diversity along religion and caste, a deep-rooted institution that divides India's Hindu society into hundreds of communities. We build the first data-driven mapping of last names to religion and caste at different levels of granularity, and use it to identify directors' castes and religions. Indian corporate boards strikingly lack diversity during 1999-2015. We show that lack of board diversity is negatively related to firm performance. We find the strongest negative results for the finest caste measure, along which mindsets are likely most similar, reinforcing the importance of the granularity at which diversity is measured.

JEL Classifications: M14, L25, G30, Z13
Keywords: Diversity; Firm Performance; Corporate Governance; Caste; Corporate Culture

Electronic copy available at: http://ssrn.com/abstract=4021311

Acknowledgements
We thank Stephan Siegel, Jared Rubin, Latika Chaudhary, Sriya Iyer, Tahir Andrabi, Asim Khwaja, Rajiv Sethi, Simi Kedia, Anand Venkateshwaran and Michelle Zemel for helpful conversations and seminar participants at IIM Bangalore, Indian School of Business, and Indian Statistical Institute, and conference participants at AALIMS 2017 and Liberal Arts Macro and Finance Workshop 2018 for their comments. Thanks to VSK Teja Konduri, Tejendra Pratap Singh, Timothy Howe, Charles Fries, Leena Kinger Hans, and Sharada PM for valuable research assistance.

1 Introduction

In recent years, the lack of diversity in corporate boards has garnered considerable attention. Several measures have been proposed to diversify corporate boards, with a key principle being that diversity of identities and backgrounds brings diversity of perspectives and mindsets.[1] However, research on the relationship between mindset diversity and firm performance faces the challenge that directors' mindset is not directly measurable. Much prior work has used gender as a proxy for mindset diversity (Adams (2016), Adams and Ferreira (2009)). Yet, individuals' mindsets could differ due to other deeply impactful cultural markers that are often more difficult to measure (Page (2007), Phillips and O'Reilly (1998)). Indeed, a newly proposed requirement by the NASDAQ that publicly traded U.S. firms diversify their boards with at least one woman and one under-represented minority (by race or sexual orientation) acknowledges that mindset diversity often involves facets other than gender.[2]

Our paper measures mindset diversity using a relatively unexplored instrument which relates deeply to people's sense of identity. In particular, we focus on diversity of religion and caste – an informal social institution that divides India's Hindu society into hundred of communities. We find that India's corporate boards are characterized by strikingly low levels of religion and caste diversity. We then ask how this lack of mindset diversity (proxied by religion and caste diversity) is associated with firm outcomes and show that firm performance is negatively related to caste homogeneity of boards. This negative association is strongest for our narrowest measure of identity. This suggests that the granularity with which board diversity is measured affects firm performance estimates.

Caste membership of board members is associated with mindset differences in the boardroom as it is a deeply rooted and temporally stable structural aspect of Indian society. Although there

[1] There are several notable examples of such measures. Since 2008, Norway has required public and state-owned companies to have 40% of their boards be constituted by females. As of the end of 2019, California requires all public companies to have at least one female director. Since 2013, India has required large public companies to have at least one female director.

[2] Refer to this link for more details.

are multiplicity of views on the precise historical origins of the caste system, there is scholarly unanimity that it is an important societal feature of contemporary India. Caste has been shown to affect education (Hnatkovska et al. (2012), Hnatkovska et al. (2013)), intergenerational occupational mobility (Munshi and Rosenzweig (2013)), lending (Fisman et al. (2017)), and more. As Munshi (2019) puts it, "Caste plays a role at every stage of an Indian's economic life." [3]

Caste can be measured coarsely as five categories (termed *varnas*) or finely as hundreds of categories (termed *jatis*). However, most previous studies on caste use a binary measure of "upper" and "lower" caste, in part, because there is no systematic documentation enabling researchers to identify individuals' coarse-grained caste (*varna*) and fine-grained caste (*jati*). We use a novel computational methodology to develop a data driven mapping of last names to religion and caste at varying levels of granularity (i.e. religion, coarse- and fine-grained caste). In doing so, we exploit the facts that individuals' last names are indicative of their caste, and that marriages are predominantly intra-religion and, among Hindus, intra-caste. We obtain data from three prominent matrimonial websites on the names of nearly six million registered users and their self-reported caste. Since the mapping between last name and religion/caste is not always one-to-one, we use these data to assign probabilities with which a last name belongs to each caste (at coarse- and fine-grained levels). This method helps us map 16,637 unique last names into eight religions, five coarse-grained castes, and 471 distinct fine-grained castes. We apply this mapping to directors of large public and private firms, whose names are taken from an annual firm-level database.

We use these data to develop a rich set of stylized facts about religion and caste diversity in Indian corporate boards during 1999-2015. We assess if mindset diversity (measured as the inverse of religion and caste Herfindahl-Hirschman index (HHI)) depends on how coarsely or finely directors' identity is defined and find extremely low diversity at all granularity levels. This low diversity is pervasive across states and industries throughout the sample period. We further show that the

[3] The strong influence of caste carries beyond India. In June 2020, California's Department of Fair Employment and Housing charged two Indian-origin managers at Cisco Systems, Inc. for discriminating against another Indian-origin engineer on the basis of caste.

low diversity in boards is not simply coincidental or driven by low diversity in the supply pool of directors. However, it varies systematically across firm types indicating that higher performance and better corporate governance are associated with greater diversity in boards.

Next, we investigate the relationship between diversity of boards and firm performance. Ex ante, this relationship is unclear. On one hand, homogenous board members may not bring a wide range of perspectives to bear upon the decisions they make for the firm, worsening their advisory and monitoring roles. They may also be more prone to cronyism, hurting the firms they serve. On the other hand, socially homogeneous directors may have greater trust or fewer differences in opinions, reducing conflicts in the boardroom and improving firm performance.

Regression analysis, therefore, provides us with estimates of the net effect of these mechanisms. We use several instrumental variable strategies to examine how religion and caste diversity of boards affects key accounting measures of firm performance (operating income, operating cash flow, profitability) and market related variables (Tobin's Q and volatility). To instrument for diversity, we use the diversity of a firm's director supply pools, measured as the sets of directors in the firm's state and industry. In a second approach, we additionally use as instruments the distance of a board's religion/caste composition from that of the supply pool composition. In a third strategy, we exploit a change in corporate governance requirements that induced changes in board memberships during our sample period. In a fourth strategy, we use as an instrument the fraction of a board's directors who serve on at least one other board whose dominant caste is different from the directors' own caste. This instrument, inspired byAdams and Ferreira (2009), captures the extent to which a board's directors are exposed to castes other than their own. Results from all four analyses show that diversity on corporate boards is negatively associated with firm value and performance. Importantly, we find that the strongest negative coeffcients for the fine-grained caste measure, followed by the coarse-grained caste measure. The results are weakest for religion diversity.[4] This is consistent with the reality of India's social fabric, wherein marriages, residence, occupations, voting patterns, public good provision, etc. are all influenced by the fine-grained measure (Joshi

[4]This result may be driven by the fact that we observe low levels of variation in religion diversity of boards.

et al. (2018), Kumar et al. (2017), Beteille (1996), Srinivas (1995)). We take this finding to indicate that, as researchers investigating the relation between diversity and economic outcomes, we need to consider that level of identity along which people feel the most affinity toward others. It is this identity that would most strongly influence group dynamics. We also note that there is insufficient variation in boards' gender diversity in our data.

Several studies have focused on how socio-cultural identity shapes networks (see, for example, Currarini et al. (2009)), hiring (Åslund et al. (2014), Giuliano et al. (2009), Giuliano and Ransom (2013), and Petersen et al. (2000)), and economic exchange in dyads such as lender-borrower, manager-employee, venture capitalist (VC)-entrepreneur, VC partners, research collaborators, and teacher-student (see Gompers et al. (2016), Glover et al. (2017), Shayo and Zussman (2011), Fisman et al. (2017), Bengtsson and Hsu (2015), Hegde and Tumlinson (2013), Claes and Vissa. (2020)), Freeman and Huang (2015), Dee (2005) and Fairlie et al. (2014)). Our paper is different from all of these in that we analyze the effect of mindset diversity in teams of multiple agents and their joint decisions in high stakes economic settings.

The literature on corporate governance has also analyzed the effects of board diversity, but the dimension examined is almost exclusively gender (see, for example, Adams and Ferreira (2009), Ahern and Dittmar (2012), Kim and Starks (2016), Sila et al. (2016), and Bertrand et al. (2018). Terjesen et al. (2009) provide an excellent review. A few exceptions include Giannetti and Zhao (2019) who use ancestral origins, Arnaboldi et al. (2018) and Bernile et al. (2018) who develop a multidimensional diversity index, and Kramarz and Thesmar (2013) who use shared alma maters of directors. We add to this body of work by considering the mindset diversity of directors, as measured by the traditional institutions of caste and religion. The impact of such deep rooted traditional institutions on board composition and firm performance has not been previously explored.

This paper also contributes to the literature examining the economic effects of caste. Previous studies mainly compare socio-economic outcomes of disadvantaged castes to those of advantaged upper castes (see, among others, Hnatkovska et al. (2012), Hnatkovska et al. (2013), Iyer et al. (2013), Ghani et al. (2014), Damodaran (2008), Thorat and Neuman (2012), Jodhka (2010), and

Varshney et al. (2012)). However, we approach the economic effects of caste through a different lens – does shared caste identity influence economic outcomes, regardless of whether the caste itself is underprivileged or not? Only a few studies have taken a similar approach (see Fisman et al. (2017) and Munshi and Rosenzweig (2013, 2016)). Besides examining a different economic outcome, we differ from these studies in a few important respects. First, while they focus on rural areas, specific cases, or traditional businesses, we show that caste influences economic outcomes nationally, even in urban, elite, and high-stakes corporate environments. Our focus on board composition and firm performance also distinguishes us from Chen et al. (2015) who consider caste proximity between equity analysts and CEOs and Damaraju and Makhija (2018) who consider caste proximity between CEOs and firm owners or chairpersons.

The rest of the paper is organized as follows. In Section 2, we describe our data. Section 3 presents stylized facts about religion and caste diversity in corporate boards. In Section 4, we describe our empirical strategy to estimate the relationship between religion and caste diversity and firm performance. Results are presented in Section 5. Section 6 explores the mechanisms underlying our results. Section 7 concludes.

2 Building the Database

We combine data on names and religion and caste identity from matrimonial websites with data on Indian firms and their boards, as described below.

2.1 Data from Matrimonial Websites

We develop a novel dataset mapping Indian last names to religions and castes.[5] This enterprise exploits two features of Indian society – last names are indicative of religion and caste (Dumont (1980)), and marriages are overwhelmingly intra-religion, with Hindu marriages being intra-caste.

[5]Our work is the first academic exercise that links individual Indian last names to a likelihood distribution over religion, broad (and fine) caste. Online Appendix Section 1 lists other data sources that collect caste information.

Both aspects are reflected in individual profiles on matrimonial websites, where prospective suitors self-report their names, religions, and castes (specifically, fine-grained castes).[6] We obtain data on over six million such profiles from three popular matrimonial sites – *Shaadi*.com, *Jeevansathi*.com and *Bharatmatrimony.com*.[7] Specifically, we have information on users' first and last names, native language, religion, and fine-grained caste.[8] In our final cleaned data, we have 5,447,129 profiles, spanning 16,637 unique last names, eight religions, five coarse-grained castes, and 471 fine-grained castes.

The same last name may be associated with more than one religion or caste, often depending on the geographical region. We exploit the fact that different regions of India have different dominant languages. Therefore, our mapping from last name to caste is conditioned on users' self-reported first language.

We also account for the possibility that the same last name may have different spelling variations. The vast majority of names in our dataset are words from Indian vernacular languages like *Hindi, Tamil, Marathi, etc.*, whereas the websites from which our data are culled use the English language. Therefore, we have an English equivalent (not translation) of last names in Indian languages. This may result in multiple English spelling variations of the same last name. To address this, we use two different word matching algorithms (Levenshtein distance and Ratcliff/Obershelp pattern matching) to predict the similarity of differently spelled last names. If the similarity predictions from both algorithms are above a certain threshold, then the two last names are considered to be the same and their caste mappings are combined.[9] This name matching results in groups of

[6]Since people want to marry within caste, users who do report their caste have an incentive to report it truthfully.

[7]In India's fast evolving socio-economic setting, where the role of traditional kinship networks and local matchmakers is rapidly diminishing, matrimonial websites act as an alternative to traditional marriage brokers by nationalizing the pool of prospective spouses. According to a study conducted by KPMG in 2016, the Indian online matrimonial market was expected to rise from USD 0.11 bn to USD 0.26 bn by the end of 2022, growing at an annual growth rate of 19%.

[8]To build a robust mapping, we drop all last names that appear only once in the database.

[9]Our first algorithm is a modified Levenshtein distance algorithm. In this method, a distance measure between two strings is calculated using a dynamic programming approach, with each replacement letter adding one to the distance

similar sounding last names that have different spellings.

Following this approach, we assign all spelling variants of a same last name (conditional on same language) a frequency equal to the number of records with that last name in our data.[10] We take all religions and fine-grained castes reported for a last name across all corresponding records and count their respective occurrences. Dividing the number of times a last name is associated with a particular religion and caste by the total number of times the last name appears for a given language in the database gives us the probability with which the last name belongs to a particular religion and caste. Doing so over all religions/castes associated with a last name gives us the probabilistic mapping of each last name to religions and fine-grained castes.[11] For names identified as non-Hindu (3070 in number), we assign the corresponding religion instead of caste.

measure. The standard Levenshtein distance measure is appropriate for English words, whereby each difference in letters between two words contributes equally to the distance measure. However, since we want to match the phonetic translation of Indian languages, we develop a modified Levenshtein distance algorithm. In this method, differences in letters that constitute the same sound in Indian languages are assigned a zero distance measure. The final distance calculated using this algorithm provides a measure of how similar two words are in an Indian language. In our approach, two words that have a distance of less than three are deemed to have the same spelling in the Indian language and, hence, be the same name. Our second method is a modified version of the Ratcliff/Obershelp pattern matching algorithm. This algorithm looks for matches in the longest contiguous matching subsequence of letters in two words and assigns a matching score. The original algorithm was developed to find sequence matches between two sentences, and was found to be appropriate for matching words that are phonetic translations from other languages. For our purposes, if the algorithm provides a match score of greater than 85%, the two last names are deemed to be matched. As a final step, we consider two last names to be fully matched if both algorithms predict that the corresponding strings are matched. This approach is borrowed from the concept of bagging used in the machine learning based classification literature where votes from different classifiers are used together to increase the robustness of the final prediction (Friedman et al. (2001)).

