Cultural Homophily in Corporate Boards and Firm Outcomes

Suresh Bhagavatula* Manaswini Bhalla†
IIM, Bangalore IIM, Bangalore

Manisha Goel‡ Balagopal Vissa§
Pomona College INSEAD

March 2018

Abstract

We show that boards of directors of large Indian firms are characterized by high levels of cultural proximity, with members on a board belonging overwhelmingly to the same religion or caste. Using a unique database of self-reported religions and caste from matrimonial websites, we develop a novel methodology to probabilistically map individuals' last names to religions and castes. We also develop a new homophily index to measure cultural proximity of board members. Results show few signs of increase in cultural diversity on boards during 1999-2012. Modest heterogeneity exists across firms, sectors, and states, however. Better performing firms have more diverse boards. Board diversity also increased in sectors and states that witnessed the largest increases in output. Rigorous analysis demonstrates that lack of diversity on boards is causally associated with lower firm performance.

JEL Classifications: Z13, G34, M14, L14, J44

Keywords: Corporate Culture, Corporate Governance, Boards of Directors, Caste

*sureshbh@iimb.ernet.in

†manaswinib@iimb.ernet.in

‡manisha.goel@pomona.edu (Corresponding Author)

§Balagopal.vissa@insead.edu. We gratefully acknowledge the financial support of Pomona College and Indian Institute of Management, Bangalore. We thank Tahir Andrabi, Asim Khwaja, Rajiv Sethi, Simi Kedia, and Michelle Zemel for helpful conversations and seminar participants at AALIMS 2017, IIM Bangalore, Indian School of Business, and Indian Statistical Institute for their comments. Thanks to VSK Teja Konduri, Timothy Howe, Charles Fries, Leena Kinger Hans, and Sharada PM for valuable research assistance.

1 Introduction

Social interactions are characterized by cultural homophily, along shared traits such as language, race, gender, ethnicity, or tribe.[1] A growing literature shows that cultural proximity between agents can influence their economic decisions. For instance, shared ethnicity can be a factor underlying the choice of research collaborators, and shared race can influence hiring decisions.[2] In this paper, we study the prevalence and intensity of such homophily in teams with several agents engaged in joint economic decision making. In particular, we examine whether, and to what extent, directors on boards of large firms in India tend to be culturally proximate. We find that corporate boards are characterized by extremely high levels of homophily over time, across regions, and across industries. We further assess if the extent of homophily depends on how granularly directors' cultural identity is identified, and find a striking persistence of homophily at all aggregation levels, even for narrowly defined groups. Finally, we ask how this homophily impacts firm outcomes, and show that firm value and performance are perversely affected by high cultural homogeneity of boards.

A major challenge faced by studies focusing on cultural proximity is the identification of individuals' cultural groups. As a result, studies thus far have considered homophily based on easily identifiable coarse groups at high level of aggregation, such as gender, race, or country of origin. India offers a uniquely useful cultural context to overcome this challenge. The country has a multi-religious society, and Hindus, the dominant religious group, are divided into hundreds of hierarchical subgroups by the caste system, a traditional institution that has endured since c. 1300 B.C.[3]

Using data from India, and a novel computational methodology, we develop a data driven mapping of corporate directors into distinct cultural groups at varying levels of aggregation. Data on firm board members and other characteristics are taken from

[1] McPherson et al. (2001).

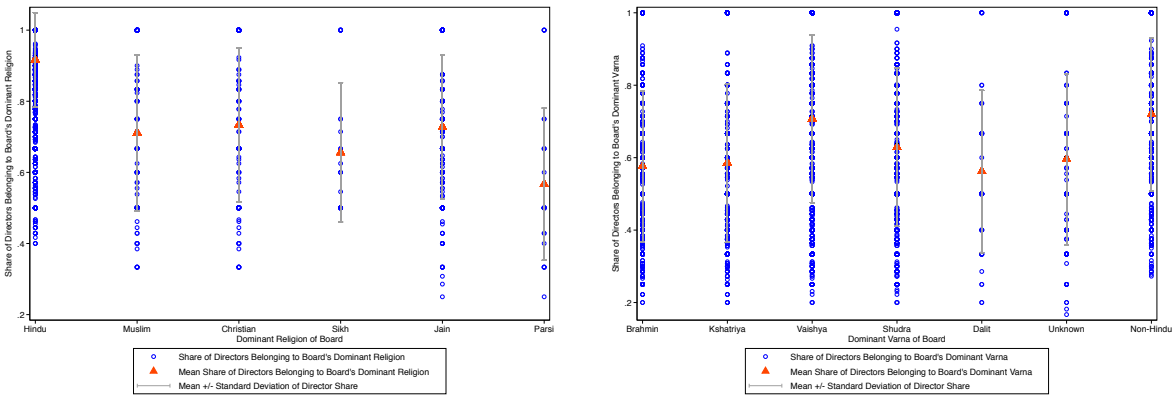
[2] See, for example, Freeman and Huang (2015), Aslund et al. (2012), Giuliano and Ransom (2011), Giuliano et al. (2009), and Petersen et al. (2000).