[10]The resulting average total count for each last name is 362.68.

[11]In one region of southern India (the state of Tamil Nadu), the last name of an individual is often their father's first name. Hence, the last name changes across generational cohorts of a family. For these cases, we map the first name of the individual to their caste instead of the last name. Though Tamil first names do vary systematically by caste, we expect the caste mapping of Tamil names to be noisier than the mapping based on last names for the rest of India, since the latter remain invariant across generations.

We also map last names to coarse-grained caste. For this, we rely on Government of India reports, Wikipedia entries, and several other sources to build a mapping from fine- to coarse-grained caste.[12] In our final mapping, last names are associated with an average of 1.57 and a maximum of 6 religions, with 56.59% of last names being associated with a single religion. The probabilistic mass is concentrated in the top two most likely religions, with them jointly accounting for about 99% of the total likelihood on average. As for caste, last names are associated with an average of 3.6 (10.4) and a maximum of 11 (138) coarse (fine)-grained castes.[13] Table 6 shows the most likely religion, coarse- and fine-grained caste composition in the final matrimonial sample. Hindus, at 80.69%, form the overwhelming majority of names. Muslims and Christians, respectively, account for the next highest proportions. Other religions together constitute about 5% of the sample. This distribution is not far from the religion composition of the aggregate population of India.[14]

2.2 Firm Level Data

We obtain firm-level data for 1999-2015 from Prowess, a database provided by the Centre for Monitoring the Indian Economy (CMIE). The data provide information on financials and corporate governance of large firms, as reported in their annual reports, quarterly financial statements, and

[12]The caste system divides Hindu society into four hierarchical categories – *Brahmin*, *Kshatriya*, *Vaishya*, *Shudra*. There is an additional de facto fifth category of *Dalit*. Our mapping assigns every last name a probability mapping over these five categories. For Non-Hindu last names, we assign the corresponding religion as their coarse- and fine-grained caste. Additionally, we assign "unknown" to names for which we are unable to find a fine- to coarse-grained caste mapping, or if a fine-grained caste maps to multiple coarse-grained castes.

[13]The coarse- and fine-grained classification of individuals is a practice primarily followed by Hindus. While building the mapping of last names to coarse - and fine-grained castes, we do not drop non-Hindu last names. Instead, we consider non-Hindu religions as additional distinct categories for our last names to caste mappings.

[14]According to the 2011 census, Hindus constitute 80% of the population, Muslims are the next largest group (14.2%), followed by 2.3% Christians. Other religions together account for just over 5% of the total population. The census does not provide population composition by coarse- or fine-grained caste. In our data, however, we see the maximum representation of the coarse caste category of *Shudra* followed by *Brahmin*. For space considerations, table 6 only reports the percentages of last names for the top eleven most frequently occurring fine-grained castes in the data.

profit and loss accounts. The database includes large publicly listed and unlisted firms.[15] While the database does not cover the universe of all firms, those included account for a substantial proportion of economic activity; in 2009, they contributed 84% of GDP, 55% of exports, 70% of imports, 47% of the total output of non-agricultural and non-government services sector, and 58% of all corporate taxes and all excise taxes collected by the government. Although the database follows firms longitudinally, most firms appear in the data only for a few years. Thus, for most of our analyses, we treat these annual data as repeated cross sections of firms. We deflate all nominal variables by all-India CPI (2001=100).

To identify the religions and castes of firms' directors, we match their last names to those in the matrimonial data, assigning each matched director last name the same probabilistic distribution over religions and castes as that constructed using the matrimonial data and as described in Section 2.1. We retain only those firm-year observations for which we can (probabilistically) identify the religion/caste of all board members.[16] We also consider only those firm-years that have at least two directors serving on their boards. In our matched sample, we have 23,819 unique firms with a total of 576,579 directorships.[17]

Table 5 profiles the religion and caste mapping for directors. For religion, we see that 69% of directors' last names are associated with a single religion, and 92% names are associated with up to two religions. As expected, the mapping for coarse- and fine-grained caste is noisier, with 45% (40%) last names associated with a single coarse- (fine-) grained caste and 85% (65.5%) last names associated with up to five coarse- (fine-) grained castes. Although the mapping is probabilistic, the probabilities are front loaded. The top two most likely religions account for 99.5% of the total

[15] As of 31 March, 2009, 7,86,774 companies were registered with the Registrar of Companies, an administrative arm of the Ministry of Corporate Affairs. Of these, 82,058 were public limited companies and 704,716 were private limited companies. Of the 82,058 public limited companies, Prowess contains information on about 24,000 companies.

[16] There are a few directors whose names suggest that they may not be of Indian origin. We are unable to match these names with those in the matrimonial database. These directors are assigned a religion, coarse-, and fine-grained caste category of "NA."

[17] As a percentage of firms and directors in the Prowess database over 1999-2015, we are able to fully match 63.52% of firm-year boards and 57.11% of all directorships.

likelihood, on average. Similarly, the top five most likely coarse- (fine-) grained castes account for 99.3% and 89% of the total likelihood, on average. Table 1 of the Online Appendix provides the religion and caste composition of directors for the first and last years of our sample.[18]

A few key firm-level characteristics are provided in Table 2. We note that the highest proportions of firms at both the start and end of our sample period belong to manufacturing, and finance, insurance and real estate sectors. Under half of them are listed on India's stock exchanges and about half are members of business groups. The mean real assets of these firms were about Rs. 18 million in 1999 and Rs. 23.3 million in 2015. The average board size is about 5.4. Note that although we report the percentages of firms with dual CEOs and the mean percentages of independent directors, this information is sparse. Specifically, we can identify whether CEOs are dual for only 46.14% of our sample, and calculate the percentage of independent directors on boards for only 21.38% of the sample. Therefore, we are unable to use these board characteristics in our main empirical analyses.

Homophily Index: To measure the degree of diversity we calculate the HHI for religion, coarse- and fine-grained caste for every board. The religion/caste HHI for a board is the sum of squared shares of directors belonging to the various identities represented on the board. For example, consider a board with five board members – three Hindus and one each with the most likely religion as Muslim and Christian. The religion HHI of this board is $0.44(= (0.6)^2 + 2 * (0.2)^2)$. Continuing with this example, suppose that of the three Hindu board members, the most likely coarse-grained caste of two of them is *Brahmin* and the other is *Kshatriya*. The coarse-grained caste HHI (replacing coarse-grained caste by the religions of non-Hindu directors) for this board is $0.28(= (0.4)^2 + (0.2)^2 + 2(0.2)^2)$. Finally, suppose that the two *Brahmin* directors have the fine-grained castes *Brahmin Iyer* and *Pandey*, and the *Kshatriya* director has the fine-grained caste

[18]Looking at coarse-grained caste, in both years, *Vaishyas* dominate the sample, accounting for 26.6% and 28% of directors, respectively. Looking at fine-grained castes, *Agarwals*, who belong to the coarse caste category of *Vaishyas*, are the most represented, constituting just over 11% of directors in both years.

Khatri. In this case, the fine-grained caste HHI (replacing fine-grained caste by the religions of non-Hindu directors) for this board is $0.2 (= 5 * (0.2)^2)$.

A higher HHI with respect to an identity represents lower diversity, or greater homophily, in a board. This example illustrates that for the same board composition, the measure of concentration reduces as the lens of identity becomes finer from religion to fine-grained caste.

2.3 Indian Boards Database

Prowess does not identify unique individuals serving as directors. Since more than one director may have the same name, Prowess cannot be used to identify unique directors. To do this, we use the Indian Boards Database, maintained by the Prime database group, which provides a unique identification code for each individual serving as a director, along with the firms on which he/she serves, for the period 2012-2015. This helps us measure the degree of interlocks across boards as well as construct an instrument that closely follows Adams and Ferreira (1999). Using these data, we have information on 17,608 unique directors across 1,501 firms.

3 Diversity in India's Corporate Boards

In this section, we present several stylized facts about diversity in Indian corporate boards for fine-grained caste. Those for religion and coarse-grained caste are similar and included in the Online Appendix.

3.1 Diversity in Corporate Boards is Systematically Low

India's corporate boards are not diverse – the average homophily index is high, at 0.87 for religion, 0.56 for coarse-grained caste and 0.45 for fine-grained caste. To assess if this is simply a result of the caste or religion composition of the entire pool of directors, we compare the observed diversity levels to those in several random simulated samples. In the first method, we consider all directors across all firms in a year as the potential pool of directors available to each firm in that year.

From this "supply pool" of directors, we randomly assign directors to each firm, equal in number to its observed board size. We create hundred such simulated samples of boards for each year, calculate the mean board homophily for all boards across the hundred iterations, and compare it to the corresponding mean in the observed data. In the second (third) method, we define the director supply pool for a firm in a year as the set of all directors in that year across all firms in the same state (two-digit industry) as that firm.[19]

Figure 1 shows the yearly means of firms' fine-grained caste homophily for the observed and simulated samples. For the simulated means, we also present the 5% confidence intervals. The figure presents these means for all three approaches described above: unconditional, conditional on firms' state, and conditional on firms' industry. In all cases, we see that the mean observed fine-grained caste homophily of boards is significantly higher than the corresponding simulated mean in every year. Similar results are presented for coarse-grained caste and religion in Figures 1 and 2 of the Online Appendix. Table 2 of the Online Appendix also presents hypothesis tests for comparisons of observed and simulated homophily means. The t-statistics are large, indicating that the observed mean homophily is significantly different from the simulated means.

Further results presented in the Online Appendix show that diversity in boards has been persistently low over time across states and sectors. Figures 3, 4, and 5 of the Online Appendix respectively present state-wise mean fine-grained caste, coarse-grained caste, and religion HHIs for the first and last years of our sample. While the average homophily stayed high in most states over the sample period, states did change their relative quartile positions in the overall distribution. Sectoral average board diversity also stayed low over the sample period (Figures 6 and 7 of the Online Appendix). The sectors of professional, technical, and administrative services, and arts and recreation have the highest homophily levels, while the information, communication, and real

[19] As mentioned earlier, we cannot identify individual directors since we do not have unique identification codes for them. So, we do the simulations by defining the supply pool in two ways. In one approach, we consider every name as a distinct director, i.e., we consider directorships rather than directors. Alternatively, we consider all occurrences of the same name as the same individual director. We present results from the first approach in the paper. Results from the second approach are extremely close and available upon request.

estate, diversified, and health and education sectors have the least homophily.

3.2 Diversity is Higher in Higher Quality Firms

Figure 2 shows that larger firms have more diverse boards. Panels (a), (b), and (c) of Figure 2 show the evolution of average fine-grained caste homophily in firms across assets, sales, and profits quartiles, respectively. We see a systematic pattern of higher diversity (lower homophily) as we move from lower to higher quartiles of assets and sales. For profits, the second quartile firms have lower diversity (higher homophily) than firms in the first quartile. As before, the differences across quartiles in all three panels are small in magnitude, and the fluctuations over time within each quartile are negligible. Similar results hold for diversity along coarse-grained caste and religion (see Figure 8 of the Online Appendix).

Figure 3 demonstrates that older firms have board members from more diverse fine-grained caste backgrounds. Figure 4 shows that exporting firms have significantly more diverse boards, on average, than non-exporting firms. Finally, we observe lower fine-grained caste diversity, on average, among firms that do not belong to business groups relative to those that do. This is noteworthy since one might expect that business groups in India, that are often dominated by a single extended family, would tend to hire directors from among their kin. In that case, family ties would drive the lower caste diversity on the board. However, we see the opposite. This may be suggestive of the greater productivity, size, and prominence that is associated with business groups, enabling or incentivizing these firms to have more diverse boards. We see little difference in average board homophily across other firm characteristics such as public versus private firms and government versus non-government firms. Results for coarse-grained caste and religion homophily are presented in the Online Appendix (Figure 9, 10, and 11).

3.3 Corporate Governance and Diversity

Figure 5 shows the association between fine-grained caste and two features of corporate governance: size of the board and proportion of independent directors on the board. Figure 5(a) presents

average fine-grained homophily for firms with different board sizes, grouped into four quartiles. We see that firms with larger corporate boards have higher caste diversity among their directors. Figure 5(b) shows a negative association between average board caste homophily and the average proportion of independent directors across listed firms in one-digit sectors in the year 2015. Sectors with the highest fine-grained caste homophily levels such as trade and finance also have the lowest proportions of independent directors in their corporate boards. Results for coarse-grained caste and religion homophily are presented in the Online Appendix (Figure 12 and 13).

4 Empirical Strategy

Next, we examine whether and to what extent low caste diversity in firm boards impacts key measures of firm performance. In this section, we describe our empirical strategy.

Consider the following regression equation:

$$P_{it} = \beta_0 + \beta_1 H_{it} + \beta_2 X_{it} + \beta_3 B_{it} + \beta_4 I_{jt} + \delta_1 I_j + \delta_2 T_t + \varepsilon_{ijt} \tag{4.1}$$

where P_{it} denotes the value or performance of firm i in year t, H_{it}, the key variable of interest, is the religion, coarse- or fine-grained caste homophily of firm i's board in year t, X_{it} is a vector of time varying firm characteristics, B_{it} is a vector of time varying board characteristics, I_{jt} denotes a vector of time varying industry characteristics, I_j denotes a vector of two-digit industry (National Industrial Classification (NIC) 2008) fixed effects, and T_t is a vector of year fixed effects. We cluster the standard errors by year and industry (two-digit) and correct them for arbitrary heteroskedasticity.

Our dependent variables (P_{it}) are accounting measures of performance including operating income, operating cash flow, and profits (all relative to sales and in natural logs), and market based indicators including Tobin's Q, and firm volatility. Operating income is defined as the difference between sales and operating expenses. Operating cash flow is the cash flow from operating activities before depreciation. Tobin's Q is calculated as the market value of a firm divided by the

replacement value of the firm's assets. Volatility is measured as the standard deviation of the returns on a firm's security over a year.[20]

Firm level control variables (X_{it}) include firm age, firm size, tangibility, book leverage, and indicators for whether the firm is listed on the stock market, whether it belongs to a business group, and whether it is an exporter. We define a firm's age as the number of years since incorporation of the firm. Firm size is measured by real assets (in natural logs). Tangibility is defined as the fraction of tangible assets in the total assets of a firm. Book leverage is calculated as the ratio of the total debt of a firm to its total assets.[21]

Board characteristics (B_{it}) include size of the board and its square, where size of the board is defined as number of directors on the board of a firm in a year. Coles and Naveen. (2008) et al. find that big firms which have greater advising requirements have larger boards than small firms. They also find that the relation between Tobin's Q and board size is U-shaped. Therefore, we include a quadratic in board-size across all regressions in our analysis. Different social groups have different levels of cohesion or sense of affinity. To account for this we include a vector of indicator variables to control for the dominant coarse-grained caste (for Hindus) and religion (for non-Hindus) on the firm's board. The dominant religion/caste of the board is the most likely religion/caste for the largest number of board members.