[3] Macdonell (1914). The Hindu society is divided into four hierarchical *varnas – Brahmins, Kshatriyas, Vaishyas*, and *Shudras*, in that order. There is a fifth, de facto, "*varna*" of "Dalits," a term that was used by the government to describe those disadvantaged groups who were considered "untouchables" and formally outside the caste system. The *Shudras* and *Dalits* historically faced significant discrimination. Today they are respectively designated "other backward castes (OBC)" and "scheduled castes (SC)" by the Indian government and are beneficiaries of several affirmative action programs. Affirmative actions are also aimed at uplifting some tribal populations that are classified as "Scheduled Tribes (ST)". Within the five *varnas*, there are thousands of sub-castes or *jatis*. An individual belongs to a certain *jati* and *varna* based on her lineage. Historically, castes are endogamous and have been associated with occupations, although there is some flexibility (Srinivas (1966)). As a result, caste, albeit a cultural construct, has influenced the socio-economic status of individuals.

Prowess – a database of financial information on large public and private firms. To identify the castes of directors, we exploit the facts that individuals' last names are indicative of their religion and caste, and that Indian marriages are overwhelmingly intra-religion, with Hindu marriages, in particular, being predominantly intra-caste. We obtain information from three prominent matrimonial websites on the first and last names of nearly six million registered users and their self-reported religions and castes. Since the mapping between last name and caste/religion is not always one-to-one, we use these data to assign probabilities with which a last name belongs to each religion, and for Hindu names, each caste (*varna*) and subcaste (*jati*). We apply this mapping to the corporate directors in our sample. Thus, the cultural groups identified for each director range from a broad classification into six religions to an extremely granular classification that groups people into 472 distinct subcastes.

Having built this innovative dataset, we use it first to document a rich set of stylized facts about the levels and trends in religion and caste homophily in Indian corporate boards during 1999-2012. We show that corporate boards in India are characterized by high levels of cultural proximity among directors. Figures 1(a) and 1(b) show the proportions of boards constituted by directors belonging to the boards' dominant religion and *varna* using data for 2004. We see that each religion (*varna*) finds majority representation across boards, and that majority tends to be high, ranging from 55% (parsis) to 90% (hindus) for religion and 50% (*dalits*) to 65% (*vaishyas*) for (*varna*). In addition to displaying high caste proximity, Figure 1(b) also shows that individuals of all *varnas* tend to associate with those of their own *varna*, even when the *varna* is a socially-disadvantaged one. We also find that homophily varies systematically across states, sectors, and firm types indicating that higher economic growth or performance are associated with lower cultural homophily on boards. Finally, low directorial cultural homophily is associated with better corporate governance indicators such as high proportion of independent directors and no CEO duality.

These stylized facts demonstrate that cultural institutions of caste and religion strongly influence corporate board membership in India. We next examine whether the resulting lack of cultural diversity among board members impacts firm performance. The direction of this impact, if any, is unclear ex ante and depends on the factors underlying the mechanisms through which cultural considerations influence incumbent directors' decisions about whom to hire as new directors. High cultural homophily on a corporate board may result from a few underlying factors. (a) Hiring a new director in the presence of information asymmetry about their skills, may lead incumbent directors to rely on their social networks, which tend to be culturally homogeneous, as shown previously (see



(a) (b)

Figure 1: Shares of Directors Belonging to Boards' Dominant Cultural Group[a]

[a]Source: Prowess, authors' mapping of last names to religion and *varna* using matrimonial data. Figure 1(a) (1(b)) shows the proportion of directors that belong to the religion (*varna*) with the highest representation on a board, as well as the overall average across boards, for each religion (*varna*). Both figures use data for the year 1999.

McPherson et al. (2001)). In this case, the new firm director may contribute to increasing the religion or caste homophily of the board. (b) Alternatively, high cultural homophily of boards may reflect favoritism being exercised by incumbents in hiring new members, i.e., preferring potential new directors belonging to similar cultural backgrounds as themselves, with less regard to their formal qualifications for the position. This may, in part, be driven by greater trust of members of the same community as oneself. (c) Yet another possibility is that directors are hired solely on the basis of their human capital and they happen to be of the same caste or religion as the incumbents. This can happen either (c1) through pure coincidence or (c2) because the socio-economic status of different cultural groups implies that qualified candidates tend to be of a given cultural background.

Let us consider what each of these scenarios implies about the effect of homophily on firm performance. In scenario (a), the association between board homophily and firm performance is unclear. Reliance on social networks when hiring directors can lead to new members either having similar opinions as the incumbents' or being unable to express their independent views. In this case, higher board homophily can negatively affect firm outcomes. On the other hand, having board members from the same social network can increase trust in the group thereby reducing information frictions and allowing for more coordinated actions (Kramarz and Thesmar (2013)). In the latter case, firm outcomes may be positively influenced by greater cultural proximity among its directors. Under scenario (b), firm performance is likely to be negatively impacted by higher levels of

homophily on the board. If high observed homophily is coincidental, as in (c1) above, then we expect to see little to no effect on firms' outcomes. Finally, under scenario (c2), higher homophily may have a positive effect on firm outcomes.

Of the above potential mechanisms, we can rule out homophily purely by coincidence since board homophily varies systematically across firms and sectors as shown in Figure 1 and in section 3. We can also rule out the possibility that individuals of some religions or castes tend to have higher human capital for two reasons. First, individuals serving as directors on firms presumably already have a high socio-economic status and human capital, even if they belong to historically disadvantaged groups. Second, as shown in section 3, we observe homophily even among groups that we would expect to be socially disadvantaged so that some boards are dominated by disadvantaged groups. Yet, these firms' performance is not systematically any different from that of firms whose boards are dominated by socially privileged groups. However, we cannot distinguish between mechanisms (a) and (b). Thus, the direction of the causal association between cultural homophily of boards and firm performance is unclear ex ante. Regression analysis, therefore, provides us with estimates of the net effect of these mechanisms. Results from this analysis show that higher cultural homophily on corporate boards (using all measures of cultural identity) negatively affects firm performance (return on assets and profitability) and firm value (Tobin's Q).

Our paper relates to the rapidly growing literature on the interplay between culture and economic outcomes. Studies have examined the influence of culture on a wide range of outcomes. We highlight the focal points of this research and a few salient studies. Many papers have examined how cultural attitudes and assimilation, religion, and family values impact economic growth (see Guiso et al. (2003), Barro and McCleary (2003, 2006), Noland (2005), Ashraf and Galor (2007), Tabellini (2010), Fernandez (2010), Alesina and Giuliano (2010), and Campante and Yanagizawa-Drott (2015)) and a range of other economic outcomes.[4] One strand of this literature focuses on how cultural identity of individuals shapes social and economic networks (see, for example, Currarini, Jackson, and Pin (2009)). Studies have also analyzed how cultural proximity in dyads impacts economic decisions. For example, studies have documented that shared ethnicity between entrepreneur and venture capitalist increases the likelihood of the VC financing

[4]Other outcomes studied include institutions (see Alesina and Giuliano (2014) for an excellent review), public good provision (Alesina et al. (2017), Benjamin et al. (2012), Alesina et al. (1999)), management practices and organization of firms (Bloom et al. (2012), Bloom et al. (2014) and Bloom et al. (2016)), trade and economic exchange (Anderson (2011)), nutrition (Atkin (2016)), female labor force participation (Alesina et al. (2013), Fernandez (2013, 2007), Fernandez and Fogli (2009, 2006), among others), inequality (Alesina et al. (2016)) and political outcomes (Alesina (2011), Alesina and Giuliano (2011), Gorodnichenko and Roland (2015)).

the startup (see Bengtsson and Hsu (2015) and Hegde and Tumlinson (2012)). Cultural proximity has also been shown to affect hiring and firing decisions (for example, Aslund et al. (2012), Giuliano and Ranson (2011), Giuliano et al. (2009), and Petersen et al. (2000)), research collaborations (Freeman and Huang (2015), loan outcomes (Fisman et al. (2017)), and educational outcomes (Dee (2005) and Fairlie (2014)). Our paper is different in that we analyze the effect of cultural identity on *teams* of multiple agents and their joint economic decision making. We also contribute to the broader literature on culture and economic outcomes in several ways. We develop an innovative methodology that maps the same individuals into three hierarchical cultural groups at different aggregation levels. This enables us to measure cultural identity at aggregated as well extremely granular levels. Our evidence shows high and persistent homophily at all aggregation levels. This suggests that even though members of different narrowly defined cultural groups may in fact be quite similar to each other, they still tend to associate with members of their own groups. Further, our results demonstrate that such cultural homophily on corporate boards negatively affects firm value and performance.