Time varying industry characteristics (I_{jt}) include market-concentration of the industry (measured as the three-digit NIC HHI at the national level), proportion of group firms in the industry (three-digit NIC), and the shares of firms in the industry (three-digit NIC) with a particular dominant religion/coarse-grained caste.[22] Including the industry's market-concentration helps us con-

[20] We measure Tobin's Q and firm volatility in two ways. One measure uses stock prices of the entire year between two annual reports and the other uses stock prices for a month around the reporting date. We present results for the former but results remain close for the latter method. Some firms' stocks are traded on the National Stock Exchange (NSE), and some are traded on the Bombay Stock Exchange (BSE), and some on both. Throughout the paper, we use firm returns at NSE when their stocks are traded on only NSE or both, and BSE when they are only traded on BSE. Results using BSE returns for firms whose stocks are traded on both markets are close.

[21] All financial variables are winsorized at 1% and 99% for the entire sample period.

[22] Alternatively, we also estimate industry HHI at two-digit NIC-state level. Results remain close.

trol for the competitiveness of the industry which can influence a firm's performance. Controlling for the proportion of group firms in the industry allows us to account for the possibility that there may be firms with close ties that adopt similar practices, which, in turn, affect a firm's performance. Controlling for the shares of firms in an industry with various dominant religions or coarse-grained castes allows us to account for the caste/religion composition of firms in the industry.

In the above regression, β_1 captures the association between religion or caste homophily and firm performance. However, this coefficient is not a causal estimate since homophily is an endogenous regressor. The endogeneity can result from both omitted variable bias and reverse causality. An unobservable time varying firm characteristic (for example, adoption of new management practices) can drive both homophily and firm outcomes. Firm performance can also influence homophily. For instance, as a firm's value grows, it may become increasingly prestigious for directors to serve on its board. This can influence board composition.

To overcome this endogeneity, we employ four instrumental variable strategies. In the first approach, we use two instruments: (1) the religion/caste HHI for all directors in the two-digit industry that the firm belongs to,[23] (2) the religion/caste HHI for all directors in the state where the firm is located.[24] In the second approach, we augment our list of excluded instrumental variables with: (3) the Euclidean distance of the vector representing the board's religion/caste composition

[23] A more disaggregated classification is unsuitable for two reasons. First, directors may not serve on closely competing firms' boards due to conflict of interest. Second, the narrower the classification level, the fewer the number of firms in each industry so that the influence of each firm in determining the overall pool of directors in the full industry may be high, invalidating the instrument. A less disaggregated classification level, on the other hand, is undesirable as it will not yield enough variation in the industry level homophily index.

[24] We measure homophily of state and industry level directors in two ways. In the first approach, each name is considered to represent a distinct director. In doing so, we effectively measure the homophily of directorships rather than unique directors. In the second method, we assume that all occurrences of the same name represent the same unique director and measure homophily using unique names in a state/industry. In the paper, we present results using the first approach. Results using the second approach are extremely close. These two variables provide us a measure of the religion/caste composition of the set of directors that constitute the firm's "supply pool," as described in Section 3.1.

from that of the full set of directors in the corresponding industry, (4) the Euclidean distance of a board's religion/caste composition vector from that of the full set of directors in the corresponding state. In the third approach, we construct an instrument that exploits changes in board memberships necessitated by a set of requirements announced by the Securities and Exchange Board of India (SEBI), commonly referred to as "Clause 49." In the fourth approach, similar to Adams and Ferreira (2009), we instrument for board homophily with the fraction of the firm's directors with board membership in at least one other firm whose dominant caste is different from that of the directors' own caste.

Approach 1: Homophily of State and Industry Director Supply Pools

The intuition for using state and industry director caste HHIs as the excluded instruments in the first two approaches is that a firm's board composition may be similar to that of other firms in the same industry or state. Previous studies have shown that both geography and industry influence the supply of directors that firms can choose from (see Knyazeva et al. (2013) and Dass et al. (2013)). We show that this holds in our setting too by documenting that (a) a non-negligible proportion of directors on a board are also directors of other firm(s) in the same industry and state, and (b) the religion/caste composition of directors on firm boards is very similar to that in the industry or state. Table 3 documents within-industry board interlocks for one-digit industries for the year 2015.[25] To identify these interlocks, we use the Indian Boards Database which, unlike Prowess, allows us to identify unique directors, albeit for a smaller sample of firms. Using these data, we identify a within-industry interlock if a director on a firm is currently, or has been in the past, a director on at least one other firm that belongs to the same industry. We then calculate the percentage of all directors in a firm that are interlocked within-industry. We observe that the average interlocks range from 0% to 31% across these broad industries, while the maximum degree of interlocks can be as high as 100%. Looking at two-digit and three-digit industries, we see that interlocks are

[25]For space considerations, we present this evidence only for 2015. Similar results hold in other years of our sample period.

present even at these narrower levels, albeit to a smaller degree. The mean interlock in two-digit industries in 2015 is 5.2% (3.4% in three-digit), although the maximum interlock is over 80% in many industries.

Nonetheless, there are several firms with no directors that serve (or have served in the past) on other firm(s) in the same broad industry. However, even across these firms, the religion/caste composition of directors is similar to that of directors in the industry. We show this by comparing the distribution of directors in firms that have below (and above) median interlocks to that of the industry using two Kolmogorov-Smirnov tests (K-S tests).[26] In the first test, we compare the distribution of the top religion/caste of all the unique directors in the pool of firms with interlocks below (and above) median to that of the entire industry. In the second set of K-S tests we compare the distribution of the dominant religion/caste of firms with below (and above) median interlocks to that of all the firms in the entire industry. Results for fine-grained caste composition from these tests are presented in Table 3 of the Online Appendix. The table shows that for one-digit industry, we are unable to reject the null hypothesis that the samples of directors in firms below (and above) median and the aggregate industry are drawn from the same distribution. The same conclusion is reached when we alternatively look at the samples of firms according to their dominant coarse-grained caste and religion.

The validity of using state and industry director caste HHIs as the excluded instruments is plausible for several reasons. First, to the extent that industry and state level homophily indices are associated with some unobservable characteristics of the industry or state that can have an independent effect on firm performance, that possibility is controlled for by including state and industry fixed effects. Note, however, that we are unable to include both sets of fixed effects simultaneously, in addition to year fixed effects and other time-invariant firm characteristics including listing and export status. This is because, the number of firms within the resulting cells is often small so that we do not have enough variation left in a large proportion of cells in the samples.

[26]The Kolmogorov-Smirnov test (K-S test) examines the null hypothesis that two samples are drawn from the same continuous, one dimensional probability distribution.

Second, as explained above, we define the industry broadly at the two-digit level. The number of firms in a two-digit industry tends to be large, so that any single firm is unlikely to strongly influence homophily among the set of directors in the entire industry. Analogous intuition applies to the state-level homophily index.

Third, we control for several other firm and industry characteristics in our regression to ensure that the excluded instruments only affect firm performance through board diversity or through their correlation with included controls. Directors and firms could match in an assortative manner based on factors such as reputation, causing an omitted variable bias invalidating the instruments. We measure firm performance variables relative to size (as proxied by firm sales) and additionally control for firm assets. We also include other firm characteristics – it's listing and exporting status, leverage, tangibility, age, whether it belongs to a business group and the dominant coarse-grained caste of the firm. This last firm characteristic also helps account for the possibility that some caste/religion groups are more cohesive than others such that diversity may have different implications for firm performance. In a robustness check, we split our sample into sub-groups of firms dominated by various coarse-grained castes (for Hindus) and religions (for non-Hindus). Results are presented in Section 5. [27]

We also control for industry characteristics to account for the possibility that the industry is dominated by similar castes or closely related firms or has low competition. All of these could be related to both supply pools and individual firm performance. In particular, we control for industry's market competition, proportion of business group firms, and the shares of firms dominated by various castes or religions. In another robustness check, we adjust for industry by taking differences of firm performance variables from the corresponding industry median and regressing them on the above-mentioned control variables. Results are presented in Section 5.

[27] To control for firm ownership beyond whether it belongs to a business group, in one specification we additionally include whether the firm is a family firm, percentage of major equity investments coming from institutional investors, and the proportion of promoters on the board. Due to data limitations these variables are measurable only for a considerably smaller number of firms. Second stage results for the second IV approach show that lower diversity in boards leads to significantly worse firm performance. We find that CEO duality is negatively associated with firm performance and that a firm that has greater institutional inverstors is associated with better firm performance. A one unit increase in fine-grained caste homophily leads to almost 0.28 log points drop in operating income and 0.22 point drop in Tobin's Q.

Approach 2: Euclidean Distance between Caste Composition of Board and State and Industry Director Supply Pools

In our second instrumental variable approach, we additionally use the distance between a firm and its industry/region with regard to their religion/caste composition. Note that several different compositions can yield the same homophily index. So whether a firm's board composition is similar to that in its industry/state can be determined not only by comparing its overall homophily index with that of the industry/state but also its underlying composition. The larger this distance, the less similar is the firm's director composition to that in the industry. Since these additional Euclidean distance based measures vary across firms (and over time), instead of only across industries or states, the relevance of our set of instruments also increases. Table 4 demonstrates that Euclidean distances between firms and industry/state level fine-grained caste composition of directors vary considerably, but are generally quite small. Panel A of the table shows key moments of the distances between firms' director composition and industry director composition for four years over the sample period. We see that the distribution of these distances is quite stable over time. In all years, the mean distance is slightly larger than the median, indicating that the distribution has a heavier right tail. Relative to the magnitudes of these distances, the standard deviation is quite large, suggesting considerable variation within years. Similar patterns are evident for distances between firms' and state director compositions (Panel B).

The Euclidean distance between a board's religion/caste composition and that of the aggregate set of directors in the corresponding industry or state also meets the exclusion criterion. The reasons described above for the validity of the state and industry-level homophily indices also apply to the distance measures. We do an additional robustness check to further account for the possibility that firm performance itself may influence Euclidean distance of its board from the supply pools. We split the sample into firms that are above and below the industry median of a performance measure and estimate the same regression for these sub-samples. Results are presented in Section 5.

Approach 3: Clause 49

In a third approach, we exploit board membership changes induced by firms complying with Clause 49 of a new set of corporate governance regulations announced by the Securities and Exchange Board of India (SEBI) that went into effect in February 2000. Among other things, the new requirement was for firms to have at least 50% of their board be comprised of non-executive members. The compliance deadlines differed for different groups of firms; March 31, 2001 for Group 1 – the largest ("Group A") companies listed on the Bombay Stock Exchange (BSE) and National Stock Exchange (NSE) S&P CNX Nifty Index companies, March 31, 2002 for Group 2 – other companies with paid-up share capital of at least Rs. 100 million, or net worth of at least Rs. 250 million, at any time in the firm's history, March 31, 2003 for Group 3 – firms with paid-up share capital of at least Rs. 30 million, and Group 4 – any newly listed or re-listed firms at the time they get listed. Clause 49 did not apply to older firms with lower net worth or paid-up capital. To construct the instrument, we closely follow Helmers et al. (2017) and use a regression discontinuity design where we compare firms just below and above the threshold of 50% board members being non-executive along with whether they were required to comply with the clause requirements. Therefore, the treatment depends on the interaction of whether a firm was required to comply and whether a firm was below the threshold of 50% non-executive directors. Similar to Helmers et al. (2017), our first stage equation is:

$$
\begin{aligned}
H_{it} &= \alpha_0 + \alpha_1 BT_{it} * I(Group1)_i + \alpha_2 BT_{it} * I(Group2)_i + \alpha_3 BT_{it} * I(Group3)_i + \alpha_4 BT_{it} * I(Group4)_i \\
&+ \alpha_5 BT_{it} + \alpha_6 I(Group1)_i + \alpha_7 I(Group2)_i + \alpha_8 I(Group3)_i + \alpha_9 I(Group4)_i \\
&+ \beta_1 NonExecProp_{it} + \beta_2 (NonExecProp_{it})^2 + \beta_3 ShareCapital_{it} + \beta_4 (ShareCapital_{it})^2 \\
&+ \beta_5 NetWorth_{it} + \beta_6 (NetWorth_{it})^2 + \beta_7 (NonExecProp_{it} * ShareCapital_{it}) \\
&+ \beta_8 (NonExecProp_{it} * ShareCapital_{it})^2 + \beta_9 (NonExecProp_{it} * NetWorth_{it}) \\
&+ \beta_{10} (NonExecProp_{it} * NetWorth_{it})^2 + \delta_1 X_{it} + \delta_2 I_{jt} + \delta_3 I_j + \delta_4 T_t + \varepsilon_{ijt} \qquad (4.2)
\end{aligned}
$$

where BT_{it} denotes a binary variable that equals 1 if firm i is below the 50% non-executive directors threshold in year t and 0 otherwise, $I(GroupN)$ are binary variables for whether a firm i is in a group of firms required to comply with Clause 49, for $N = 1,2,3,4$, $NonExecProp_{it}$ is a continuous variable measuring the proportion of non-executive members on firm i's board in year t, $NetWorth_{it}$ is the net worth of firm i in year t and $ShareCapital_{it}$ is firm i's paid up share capital in year t.[28] The remaining control variables are as in equation (4.1). Note that the coefficients on the excluded instruments indicate the intent-to-treat since there may not be perfect compliance with the regulation. To estimate this regression, we build a longitudinal sample of firms for the period 1999-2007 such that we can observe a firm for at least two consecutive years. We cut off the sample period in 2007 since in 2008, a new requirement around independent directors was included in the amended Clause 49 which would induce other changes in board membership that would be correlated with the changes we are focusing on. Following Lee and Lemieux (2010) and Helmers et al. (2017), we include all firms in the sample but additionally include polynomials of the continuous running variables and their interactions as shown in the equation above. We report results for a quadratic polynomial.[29]

The regulation only specified the percentage of non-executive directors that a firm must have. So the firms were free to choose directors of any caste or religion. The directors that they hired to comply with Clause 49 would influence the board caste/religion homophily, our endogenous regressor. Exclusion restriction on this set of instruments is also plausible given the regression discontinuity design. We further include controls to ensure the validity of the instrument. The choice of who to hire may be correlated to the industry caste composition, industry competition, and other observable or unobservable industry characteristics which also matter for firm performance. Thus, we include industry fixed effects, industry HHI, proportion of business group firms in the industry and shares of firms dominated by various coarse-grained castes/religions. Board hires may also be

[28] Note that Group 2 firms were those with at least Rs. 250 million in net worth at any time in their histories to comply. Thus, for Group 2 firms, for the first year in our sample period, we look into each firm's records in all previous years and take the maximum net worth across those years.

[29] We also estimate results for firms falling in narrow bandwidths around the cutoff of 50% non-executive directors. The main coefficient of interest (on board homophily) remains qualitatively similar but the instruments are weak for some of these smaller samples for some performance variables.

impacted by several firm characteristics that correlate with performance. We also control for these firm characteristics as described earlier for equation (4.1).

Approach 4: Exposure to Diversity Outside the Board

In our last strategy, we construct an instrument similar in spirit to the one developed by Adams and Ferreira (2009). In that study, the authors examine the firm performance effects of gender diversity on the board, an endogenous regressor. They instrument for the latter by calculating the fraction of a firm's male directors that serve on other boards that have female directors. The intuition is that connections of a firm's male directors with female directors on other boards is likely positively correlated with gender-diversity on the firm's board. Following this logic, we measure the fraction of a firm's directors who serve on at least one other board whose dominant fine-grained caste is different from their own. We expect such exposure to caste diversity to be negatively correlated with caste homophily on the firm's board, our endogenous regressor.

Several control variables help us ensure the validity of this instrument. Directors' exposure to other castes may itself be influenced by the caste composition of the industry or other characteristics of the industry, both of which correlate with firm performance. To account for these possibilities, we include industry fixed effects, industry's HHI, the share of group firms in the industry, and a vector of shares of firms in the industry dominated by various coarse-grained castes (for Hindus) and non-Hindu religions. Several firm characteristics that influence firm performance may also drive the connectedness of the board. Thus, we include a vector of several firm characteristics as described earlier for equation (4.1) and, additionally, control for the total other directorships held by the directors of the firm.

Note that construction of this instrument relies on information in the Indian Boards Database. Using these data we can identify which other firms' boards a director serves on and then measure the dominant castes of those boards. Based on the information available, we are able to calculate this instrument only for 936 publicly traded firms.

5 Homophily and Firm Performance

5.1 Fixed Effects Regressions

We present fixed effects regression results to assess how firm performance relates to fine-grained caste homophily of boards of directors in Table 7. Columns (1)-(3) show results for the association between homophily of boards and accounting measures of firm performance – log (operating income/sales), log (operating cash flow/sales), and log (profits/sales), respectively. Columns (4)-(5) present analogous results for homophily and market based measures – Tobin's Q, and volatility, respectively. We observe that firm performance measures are worse for firms with less diverse boards, i.e., those with higher homophily. A one unit increase in fine-grained caste homophily is associated with 0.26 and 0.225 log points fall in operating income and profits, respectively, relative to sales. Tobin's Q is also negatively associated with homophily. Higher board homophily is also significantly correlated with greater stock market volatility for the firm. A one unit increase in fine-grained caste homophily is associated with a 0.008 increase in the standard deviation of the firm's stock market returns, on average. The corresponding associations between religion and coarse-grained caste homophily and firm outcomes are similar, although the estimated coefficients are smaller in magnitude (see Tables 4 and 6 in the Online Appendix.)

5.2 Instrumental Variable Analysis

In this section, we present results for our four instrumental variable strategies.

Approach 1: Homophily of State and Industry Director Supply Pools

Table 8, columns (1) and (2), present the first stage results for the first instrumental variable (IV) approach. Since the samples differ somewhat due to missing observations for the second-stage dependent variables, there is a different first stage regression equation estimated for each of the five dependent variables that we consider. For space considerations, we do not show all first stage equations. Instead, Table 8 presents the first stage results for two dependent variables: log(profits/sales)

(column 1) and volatility (column 2).[30] The table also shows coefficients only for the excluded instruments. The excluded instruments in columns 1 and 2 are the fine-grained caste homophily levels of the industry and state director supply pools for a given firm. The table shows that both instruments are strongly and significantly positively associated with an average firm's board homophily. The corresponding first stage F-statistics are well above 10, indicating that the instruments explain a significant proportion of variation in the endogenous regressor. Tables 5(B) and 7(B), columns (1) and (2), in the Online Appendix present similar first stage results when the first stage dependent variable is religion and coarse-grained caste homophily of boards, respectively.

Table 9 presents second stage results for the first IV approach. The estimated coefficients on fine-grained caste homophily show that all performance measures are worse when boards are less diverse. For example, operating cash flows fall by 1.6 log points and Tobin's Q by 2.32 points when fine-grained caste homophily increases by one unit. Volatility also increases 0.007 points, although the estimate is not statistically significant.

Approach 2: Euclidean Distance between Caste Composition of Board and State and Industry Director Supply Pools

Next, we consider results from our second instrumental variable approach, in which the excluded instruments are the homophily of state and industry level director supply pools as well as the distance between the caste composition of these supply pools and that of individual boards. Table 8, columns (3) and (4), show the first stage results for two second stage dependent variables – log(profits/sales) and volatility, respectively. We observe that greater homophily in both state and industry supply pools is significantly positively correlated with an average firm's board homophily. Greater distance from the state/industry caste composition of directors is also associated positively and significantly with an average firm's board homophily. The F statistics are also large, indicating that these excluded instruments are highly correlated with our endogenous regressor.

Second stage results are presented in Tables 10 for fine-grained caste homophily and in Tables

[30] Other results are available upon request.

5(A) and 7(A), columns (3) and (4), in the Online Appendix for religion and coarse-grained caste homophily, respectively. We find similar results as in the first IV strategy. Specifically, we see that fine-grained caste homophily is associated with significantly worse accounting and market based measures of firm performance and significantly greater firm risk. Greater coarse-grained caste homophily is also associated with worse firm outcomes. Similar results are observed for religion homophily and firm performance, although they are less precisely estimated.

We additionally perform several robustness checks to account for possibilities that may invalidate our instruments.

Industry Adjusted Regressions: In our main regressions, to ensure validity of our instruments we control for several industry characteristics that can influence both our instruments and firm performance. We conduct an alternative robustness check wherein we use industry-adjusted firm performance measures as our dependent variables. For this purpose, we take the difference of all firm performance variables in each year from the corresponding industry (two-digit NIC) medians in that year and regress these on the same regressors as in equation (4.1). Note that we drop industry-fixed effects in this specification since fixed effects would de-mean all variables by the industry mean, invariant across years.

Table13 presents the first and second stage results for the first IV approach, where the excluded instruments are the fine-grained caste homophily levels of the state and industry director supply pools.[31] Table13 Panel (B), shows that homophily of state and industry supply pools of directors are strongly and significantly positively associated with an average firm's fine-grained caste board homophily. First stage F-statistics are above 10 in all cases. Table13 Panel (A) presents the second stage results. The estimated coefficients on fine-grained caste homophily show that all performance measures are worse when boards are less diverse. For example, operating cash flows fall by 1.7 log points and Tobin's Q by 3 points when fine-grained caste homophily increases by one unit. Similar results are obtained for the second IV approach (Table14). We find that the excluded instruments are significantly positively associated with firm's fine-grained caste homophily. Sec-

[31] Results for coarse-grained caste homophily and religion are available upon request.

ond stage results for this approach are stronger and corroborate our finding that firm performance is negatively associated with board caste homophily.

Above and Below Median Firms: We include several control variables to ensure exclusion restriction of the Euclidean distance based IVs. It is still possible that poorly performing firms are unable to hire directors from beyond their already represented castes/religions, thereby affecting the distance between the boards' and supply pools' caste compositions. In that case, the exclusion restriction would not hold. To address this possibility, we split our sample into two sub-groups of high- and low-performing firms identified as those above and below the industry (NIC two digit) median for that performance measure in that year. Table15 presents second stage results for firms above and below the median for the second IV strategy. The table also shows the F statistics from the first stage regressions. We observe that the excluded IVs are highly correlated with board caste homophily in both sub-groups (above and below median). We also find that for both high- and low-performing firms, all measures of performance are worse when boards are less diverse. The magnitude of coefficients on board homophily are higher for high-performing firms. These findings demonstrate that our results hold even after accounting for the possibility that our excluded instruments may be correlated to both performance and diversity.

Caste Sub-Samples: We recognize the possibility that different social groups may have different levels of cohesion or sense of affinity. This has two implications for our measure of board diversity. First, the same HHI number may reflect different levels of mindset diversity since more cohesive caste or religious groups may have less mindset diversity than others. Second, greater cohesion within a caste group may result in less variation in caste homophily measures for boards with that as dominant caste. It is then possible that the results that we obtain may be driven by firms dominated by those castes that are less cohesive so as to have more variation in board homophily.

To address both possibilities, we divide our sample into eleven sub-samples of firms dominated by various coarse-grained castes (for Hindus) and religions (for non-Hindus). We present results for four coarse-grained castes and one non-Hindu religion (Jain). These groups have substantial numbers of firms in our data. The remaining groups do not have a large enough number of firms to

yield precise estimates or high first-stage F statistics in our second instrumental variable strategy. Tables 16 and 17 show the first and second stage results for this strategy for the five sub-samples. We observe that the excluded instruments are positively and significantly associated with board fine-grained caste homophily across all sub-samples and firm outcomes. For each sub-sample, we also observe that operating income, operating cash flows, profits and Tobin's Q are significantly negatively and volatility is significantly positively associated with fine-grained caste homophily. These results demonstrate that even though there may be heterogeneously cohesive caste and religious groups, our main results are not driven by any particular major group(s).

Approach 3: Clause 49

Next, we discuss results from the third identification strategy, which exploits firms' board composition changes resulting from SEC's Clause 49 requirements as explained in Section 4. Table11 presents the second stage results and the first stage F statistics. The regression discontinuity based set of excluded instruments is highly correlated with the board's fine-grained caste homophily, as evident from the high F statistics for all first stages corresponding to each second-stage dependent variable. Estimates of our main coefficients of interest show that accounting and market-based measures of firm performance are worse when boards are less diverse. Operating income, operating cash flow, and profits (all relative to sales) decline significantly by over 2 log points when homophily increases by one unit. For the same change in homophily, Tobin's Q falls and volatility increases significantly by 4 and 0.2 points, respectively.

Approach 4: Exposure to Diversity Outside the Board

Inspired by Adams and Ferreira (2009), we construct a variable that measures the fraction of a firm's directors that serve on the board of at least one other firm whose dominant fine-grained caste is different from the directors' own fine-grained caste. As explained in Section 4, we are able to measure this instrument for only 936 large publicly traded firms. Table12 (Panel B) presents the first stage results using this as an instrument. The table shows that the instrument is strongly

negatively correlated with firm's board homophily and that the first-stage F statistic is also large across all regressions corresponding to the various second stage dependent variables.

Table12 (Panel A) presents the second stage results for this IV approach. We observe that operating income and Tobin's Q are negatively associated with caste homophily of the board, although the coefficients are imprecisely estimated. Firm volatility significantly increases by 0.047 points with a one unit increase in board caste homophily. However, firm operating cash flow and profits are positively associated with board caste homophily, with the former significantly so.

On the basis of our detailed and rigorous empirical analyses, we conclude that lack of diversity on firm boards is negatively associated with key firm outcomes. An understanding of how important diversity may be for firms, however, depends on the granularity with which diversity is measured. Our results show that the strongest negative associations with performance are witnessed when diversity is measured along our narrowest measure, fine-grained caste, followed by coarse-grained caste, in turn followed by religion.

6 Mechanisms

Diversity among directors on a board affects boardroom group dynamics which, in turn, affect firm performance. On one hand, homogeneous directors may get along better and have fewer interpersonal conflicts (O'Reilly et al. (1993), Smith et al. (1994)). This may help boards in their decision making, benefitting firm performance. On the other hand, homogenous boards may be characterized by cronyism, worsening board decisions and firm performance. Additionally, they may also have less access to novel information and may not monitor the management well, which can also worsen their decisions and firm outcomes. The relative strengths of these channels determine the net effect of board diversity on firm performance. In this section, we present evidence that can be viewed as consistent with both cronyism and fewer interpersonal conflicts. Data limitations prevent us from investigating access to information and monitoring by directors.[32] Nonetheless,

[32] As a board's diversity increases, the novelty of information available to the board increases because directors that are drawn from diverse backgrounds have access to non-redundant information. Novel information enables experi-

our results in the previous section demonstrate that the net association of board homophily with performance is negative for firms, indicating that cronyism, along with lack of superior monitoring and absence to novel information are stronger channels.

Figure 7 shows that boards with high homophily (low diversity) have higher meeting attendance rates. In Figure 8(a), we further show that directors that share their fine-grained caste with the dominant fine-grained caste of the board have greater attendance than the directors that do not. Figure 8(b) indicates that the difference in attendance of the dominant and non-dominant fine-grained caste directors increases with board homophily. These patterns suggest that directors belonging to dominant castes are more willing to attend meetings. These findings are consistent with cronyism as well as directors getting along better in less diverse boards.

We also find that boards where at least one director has resigned have lower homophily (i.e., are more diverse) than the ones with no director resignations (Figure 9(a)). Additionally, less diverse boards see a smaller fraction of directors resigning (Figure 9(b)). Both findings again suggest that more homogeneous boards are either more prone to cronyism or have less conflict.

Next, we assess if directors of dominant castes on boards serve disproportionately more on important sub-committees of directors, as indicative of cronyism or in-group favoritism. Table18 show some evidence of this. The first row shows the percentage share of directors that belong to the most, second-most, and third-most dominant caste on the board, averaged over all firms for the sample period. The subsequent rows depict the ratios of the share of committee chair positions or memberships held by directors belonging to the dominant fine-grained caste relative

mentation and complex problem solving. To test whether diverse boards encourage management to undertake riskier projects or engender more innovative outcomes, studies have used measures such as the number of patents. We do not have similar information in our data. An R&D variable is sparsely populated; of the firms in our sample, 92% do not have information on R&D, rendering any meaningful analysis impossible. More diverse boards may also be better able to monitor the management, reducing fraud and improving performance. The literature typically measures the monitoring function of the board by looking at CEO turnover and compensation sensitivity to firm performance. In our data, of over 20,000 firms and CEOs, only 181 CEOs resign, of which only 8 resign after poor firm performance (measured as firm sales below average of the industry in the past one or three years). We do not have compensation data for the majority of CEOs in our sample.

to the share of the entire board constituted by that fine-grained caste. We observe that this ratio is 0.99 for the chairs of all committees belonging to the most dominant fine-grained caste. Being less than 1, this does not indicate a disproportionality. However, the fraction is indeed greater than 1 for the second- and third-most dominant fine-grained castes on the board, indicating that they are over-represented in chair positions. We find a similar over-representation when we consider memberships in a few important committees in particular – audit, remuneration, and shareholder grievance. Similar results for coarse-grained caste are available in Table 8 of the Online appendix.