Several studies have analyzed the effects of board composition on corporate governance and firm outcomes. For instance, Knyazeva et al. (2013) and Dass et al. (2013 examine the effects of independent directors and director experience in related industries on firm performance, respectively. An important strand of this literature focuses on board diversity, but the dimension examined is almost exclusively gender (see, for example, Ahern and Dittmar (2012) and Bertrand et al. (2014). Terjesen et al. (2009) provide an excellent review.). A few exceptions include Bernile et al. (2016) who develop a multidimensional diversity index and Kramarz and Thesmar (2013) who show that social networks formed through shared alma maters strongly influence memberships on firm boards in France and adversely affect corporate governance. We add to this body of work by considering the *cultural* proximity of directors, as measured by the traditional institutions of religion and caste. The impact of such traditional institutions on board composition has not been previously explored. Note that cultural proximity, the focus of our paper, is inherently different from *social* ties between people. As described in Fisman et al. (2017), while social ties are formed endogenously by conscious choice, culture is inherited at birth. Individuals may feel affinity to those who are culturally proximate, even if they have never met or interacted before. This is corroborated by our finding that boards systematically display high religion and caste homophily. This finding is also remarkable since it shows that the traditional constructs of caste and religion continue to significantly influence a modern institution of corporate boards.

This paper also contributes significantly to the literature examining the economic

effects of caste in India. The majority of previous studies compare secular trends in socio-economic outcomes of disadvantaged castes, such as SC, ST, and OBC, to those of advantaged upper castes, in the context of affirmative actions or discrimination. For example, Hnatkovska, Lahiri, and Paul (2012, 2013) find that educational and occupational composition and mobility of backward caste groups in India have improved over time, converging to that of upper caste households. Iyer et al. (2013) show that entry into entrepreneurship among backwards castes continues to lag behind that of upper castes.[5]

However, we approach the economic effects of caste through a different lens – does caste proximity between individuals influence economic decisions and outcomes, regardless of whether the caste itself is underprivileged or not? Only a few studies have taken a similar approach previously. Fisman et al. (2017) show that when borrowers and loan officers belong to the same caste or religion dyad, it increases the likelihood of a loan being made and being repaid in the future. Thus, cultural considerations yield a positive economic outcome. Munshi (2013) documents a different mechanism through which caste can have a positive economic impact – relying on caste networks increases intergenerational occupational mobility. However, Munshi and Rosenzweig (2016) argue that caste networks can also have a negative economic effect by showing that reliance on these networks as informal insurance mechanisms prevents rural workers from earning a higher wage by migrating to urban areas, lest such insurance is lost.[6] Besides examining an entirely different economic outcome of caste and religion proximity, we differ from these studies in a few important respects. First, while these previous papers focus on rural areas, particular cases, or traditional businesses, we show that cultural proximity shapes economic behavior and outcomes even in highly urban and corporate environments all across the country. Second, in the absence of self-identified caste/religion, we develop a data-driven approach to probabilistically map individuals' last names to caste and religion that can be used for future work. This data-driven approach and focus on corporate boards also distinguishes us from Chen, Chittoor and Vissa (2014) who show that caste ties between equity analysts and CEOs of Indian firms serve as conduits for information flows and hence improve the company forecasts of analysts, and Vissa (2011), who shows that entrepreneurs are likely to form economic ties with new people they meet if they belong to the same caste as themselves.

To the best of our knowledge, only one previous study by Ajit, Honker, and Saxena (2012) has examined the role of caste in membership on corporate boards in India. Our

[5]See also Ghani, Kerr, and O'Connell (2011), Damodaran (2008), Thorat, Kundu and Sadana (2010), Jodhka (2010), and Varshney (2012).

[6]See also Anderson (2011), Munshi and Rosenzweig (2006, 2013), and Banerjee and Munshi (2004).

study differs from their work in several respects. While these authors assigned caste affiliations to directors relying on their own conjectures and information gathered online, we adopt a data-driven approach and develop a probabilistic mapping of last names to caste. Moreover, they do not attempt an investigation into the implications of caste concentration on boards for firm performance. Further, while their caste groups are defined coarsely into *varnas* and SC/ST/OBC, we are also able to identify *jatis*, which are substantially finer.[7] We also use a much larger sample of firms and years.