Finally, as further evidence of cronyism, we observe that firms with less diverse boards are more likely to have the CEO belong to the dominant fine-grained caste in the board. This is evident in Figure 6 (Figure 14 of the Online Appendix) which shows that throughout the sample period, firms which have the CEO belonging to the dominant fine-grained caste (coarse-grained caste) in the rest of the board have substantially higher average caste homophily than others.

Taken together, these results paint a picture wherein less diverse boards are prone to cronyism or get along better with each other. However, results in Section 5 indicate that the potential beneficial effects of fewer boardroom conflicts on firm performance are more than offset by the negative effects of cronyism (as well as diminished information access and monitoring).

7 Conclusion

We build a unique dataset that allows us to map Indian last names to religion and caste. Since caste is deeply rooted in Indian society and influences myriad outcomes from marriages to intergenerational educational mobility, we argue that Indians' caste identities shape their mindsets. As such, we use caste diversity among directors in a firm's board as a measure of their mindset diversity. We show that Indian corporate boards persistently and systematically lack in diversity. Our results demonstrate that this lack of diversity has a negative association with key measures of firm value and performance. We find the strongest negative results for the fine-grained measure of caste, indicating the ground reality that people feel stronger affinity with those who share their caste identity

along this dimension relative to the coarse-grained dimension or religion.

Besides developing a potentially highly useful dataset for future studies and presenting novel findings about corporate governance of Indian firms, our paper has two key takeaways. First, while much research, corporate governance laws, and recommendations emphasize gender diversity in boards, our results demonstrate that socio-cultural aspects other than gender influence firm outcomes. Recent proposals that urge firms to hire diverse directors along dimensions such as race and sexual orientation (NASDAQ, 2020), and ethnicity and backgrounds (UK's Financial Reporting Council, 2018), are, therefore, welcome developments. Second, the granularity of traits along which directors differ from each other matters for group dynamics in the boardroom, and through them, firm performance. Policy recommendations about boardroom diversity should consider the dimension of identity along which individuals feel the most affinity towards others.

References

Adams, R. B. (2016). Women on boards: The superheroes of tomorrow?. *The Leadership Quarterly*, 27(3):371–386.

Adams, R. B. and Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of financial economics*, 94(2):291–309.

Ahern, K. R. and Dittmar, A. K. (2012). The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation. *The Quarterly Journal of Economics*, 127(1):137–197.

Arnaboldi, F., Casu, B., Kalotychou, E., and Sarkisyan, A. (2018). The performance effects of board heterogeneity: what works for eu banks? *The European Journal of Finance*, pages 1–28.

Åslund, O., Hensvik, L., and Skans, O. N. (2014). Seeking similarity: How immigrants and natives manage in the labor market. *Journal of Labor Economics*, 32(3):405–441.

Bengtsson, O. and Hsu, D. H. (2015). Ethnic matching in the us venture capital market. *Journal of Business Venturing*, 30(2):338–354.

Bernile, G., Bhagwat, V., and Yonker, S. (2018). Board diversity, firm risk, and corporate policies. *Journal of Financial Economics*, 127(3):588–612.

Bertrand, M., Black, S. E., Jensen, S., and Lleras-Muney, A. (2018). Breaking the glass ceiling? the effect of board quotas on female labour market outcomes in norway. *The Review of Economic Studies*, 86(1):191–239.

Beteille, A. (1996). Varna and jati. *Sociological Bulletin*, 45(1):15–27.

Chen, G., Chittoor, R., and Vissa, B. (2015). Modernizing without westernizing: Social structure and economic action in the indian financial sector. *Academy of Management Journal*, 58(2):511–537.

Claes, K. and Vissa., B. ((2020)). "does social similarity pay off? homophily and venture capitalists' deal valuation, downside risk protection, and financial returns in india.". *Organization Science*.

Coles, Jeffrey L., N. D. D. and Naveen., L. (2008). Boards: Does one size fit all?

Currarini, S., Jackson, M. O., and Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045.

Damaraju, N. L. and Makhija, A. K. (2018). The Role of Social Proximity in Professional CEO Appointments: Evidence from Caste/Religion-Based Hiring of CEOs in India. *Strategic Management Journal*, 39(7):2051–2074.

Damodaran, H. (2008). *India's New Capitalists: Caste, Business, and Industry in a Modern Nation*. Springer.

Dass, N., Kini, O., Nanda, V., Onal, B., and Wang, J. (2013). Board expertise: Do directors from related industries help bridge the information gap? *The Review of Financial Studies*, 27(5):1533–1592.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, 95(2):158–165.

Dumont, L. (1980). *Homo Hierarchicus: The Caste System and Its Implications*. University of Chicago Press.

Fairlie, R. W., Hoffmann, F., and Oreopoulos, P. (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review*, 104(8):2567–91.

Fisman, R., Paravisini, D., and Vig, V. (2017). Cultural Proximity and Loan Outcomes. *American Economic Review*, 107(2):457–92.

Freeman, R. B. and Huang, W. (2015). Collaborating with People like Me: Ethnic Coauthorship Within the United States. *Journal of Labor Economics*, 33(S1):S289–S318.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The Elements of Statistical Learning*, volume 1. Springer series in statistics New York.

Ghani, E., Kerr, W. R., and O'connell, S. (2014). Spatial Determinants of Entrepreneurship in India. *Regional Studies*, 48(6):1071–1089.

Giannetti, M. and Zhao, M. (2019). Board ancestral diversity and firm performance volatility. *Journal of Financial and Quantitative Analysis*, 54(3):1117–1155.

Giuliano, L., Levine, D. I., and Leonard, J. (2009). Manager race and the race of new hires. *Journal of Labor Economics*, 27(4):589–631.

Giuliano, L. and Ransom, M. R. (2013). Manager ethnicity and employment segregation. *ILR Review*, 66(2):346–379.

Glover, D., Pallais, A., and Pariente, W. (2017). Discrimination as a self-fulfilling prophecy: Evidence from french grocery stores. *The Quarterly Journal of Economics*, 132(3):1219–1260.

Gompers, P. A., Mukharlyamov, V., and Xuan, Y. (2016). The cost of friendship. *Journal of Financial Economics*, 119(3):626–644.

Hegde, D. and Tumlinson, J. (2013). Can Birds of a Feather Fly Together? the Payoffs of Ethnic Proximity in Us Venture Capital. *Unpublished working paper. New York University and University of Munich*.

Helmers, C., Patnam, M., and Rau, R. (2017). Do board interlocks increase innovation? evidence from a corporate governance reform in india. *Journal of Banking and Finance*, 80:51–70.

Hnatkovska, V., Lahiri, A., and Paul, S. (2012). Castes and Labor Mobility. *American Economic Journal: Applied Economics*, 4(2):274–307.

Hnatkovska, V., Lahiri, A., and Paul, S. B. (2013). Breaking the Caste Barrier Intergenerational Mobility in India. *Journal of Human Resources*, 48(2):435–473.

Iyer, L., Khanna, T., and Varshney, A. (2013). Caste and entrepreneurship in india. *Economic and Political Weekly*, pages 52–60.

Jodhka, S. S. (2010). Dalits in business: Self-employed scheduled castes in north-west india. *Economic and Political Weekly*, pages 41–48.

Joshi, S., Kochhar, N., and Rao, V. (2018). Jati inequality in rural bihar. *The World Bank*.

Kim, D. and Starks, L. T. (2016). Gender diversity on corporate boards: Do women contribute unique skills? *American Economic Review*, 106(5):267–71.

Knyazeva, A., Knyazeva, D., and Masulis, R. W. (2013). The supply of corporate directors and board independence. *The Review of Financial Studies*, 26(6):1561–1605.

Kramarz, F. and Thesmar, D. (2013). Social networks in the boardroom. *Journal of the European Economic Association*, 11(4):780–807.

Kumar, H., Somanathan, R., et al. (2017). Caste connections and government transfers: The mahadalits of bihar. *Working Paper, Centre for Development Economics, Delhi School of Economcis*.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Munshi, K. (2019). "caste and the indian economy.". *Journal of Economic Literature*, 57(4):781–834.

Munshi, K. and Rosenzweig, M. (2013). Networks, commitment, and competence: Caste in indian local politics. Technical report, National Bureau of Economic Research.

Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1):46–98.

O'Reilly, C., Snyder, R., and Boothe, J. (1993). Effects of executive team demography on organizational change. *Organizational change and redesign*, pages 147–175.

Page, S. E. (2007). "making the difference: Applying a logic of diversity.". *Academy of Management Perspectives*, 21(4):6–20.

Petersen, T., Saporta, I., and Seidel, M.-D. L. (2000). Offering a job: Meritocracy and social networks. *American Journal of Sociology*, 106(3):763–816.

Phillips, K. W. and O'Reilly, C. A. (1998). Demography and diversity in organizations: A review of 40 years of research. *Research in organizational behavior*, 20(77-140).

Shayo, M. and Zussman, A. (2011). Judicial ingroup bias in the shadow of terrorism. *The Quarterly Journal of Economics*, 126(3):1447–1484.

Sila, V., Gonzalez, A., and Hagendorff, J. (2016). Women on board: Does boardroom gender diversity affect firm risk? *Journal of Corporate Finance*, 36:26–53.

Smith, K. G., Smith, K. A., Olian, J. D., Sims Jr, H. P., O'Bannon, D. P., and Scully, J. A. (1994). Top management team demography and process: The role of social integration and communication. *Administrative science quarterly*, pages 412–438.

Srinivas, M. N. (1995). *Social Change in Modern India*. Orient Blackswan.

Terjesen, S., Sealy, R., and Singh, V. (2009). Women directors on corporate boards: A review and research agenda. *Corporate governance: an international review*, 17(3):320–337.

Thorat, S. and Neuman, K. S. (2012). *Blocked by Caste: Economic Discrimination in Modern India*. Oxford University Press.

Varshney, A., Chatterjee, P., and Katznelson, I. (2012). Two Banks of the Same River? Social Order and Entrepreneurialism in India. *Anxieties of Democracy: Tocquevillean Reflections on India and the United States*, pages 225–56.

Appendices

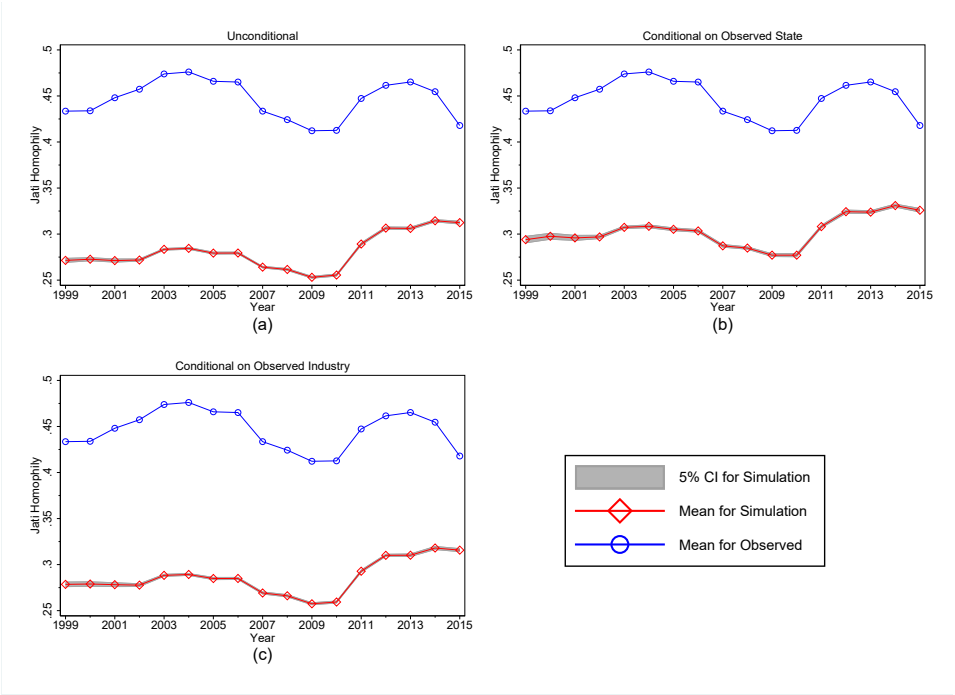


Figure 1: Observed vs. Simulated Average Fine-Grained Caste Homophily[a]

[a]Source: Prowess, matrimonial data. The three graphs in the figure present the mean fine-grained caste homophily across firms each year in the observed and simulated samples for three distinct simulation criteria: unconditional, conditional on firm's state and on firm's industry. Details about the simulation methods are provided in Section 3.1.

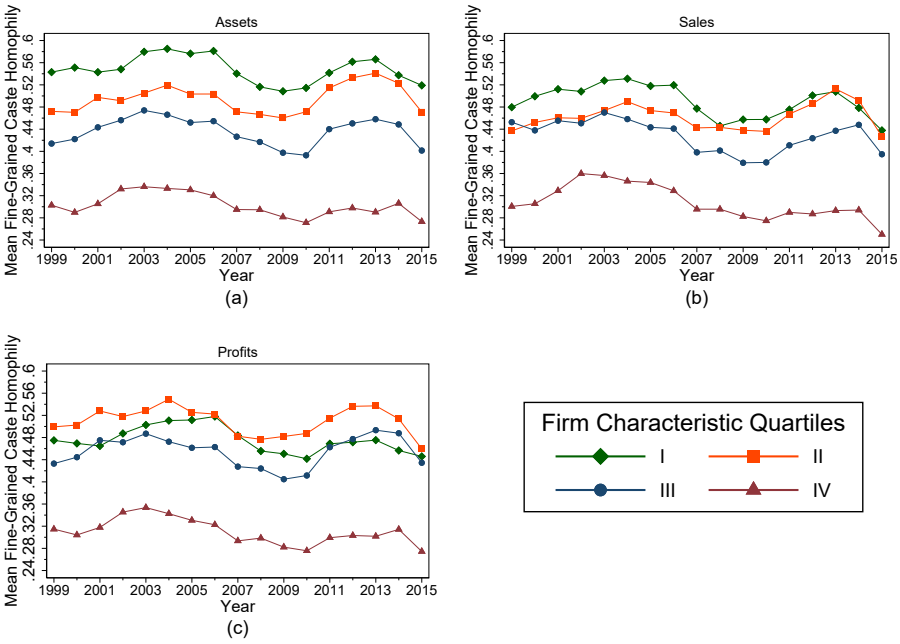


Figure 2: Average Fine-Grained Caste Homophily by Firm Size[a]

[a]Source: Prowess, matrimonial data. Figures (a), (b), and (c) show the average fine-grained caste homophily levels in firms falling in the four quartiles of assets, sales, and profits, respectively.

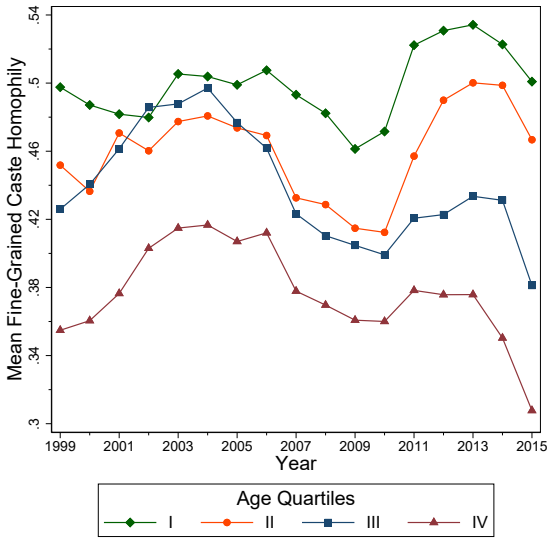
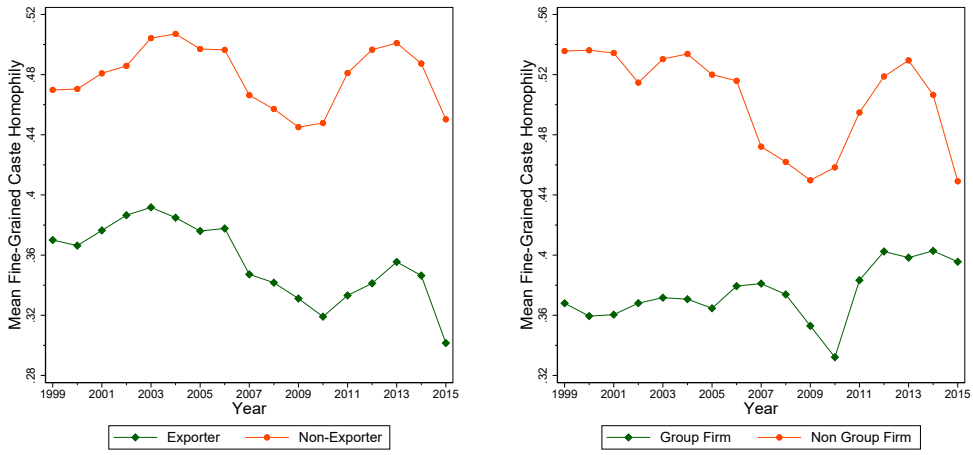


Figure 3: Average Fine-Grained Caste Homophily by Firm Age[a]

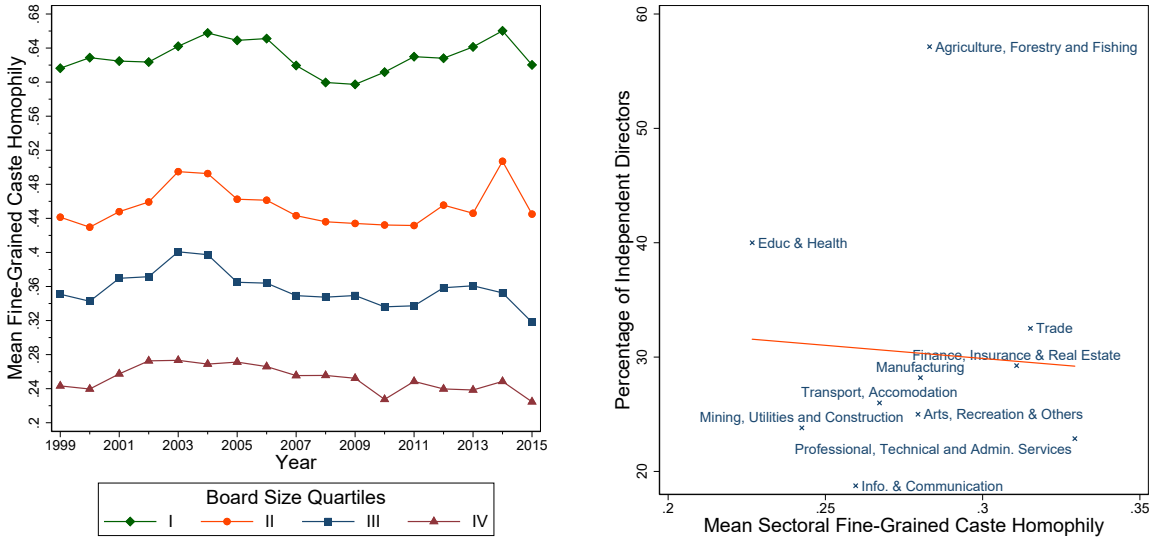
[a]Source: Prowess, matrimonial data. The figure shows the average fine-grained caste homophily levels in firms falling in the four quartiles of age.



(a) Exporting Status (b) Business Group

Figure 4: Fine-Grained Caste Homophily by Exporting Status and Business Group[a]

[a]Notes: Source: Prowess, matrimonial data. In panel (a), homophily is averaged over all exporting and non-exporting firms separately. Panel b does the same for all firms that belong to a business group and the ones that do not.



(a) Board Size (b) Pct. Independent Directors

Figure 5: Average Fine-Grained Caste Homophily by Board Size and % Independent Directors[a]

[a]Notes: Source: Prowess, matrimonial data. In panel a, homophily is averaged over all firms in each quartile of the board size distribution. Panel b plots the average sectoral fine-grained caste homophily against the average sectoral percentage of independent directors in firms' boards in the year 2015.

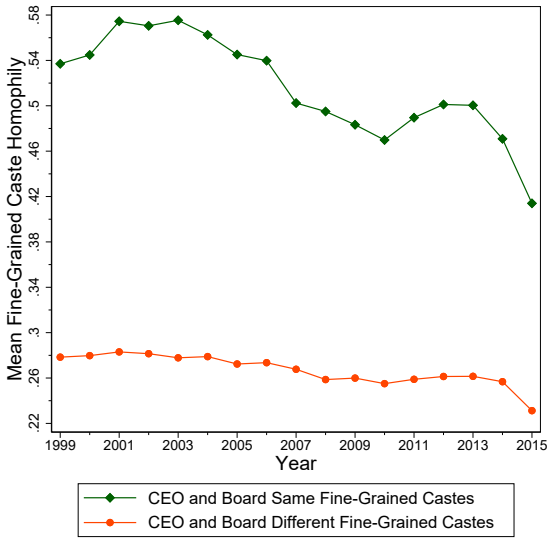


Figure 6: Caste Homophily for Firms with & without CEO of Board's Dominant Fine-Grained Caste[a]

[a]Notes: Source: Prowess, matrimonial data. Homophily is averaged over all firms in a year that fall into either of two groups: those that have their CEO belonging to the same fine-grained caste as the dominant caste in the rest of the board, and those where the CEO belongs to a different fine-grained caste.

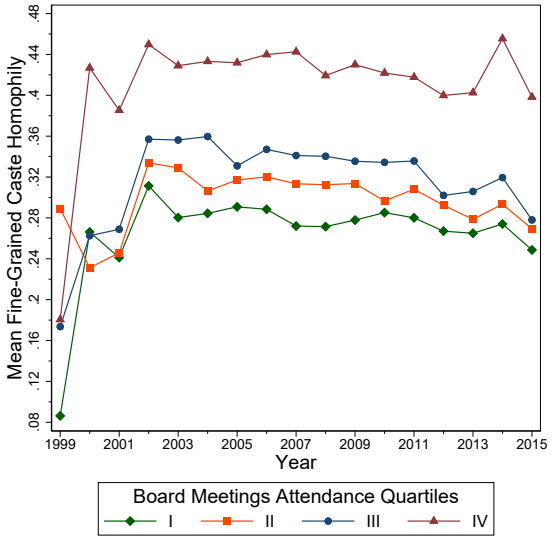
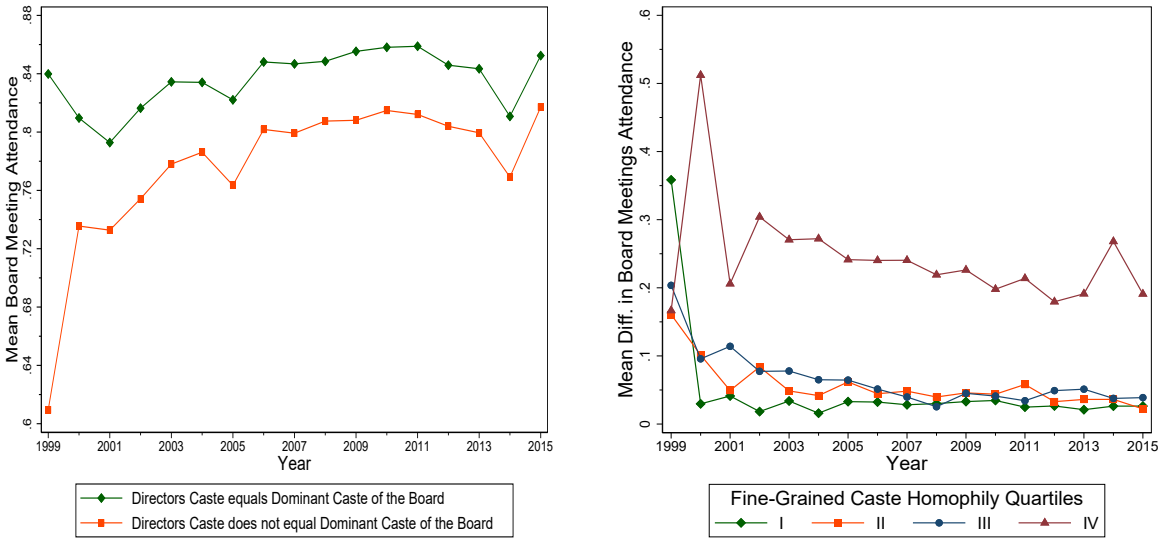


Figure 7: Fine-Grained Caste Homophily and Attendance of Board Meetings[a]

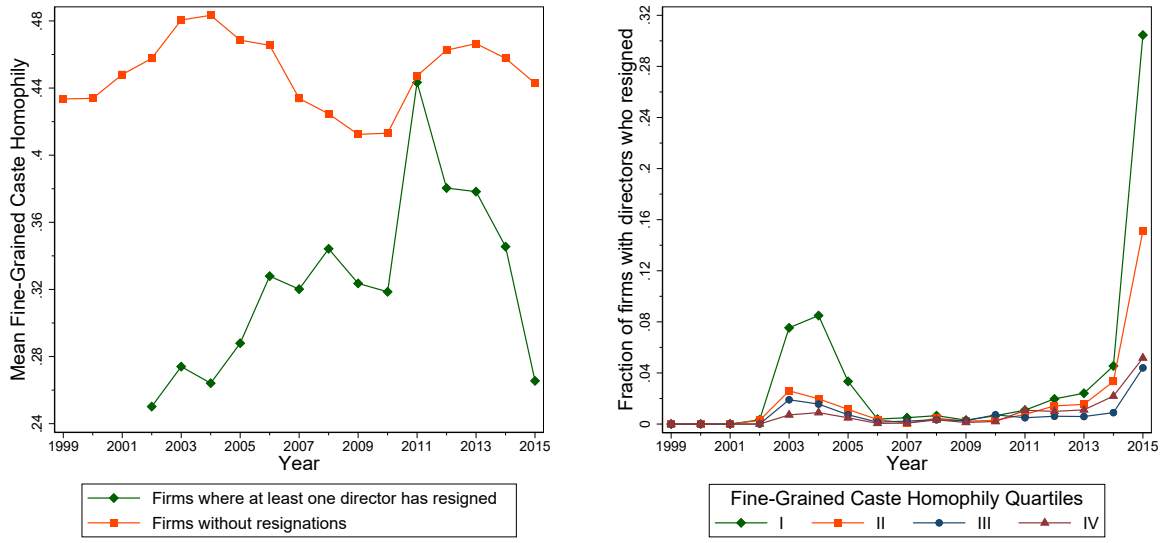
[a]Notes: Source: Prowess, authors' last name to caste mapping using matrimonial data. Homophily is averaged over all firms in each quartile of the attendance of board meetings distribution. Attendance of board meetings in a year is calculated as the average attendance of board members across all board meetings of a board in a year.



(a) (b)

Figure 8: Board Meeting Attendance by Dominant Status of Director Caste[a]

[a]Notes: Source: Prowess, authors' last name to fine-grained caste mapping using matrimonial data. In Figure (a) mean board meeting attendance is averaged over all directors with the same fine-grained caste as the dominant caste in the board and ones without. In Figure (b), difference in board meetings attendance rates between directors with the same fine-grained caste as the dominant caste in the board and ones the ones without is averaged over all firms in each quartile of the fine-grained caste homophily of the board.



(a) (b)

Figure 9: Fine-Grained Caste Homophily and Resignation of Board Members [a]

[a]Notes: Source: Prowess, authors' last name to fine-grained caste mapping using matrimonial data. In Figure (a) fine-grained caste homophily is averaged over all firms where at least one director resigned and where no director resigned.

Variable Definitions

Table 1: Variable definitions

Variables	Definitions
Panel A: Firm Variables	
Age of firm	Number of years since incorporation of firm
Export status	Indicator variable: one for exporting firms, zero otherwise
State of registration	The Indian state in which the firm is registered
Industry	Two or three digit NIC-2008 sector
Listing status	Indicator variable: one for firms listed either in the Bombay Stock Exchange (BSE) or the National Stock Exchange (NSE) at that point in time, zero otherwise
Assets	Book value of total assets in rupees million deflated by the all-India CPI (2001=100)

Sales	Total value of sales in rupees million deflated by the all-India CPI (2001=100)
Profits	Total value of profits in rupees million deflated by the all-India CPI (2001=100)
Operating cash flow	Cash flow from operating activities before depreciation
Leverage	Book value of debt over book value of total assets
Operating income	Sales less operating expenses
Tangibility	Net Property plant equipment over book value of total assets
Tobin's Q	Sum of book value of debt, book value of preferred stock and market value of common stock over book value of assets. The market value of common stock is measured in two ways - a. the latest market value available on or before the reporting date (Latest) b. the mean market value over the entire reporting period (Full Period)
Volatility	The standard deviation of stock returns of a firm in the entire reporting period
Net Worth	The net worth of a company is what it owes its equity share holders. This consists of the monies put into the company by the equity share holders in the form of equity capital and the profits generated and retained as reserves by the company.

Panel B: Board Characteristics

Board size	Number of directors in the board
Frequency of board meetings	Number of board meetings per year
Board meeting attendance	Mean number of board meetings attended by all members of a board over total number of board meetings
Resignation	Indicator variable: one if a board member resigns, zero otherwise
Board turnover 1Y	% of directors in a board who were not present in the previous year
Board turnover 3Y	% of directors in a board who were not present in the board three years prior to the current year
CEO duality	Indicator variable: one if the at least one CEO of a firm is also the chair.

Panel C: Measures of Homophily

Dominant coarse- and fine-grained caste and religion of a board	The coarse- and fine-grained caste and religion of the maximum number of directors of a board. In case of ties, the dominant coarse- and fine-grained caste and religion is chosen randomly from the tie.

Board coarse- and fine-grained caste and religion homophily	Coarse- and fine-grained caste and religion HHI, i.e., the sum of squared shares of all fine-grained castes (coarse-grained castes, religions) represented on the board.
Sector coarse- and fine-grained caste and religion homophily	Coarse- and fine-grained caste and religion HHI, i.e., the sum of squared shares of all fine-grained castes (coarse-grained castes, religions) of directors in an industry. The baseline approach considers each name as a distinct directorship, even if the name is same. The alternative approach considers all occurrences of the same name as one unique director.
State coarse- and fine-grained caste and religion homophily	Coarse- and fine-grained caste and religion HHI, i.e., the sum of squared shares of all fine-grained castes (coarse-grained castes, religions) of directors in a state.The baseline approach considers each name as a distinct directorship, even if the name is same. The alternative approach considers all occurrences of the same name as one unique director.
Sector coarse- and fine-grained caste and religion Euclidean distance of a board	Distance between the vector representing the coarse- and fine-grained caste and religion composition of directors in the industry and the corresponding vector for the firm board. The baseline approach considers each name as a distinct directorship, even if the name is same. The alternative approach considers all occurrences of the same name as one unique director.
State coarse- and fine-grained caste and religion Euclidean distance of a board	Distance between the vector representing the coarse- and fine-Grained caste and religion composition of directors in the state and the corresponding vector for the firm board. The baseline approach considers each name as a distinct directorship, even if the name is same. The alternative approach considers all occurrences of the same name as one unique director.

Table 2: Summary Statistics

Panel A: Sectoral Distribution of Firms (Percentages)		
	1999	2015
Agriculture, Forestry and Fishing	2.46	1.50
Mining, Utilities and Construction	6.47	12.02
Manufacturing	41.42	23.32
Trade	11.74	12.86
Transport, Accommodation	3.41	4.66
Information & Communication	4.01	4.90
Finance, Insurance & Real Estate	23.02	26.25
Professional, Technical and Admin. Services	3.86	6.95
Education & Health	0.50	1.29
Arts, Recreation & Others	0.85	5.23
Diversified	2.26	1.02
Panel B: Firm Characteristics		
	1999	2015
Total Firms	1994	5597
% Listed	47.94	40.16
% Exporters	35.91	21.65
% Group Firms	57.37	48.92
Mean Age	20.22	19.73
Mean Assets (Rupees Millions)	17.99	23.33
Mean Profits (Rupees Millions)	2.16	2.41
Mean Sales (Rupees Millions)	13.34	15.20
Mean Net Tangible Asset Intensity	0.32	0.25
Mean Leverage	0.45	0.45
Mean Return on Assets	0.05	0.04
Mean Asset Turnover	0.97	0.83
Mean Tobin's Q	1.02	1.72
Mean Market to Book Ratio	1.41	2.77
Mean Risk	0.11	0.04
Panel C: Board Characteristics		
	1999	2015
% with CEO Duality	15.05	15.15
Mean % Independent Directors	20.26	2.22
Mean Board Size	5.44	5.44

Source: Prowess. This table provides basic summary statistics for firms in our sample that have at least two directors and for which we can assign a caste identity for all directors on the boards.

Table 3: Within-Industry Board Interlocks

NIC	Within-Industry Board Interlocks			
	Mean	Minimum	Maximum	% Firms with interlocks
Agriculture, Forestry and Fishing	0.05	0.00	0.20	39.29
Mining, Utilities and Construction	0.10	0.00	0.83	44.53
Manufacturing	0.31	0.00	1.00	76.60
Trade	0.09	0.00	1.00	27.08
Transport, Accomodation	0.17	0.00	0.80	63.16
Information & Communication	0.13	0.00	1.00	46.67
Finance, Insurance & Real Estate	0.20	0.00	1.00	66.07
Professional, Technical and Admin. Services	0.00	0.00	0.11	4.17
Education & Health	0.03	0.00	0.31	11.76

Source: Indian Boards database. This table presents proportions of directors of firms that also serve on at least one other firm's board, currently (2015) or in the past (2012-2014), with that firm belonging to the same one-digit industry.

Table 4: Distance Between Firm and Industry/State Director Fine-Grained Caste Composition

Year	10th Percentile	50th Percentile	90th Percentile	Mean	Standard Deviation
Panel A: Distance Between Firm and Industry Director Composition					
1999	0.32	0.32	0.87	0.54	0.19
2004	0.34	0.34	0.88	0.58	0.2
2009	0.32	0.32	0.82	0.54	0.18
2015	0.31	0.31	0.83	0.54	0.19
Panel B: Distance Between Firm and State Director Composition					
1999	0.3	0.3	0.84	0.52	0.19
2004	0.32	0.32	0.88	0.55	0.2
2009	0.3	0.3	0.77	0.5	0.18
2015	0.29	0.29	0.77	0.52	0.19

Source: Matrimonial data, Prowess. This table shows moments for the Euclidean distances between the fine-grained caste composition of firms' directors and that of the set of directors in the same two-digit industry (Panel A) or state (Panel B).

Table 5: Religion, Coarse- and Fine-Grained Caste Mapping for Names of Directors

Number of classifications	% last names with up to that number of classifications	Average cumulative probability associated with classifications
	Religion	
1	68.6	95.3
2	92.2	99.5
3	98.0	99.9
4	99.6	100
5	100	100
	Coarse-Grained Caste	
1	45.1	78.5
2	58.3	90.6
3	68.5	95.8
4	76.9	98.1
5	85.1	99.3
	Fine-Grained Caste	
1	39.7	67.0
2	49.8	77.8
3	56.5	83.2
4	61.5	86.5
5	65.5	88.8

Source: Matrimonial and Prowess data. This table provides the religion, coarse- and fine-grained caste distribution of director last names. For space considerations, the table only shows the distribution for last names associated with up to five fine-grained castes.

Table 6: Religion, Coarse- and Fine-Grained Caste Composition of Matrimonial Data

Religion	% Last Names	Coarse-Grained Caste	% Last Names	Fine-Grained Caste	% Last Names
Hindu	80.69	Brahmin	18.67	Maratha	4.10
Muslim	8.00	Kshatriya	12.67	Brahmin Iyer	3.88
Christian	6.42	Vaishya	12.83	Brahmin	3.20
Jain	2.30	Shudra	32.05	Sindhi	3.02
Sikh	1.32	Dalit	1.78	Nair	2.84
Parsi	1.15	Unknown Varna	1.35	Arya Vysya	2.46
Buddhist	0.04			Agarwal	2.30
Jewish	0.01			Khatri	1.97
NA	0.08			Vannia Kula Kshatriyar	1.93
				Brahmin Deshastha	1.89
				Ezhava	1.89

Source: Matrimonial data. This table provides the religion, coarse- and fine-grained caste distribution of the last names included in the final mappings developed by the authors, as described in Section 2.1. For space considerations, the table only shows the distribution for the top ten most frequently occurring fine-grained castes out of a total of 471 distinct fine-grained castes that we can identify.

Table 7: Fine Grained Homophily and Firm Outcomes: Fixed Effects

Variables	(1) Log(Operating Income/Sales)	(2) Log(Operating Cash Flow/Sales)	(3) Log(Profits/Sales)	(4) Tobin's Q	(5) Volatility
Fine-Grained Caste Homophily	-0.263***	-0.296***	-0.225**	-0.186***	0.008***
	(0.039)	(0.087)	(0.086)	(0.046)	(0.002)
Board Size	-0.030**	-0.061	0.041	0.047**	-0.003***
	(0.014)	(0.037)	(0.036)	(0.019)	(0.001)
Board Size Squared	0.002***	0.002	-0.002	-0.001	0.000***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)
Industry HHI	0.029	0.220	0.387*	0.059	0.001
	(0.059)	(0.150)	(0.208)	(0.071)	(0.002)
I(Listed)	-0.028	0.070	-0.014	-1.390**	-0.004
	(0.025)	(0.068)	(0.074)	(0.686)	(0.003)
I(Group Firm)	0.051**	0.105***	0.155***	0.180***	-0.005***
	(0.025)	(0.040)	(0.049)	(0.034)	(0.001)
Proportion of Group Firms in Industry	0.354	0.398**	0.368*	-0.026	-0.006*
	(0.219)	(0.178)	(0.214)	(0.103)	(0.004)
Observations	46,808	34,009	57,570	27,163	28,809
R-squared	0.116	0.084	0.038	0.379	0.465

Notes: This table presents fixed effects results for firm performance variables – operating income, operating cash flows, profits, (all relative to sales), Tobin's Q, and volatility – regressed on board fine-grained caste homophily and other control variables, and year and two-digit industry fixed effects. Control variables include age, board size, board size square, leverage, real assets, tangibility, listing status, export status, HHI index measured at NIC3 level, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in the industry, and proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). Robust standard errors, clustered by industry and year, are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 8: Instrumental Variable Regression, First Stage: Fine-Grained Caste Homophily

	(1)	(2)	(3)	(4)
	State and Industry Homophily		State and Industry Homophily and Euclidean Distance	
Second Stage Dependent Variable	Log(Profits/Sales)	Volatility	Log(Profits/Sales)	Volatility
State Fine-Grained Caste Homophily	0.468***	0.447***	0.474***	0.429***
	(0.102)	(0.125)	(0.071)	(0.073))
Industry Fine-Grained Caste Homophily	0.476***	0.275**	1.173***	0.816***
	(0.136)	(0.138)	(0.103)	(0.170)
Distance from State Director Fine-Grained Caste Composition			0.307***	0.286***
			(0.059)	(0.046)
Distance from Industry Director Fine-Grained Caste Composition			1.003***	0.880***
			(0.062)	(0.050)
Observations	57,504	28,809	57,504	28,809
First Stage F-statistic	17.60	10.65	7787.37	1840.56

Notes: This table presents first stage IV results for firm performance variables (operating income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility) regressed on board fine-grained caste homophily and other control variables, and year and two-digit industry fixed effects. The excluded instruments in columns 1 and 2 are the fine-grained caste homophily levels of the state and two digit industry director supply pools for a firm. The excluded instruments in columns 3 and 4 are the fine-grained caste homophily levels of the two director supply pools and the distance of fine-grained caste composition of the firm's board from that of the two supply pools. Control variables include age, board size, board size square, leverage, real assets, tangibility, listing status, export status, three digit industry HHI, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, and the proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). The first stage F statistic is the Kleibergen-Paap rk Wald F statistic. Robust standard errors, clustered by industry and year, are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 9: Fine-Grained Caste Homophily and Firm Outcomes: Second Stage, IV Approach 1

	(1)	(2)	(3)	(4)	(5)
Variables	Log (Operating Income/Sales)	Log (Operating Cash Flow/Sales)	Log (Profits/Sales)	Tobin's Q	Volatility
Fine-Grained Caste Homophily	-2.311**	-1.618**	-2.731**	-2.320***	0.007
	(0.998)	(0.979)	(1.049)	(0.891)	(0.027)
Board Size	-0.228**	-0.174*	-0.202*	-0.082	-0.003
	(0.103)	(0.103)	(0.119)	(0.055)	(0.002)
Board Size Squared	0.010**	0.007	0.008	0.003	0.000
	(0.004)	(0.004)	(0.005)	(0.002)	(0.000)
Industry HHI	-0.008	0.189	0.331	0.028	0.001
	(0.064)	(0.156)	(0.207)	(0.071)	(0.002)
I(Listed)	-0.136**	0.024	-0.146*	-1.393*	-0.004
	(0.061)	(0.067)	(0.079)	(0.791)	(0.003)
I(Group Firm)	-0.097	0.026	-0.027	0.102*	-0.005***
	(0.066)	(0.063)	(0.080)	(0.058)	(0.002)
Proportion of Group Firms in Industry	0.267	0.358**	0.253	-0.107	-0.006*
	(0.221)	(0.176)	(0.223)	(0.104)	(0.004)
Observations	46,753	33,987	57,504	29,204	28,809

Notes: This table presents second stage IV results for firm performance variables (operating income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility) regressed on board fine-grained caste homophily and other control variables, and year and two-digit industry fixed effects. The excluded instruments are the fine-grained caste homophily levels of the state and two digit industry director supply pools for a firm. Control variables include age, board size, board size square, leverage, real assets, tangibility, listing status, export status, three digit industry HHI, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, and the proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). Robust standard errors, clustered by industry and year, are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 10: Fine-Grained Caste Homophily and Firm Outcomes: Second Stage, IV Approach 2

	(1)	(2)	(3)	(4)	(5)
Variables	Log (Operating Income/Sales)	Log (Operating Cash Flow/Sales)	Log (Profits/Sales)	Tobin's Q	Volatility
Fine-Grained Caste Homophily	-0.242***	-0.323***	-0.214***	-0.182***	0.008***
	(0.043)	(0.098)	(0.079)	(0.045)	(0.002)
Board Size	-0.028*	-0.063*	0.042	0.048**	-0.003***
	(0.015)	(0.037)	(0.035)	(0.020)	(0.001)
Board Size Squared	0.002**	0.002	-0.002	-0.001	0.000***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)
Industry HHI	0.032	0.219	0.391*	0.059	0.001
	(0.058)	(0.151)	(0.208)	(0.071)	(0.002)
I(Listed)	-0.027	0.070	-0.013	-1.390**	-0.004
	(0.025)	(0.068)	(0.075)	(0.686)	(0.003)
I(Group Firm)	0.053**	0.103***	0.156***	0.180***	-0.005***
	(0.025)	(0.039)	(0.049)	(0.035)	(0.001)
Proportion of Group Firms in Industry	0.346	0.397**	0. 365*	-0.026	-0.006*
	(0.219)	(0.177)	(0.214)	(0.103)	(0.004)
Observations	46,753	33,987	57,504	27,163	28,809

Notes: This table presents second stage IV results for firm performance variables (operating income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility) regressed on board fine-grained caste homophily and other control variables, and year and two-digit industry fixed effects. The excluded instruments are the fine-grained caste homophily levels of the two director supply pools and the distance of fine-grained caste composition of the firm's board from that of the two supply pools. Control variables include age, board size, board size square, leverage, real assets, tangibility, listing status, export status, three digit industry HHI, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, and the proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). Robust standard errors, clustered by industry and year, are in parentheses. *** p<0.01, ** p<0.05, * p<0.10..

Table 11: Instrumental Variable Regressions, First and Second Stage: Fine-Grained Caste Homophily, IV Approach 3

	(A) Second Stage Results				
	(1)	(2)	(3)	(4)	(5)
Variables	Log(Operating Income/Sales)	Log(Operating Cash Flow/Sales)	Log(Profits/Sales)	Tobin's Q	Volatility
Fine-Grained Homophily	-2.912***	-2.675***	-2.877***	-4.079***	0.180***
	(0.381)	(0.590)	(0.533)	(0.613)	(0.041)
	(B) First Stage Results				
First Stage F-statistic	20.10	22.56	10.52	10.25	10.92
Observations	94,167	84,940	108,153	113,676	115,157

Notes: This table presents first and second stage IV results for firm performance variables(operating income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility) regressed on board fine-grained caste homophily and other control variables, and year and two-digit industry fixed effects for instrumental variable approach 3 (Clause 49). Panel (B) presents the first stage results where the excluded instruments are constructed based on a regression discontinuity design as shown in equation (4.2). Panel (A) presents second stage results. Control variables include firm age, leverage, real assets, tangibility, export status, three digit industry HHI, whether the firm belongs to a business group, whether the firm is listed or not, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, and the proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). The first stage F statistic is the Kleibergen-Paap rk Wald F statistic. Robust standard errors, clustered by industry and year, are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 12: Instrumental Variable Regressions, First and Second Stage: Fine-Grained Caste Homophily, IV Approach 4

	(A) Second Stage Results				
	(1)	(2)	(3)	(4)	(5)
Variables	Log(Operating Income/Sales)	Log(Operating Cash Flow/Sales)	Log(Profits/Sales)	Tobin's Q	Volatility
Fine-Grained Homophily	-0.636	6.001***	0.603	-0.284	0.047***
	(1.202)	(1.640)	(1.211)	(1.604)	(0.014)
	(B) First Stage Results				
Share exposed to other castes	-0.061***	-0.056***	-0.058***	-0.062***	-0.061***
	(0.011)	(0.009)	(0.010)	(0.010)	(0.011)
First stage F-statistic	28.66	37.66	32.93	34.96	33.78
Observations	4,895	4,381	3,943	5,456	5,458

Notes: This table presents first and second stage IV results for firm performance variables(operating income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility) regressed on board fine-grained caste homophily and other control variables, and year and two-digit industry fixed effects. Panel (B) presents the first stage results. Excluded instrument is the fraction of a firm's directors that serve on the board of at least one other firm whose dominant fine-grained caste is different from directors' own fine-grained caste. Panel (A) presents second stage results. Control variables include age, board size, board size square, leverage, real assets, tangibility, export status, three digit industry HHI, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, proportion of group firms in the industry, and the total number of other directorships in the board. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). The first stage F statistic is the Kleibergen-Paap rk Wald F statistic. Robust standard errors, clustered by industry and year, are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 13: Instrumental Variable Regressions (Industry Adjusted): Fine-Grained Caste Homophily, IV Approach 1

	(A) Second Stage Results				
	(1)	(2)	(3)	(4)	(5)
Variables	Log(Operating Income/Sales)	Log(Operating Cash Flow/Sales)	Log(Profits/Sales)	Tobin's Q	Volatility
Fine-Grained Homophily	-1.729**	-1.234	-2.512**	-3.010***	0.011
	(0.872)	(0.932)	(1.098)	(0.974)	(0.026)
	(B) First Stage Results				
State Fine-Grained Caste Homophily	0.507***	0.532***	0.492***	0.479***	0.477***
	(0.109)	(0.145)	(0.103)	(0.120)	(0.123)
Industry Fine-Grained Caste Homophily	0.334***	0.367***	0.373***	0.197	0.186
	(0.081)	(0.136)	(0.088)	(0.144)	(0.131)
First Stage F-statistic	21.32	10.83	21.43	10.85	11.52
Observations	46,754	33,987	57,505	27,164	28,811

Notes: This table presents first and second stage IV results for the difference of all firm performance variables(operating Income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility) from the corresponding industry (two digit NIC) medians in that year, regressed on board fine-grained caste homophily and other control variables, and year fixed effects. Panel (B) presents the first stage results. Excluded instruments are the fine-grained caste homophily levels of the state and two digit industry director supply pools for a firm. Panel (A) presents the second stage results. Control variables include age, board size, board size square, leverage, real assets, tangibility, listing status, export status, three digit industry HHI, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, and the proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). The first stage F statistic is the Kleibergen-Paap rk Wald F statistic. Robust standard errors, clustered by industry and year, are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 14: Instrumental Variable Regressions (Industry Adjusted): Fine-Grained Caste Homophily, IV Approach 2

	(A) Second Stage Results				
	(1)	(2)	(3)	(4)	(5)
Variables	Log(Operating Income/Sales)	Log(Operating Cash Flow/Sales)	Log(Profits/Sales)	Tobin's Q	Volatility
Fine-Grained Homophily	-0.204***	-0.323***	-0.210**	-0.256***	0.007***
	(0.049)	(0.099)	(0.084)	(0.050)	(0.002)
	(B) First Stage Results				
State Fine-Grained Caste Homophily	0.490***	0.452***	0.481***	0.425***	0.433***
	(0.074)	(0.063)	(0.073)	(0.074)	(0.072)
Industry Fine-Grained Caste Homophily	1.177***	1.189***	1.219***	0.931***	0.934***
	(0.116)	(0.124)	(0.107)	(0.150)	(0.151)
Distance from State Director	0.326***	0.290***	0.317***	0.296***	0.296***
Fine-Grained Caste Composition	(0.059)	(0.052)	(0.060)	(0.049)	(0.047)
Distance from Industry Director	0.976***	0.965***	0.992***	0.862***	0.869***
Fine-Grained Caste Composition	(0.062)	(0.053)	(0.063)	(0.053)	(0.051)
First Stage F-statistic	8211.87	5742.46	8064.20	1887.93	1803.63
Observations	46,754	33,987	57,505	27,164	28,811

Notes: This table presents first and second stage IV results for the difference of all firm performance variables(operating Income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility) from the corresponding industry (two digit NIC) medians in that year, regressed on board fine-grained caste homophily and other control variables, and year fixed effects. Panel (B) presents the first stage results. Excluded instruments are the fine-grained caste homophily levels of the two director supply pools and the distance of fine-grained caste composition of the firm's board from that of the two supply pools. Control variables include age, board size, board size square, leverage, real assets, tangibility, listing status, export status, three digit industry HHI, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, and the proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). The first stage F statistic is the Kleibergen-Paap rk Wald F statistic. Robust standard errors, clustered by industry and year, are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 15: Instrumental Variable Regressions, First and Second Stage, Above and Below Median: Fine-Grained Caste Homophily, IV Approach 2

	(A) Second Stage Results, Below Median				
	(1)	(2)	(3)	(4)	(5)
Variables	Log(Operating Income/Sales)	Log(Operating Cash Flow/Sales)	Log(Profits/Sales)	Tobin's Q	Volatility
Fine-Grain Homophily	-0.032	-0.052	-0.010	-0.027***	0.002**
	(0.055)	(0.079)	(0.044)	(0.008)	(0.001)
First Stage F-statistic	11816.15	6106.25	7841.93	1970.31	960.76
Observations	23,753	17,382	29,063	13,590	14,410
	(B) Second Stage Results, Above Median				
Fine-Grained Caste Homophily	-0.122***	-0.019	-0.170*	-0.267***	0.003**
	(0.045)	(0.069)	(0.096)	(0.081)	(0.001)
First Stage F-statistic	4820.48	3199.22	6170.64	961.19	2335.64
Observations	23,000	16,604	28,441	13,570	14,398

Notes: This table presents first and second stage IV results for all firm performance variables (operating Income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility), regressed on board fine-grained caste homophily and other controls, and year fixed effects for two set of firms. Panel (A) presents the first and second stage results for firms with performance variable below the two digit industry median. Panel (B) presents the first and second stage results for firms with performance variable above the two digit industry median. In both panels, the excluded instruments are the fine-grained caste homophily levels of the two director supply pools and the distance of fine-grained caste composition of the firm's board from that of the two supply pools. Control variables include age, board size, board size square, leverage, real assets, tangibility, listing status, export status, three digit industry HHI, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, and the proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). The first stage F statistic is the Kleibergen-Paap rk Wald F statistic. Robust standard errors, clustered by industry and year, are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 16: Fine Grained Homophily and Firm Outcomes: Coarse-Grained Caste Subsamples, IV Approach 2

Variables	(1) Log(Operating Income/Sales)	(2) Log(Operating Cash Flow/Sales)	(3) Log(Profits/Sales)	(4) Tobin's Q	(5) Volatility
(A) Brahmin					
Fine-Grain Homophily	-0.317**	-0.535***	-0.238	-0.276*	0.012**
	(0.147)	(0.174)	(0.228)	(0.151)	(0.005)
First Stage F-Statistic	2077.39	876.12	2247.38	415.99	365.62
Observations	7,701	6,020	9,581	4,916	5,155
(B) Kshatriya					
Fine-Grain Homophily	-0.162*	0.169	-0.142*	-0.260***	0.018***
	(0.090)	(0.169)	(0.086)	(0.090)	(0.006)
First Stage F-statistic	3753.43	1302.01	3416.02	504.99	584.78
Observations	6,003	4,152	7,125	3,110	3,319
(C) Vaishya					
Fine-Grain Homophily	-0.317***	-0.495***	-0.239**	-0.183**	0.007***
	(0.058)	(0.093)	(0.100)	(0.077)	(0.003)
First Stage F-Statistic	5377.72	2895.72	5687.01	915.96	1004.04
Observations	16,526	12,345	20,682	10,465	11,116

Notes: This table presents first and second stage IV results for all firm performance variables (operating Income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility), regressed on board fine-grained caste homophily and other control variables, and year and industry fixed effects for three set of firms. Panel (A), (B), and (C) present the first and second stage results for firms whose dominant coarse-grained caste is *Brahmin*, *Kshatriya* and *Vaishya*, respectively. In all panels the excluded instruments are the fine-grained caste homophily levels of the two director supply pools and the distance of fine-grained caste composition of the firm's board from that of the two supply pools. Control variables include age, board size, board size square, leverage, real assets, tangibility, listing status, export status, three digit industry HHI, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, and the proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). The first stage F statistic is the Kleibergen-Paap rk Wald F statistic. Robust standard errors, clustered by industry and year, are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 17: Fine Grained Homophily and Firm Outcomes: Coarse-Grained Caste Subsamples, IV Approach 2

		(D) Jain			
Variables	Log(Operating Income/Sales)	Log(Operating Cash Flow/Sales)	Log(Profits/Sales)	Tobin's Q	Volatility
Fine-Grain Homophily	-0.224***	-0.503*	-0.319*	-0.240***	0.003
	(0.085)	(0.279)	(0.163)	(0.088)	(0.004)
First Stage F-statistic	4681.08	3529.37	3967.96	648.97	699.88
Observations	4,560	3,225	5,747	3,010	3,250
		(E) Shudra			
Variables	Log(Operating Income/Sales)	Log(Operating Cash Flow/Sales)	Log(Profits/Sales)	Tobin's Q	Volatility
Fine-Grain Homophily	-0.131	-0.231	-0.203*	0.048	0.001
	(0.132)	(0.244)	(0.111)	(0.173)	(0.010)
First Stage F-statistic	1686.22	1104.10	1883.80	688.21	596.09
Observations	8,623	5,917	10,230	4,115	4,337

Notes: This table presents first and second stage IV results for all firm performance variables (operating Income, operating cash flows, profits, all relative to sales, and Tobin's Q and volatility), regressed on board fine-grained caste homophily and other control variables, and year fixed effects for two sets of firms. Panel (D) and (E) presents the first and second stage results for firms whose dominant coarse-grained caste is *Shudra* and non-Hindu religion is *Jain*, respectively. In both panels, the excluded instruments are the fine-grained caste homophily levels of the two director supply pools and the distance of fine-grained caste composition of the firm's board from that of the two supply pools. Control variables include age, board size, board size square, leverage, real assets, tangibility, listing status, export status, three digit industry HHI, whether the firm belongs to a business group, dominant caste of the board, shares of firms with various dominant coarse-grained castes in three digit industry, and the proportion of group firms in the industry. All variables are defined in Table 1. All financial variables are winsorized at the 1st and 99th percentiles over the sample period (1999-2015). The first stage F statistic is the Kleibergen-Paap rk Wald F statistic. Robust standard errors, clustered by industry and year, are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 18: Committee Chair & Membership Shares Held by Dominant Fine-Grained Caste

	Dominant Caste	Second Most Dominant Caste	Third Most Dominant Caste
% of directors in nth dominant caste of board (denominator for subsequent rows)	52.29	22.15	11.93
Ratio of share of committee chairs of *n*th dominant caste to share of all directors of *n*th dominant caste	0.99	1.29	1.32
Ratio of share of audit committee members of *n*th dominant caste to share of all directors of *n*th dominant caste	0.88	1.11	1.15
Ratio of share of remuneration committee members of *n*th dominant caste to share of all directors of *n*th dominant caste	0.84	1.17	1.22
Ratio of share of grievance committee members of *n*th dominant caste to share of all directors of *n*th dominant caste	1.02	1.02	1.03

Source: Prowess, matrimonial data. The table shows, by dominance status of a fine-grained caste, proportions of particular positions occupied by directors of that caste, relative to their share in the full board. A fraction greater than 1 indicates over-representation.