The rest of the paper is organized as follows. In section 2, we describe our data sources and the construction of a homophily measure. Section 3 presents stylized facts about religion homophily in India, drawing comparisons across sectors and firm characteristics, and over time. In section 4, we describe our empirical strategy to identify the causal effect of cultural homophily on boards on firm performance. Results are presented in section 5. Section 6 concludes.

2 Data and Construction of Homophily Index

We use three main data sources: (1) Data from matrimonial websites (*Jeevansathi.com, Bharatmatrimony.com and Shaadi.com*): to probabilistically match last names to religion, *varna* and *jati*, (2) Prowess database: for information on boards of directors and other firm characteristics, and (3) Indian Boards Database: to document board interlocks and examine the relevance of our instrument strategy. The first part of the section describes these data sources. Next, we discuss the homophily index constructed to measure cultural proximity of board members with respect to their cultural identity – religion, *varna* and *jati*.

2.1 Data from Matrimonial Websites

Social and cultural practices relating to naming customs in India are complex. In general, however, the last names adopted by Indians are indicative of their religion and caste (Dumont (1980)).

We use a unique data set of self reported castes from three popular matrimonial sites of India.[8] *Shaadi*.com, *Jeevansathi*.com and Bharatmatrimony.com are three of the most

[7]Biswas (2016) also analyzes linguistic homophily of Indian boards but does not explain how she identifies linguistic affiliations of board members.

[8]In India's fast evolving socio-economic setting, where the role of traditional kinship networks and local matchmakers is fast diminishing, matrimonial websites act as an alternative to traditional marriage brokers by nationalizing the pool of prospective spouses according to different categories like religion and caste. According to The New York Times, there are over 1500 matrimony websites in India. According to

prominent and largest matrimonial services with Alexa ranks in India of 49, 289, and 1,116 respectively. The historically endogamous institution of caste continues to be a predominant factor based on which matches are determined in India. This importance of caste is reflected in such websites, where prospective brides/grooms self-report their castes and preferences for marrying outside caste.

Probabilistic Mapping of Last Names to Religion and Caste:

The raw data obtained from the three matrimonial websites include over six million self-made profiles, which provide information on an individual's first and last names, native language, religion, and caste (*jati*). To build a robust mapping, we drop all last names that only appear once in the database. After a considerable amount of cleaning, we are left with 5,447,129 profiles, spanning 14,374 unique last names, 8 unique religions, 5 unique *varnas*, and 472 unique *jatis*.

These data show that the same last name may be associated with more than one religion or caste, often depending on the geographical region. Moreover, the same last name can have a few different spelling variations. In building a concordance between last names and religion/caste, we take both of these factors into account. We describe the methodology below.

The vast majority of names are original words in Indian languages (e.g. Hindi, Tamil, Marathi, etc.), whereas the websites from which the data is culled are in English. Therefore, an English equivalent (not translation) of these names are represented in the data. Thus, in many cases, multiple English spellings of the same Indian last name are represented. To accurately map last names to castes, we need to collate all the alternative spellings of the same last name. To do so, we use two different word matching algorithms to predict the similarity of different last names. If the similarity predictions from both the algorithms are above a certain threshold, then the two last names are considered to be the same and their caste mappings are combined.

The first word matching algorithm is a modified Levinstein distance algorithm. In this algorithm, a distance measure between two stings is calculated using a dynamic programming approach, with each replacement alphabet adding one to the distance measure. The standard Levienstein distance measure is appropriate for English words, whereby each difference in alphabets between two words contributes equally to the distance measure. However, since we want to match the phonetic translation of Indian languages, we de-

the Associated Chambers of Commerce and Industry in India, the online matrimony business is expected to be worth $250 Million by 2017 (Titzmann (2013).

velop a modified Levinstein distance algorithm. In this method, changes in alphabets that constitute a sound in Indian languages are bundled together and assigned a unit distance measure. In addition, alphabets that have the same sound are assigned zero distance. The final distance calculated using this algorithm provides a measure of how similar two words are in an Indian language. In the current application, 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 word.

The last names are also matched using a modified version of the Ratcliff-Obershelp algorithm. This algorithm looks for matches in the longest contiguous matching subsequence of two words and assigns a matching score. The original algorithm was developed to find sequence matches between two sentences, and it was found to be appropriate for matching words that are phonetic translations from other languages. A detailed description is available in the Python library. For our application, if the algorithm provided a match score of greater than 85%, the two last names were deemed to be matched.

As a final step, we deem two last names to be fully matched if both the algorithms project that the corresponding strings are matched. This approach is borrowed from the concept of bagging used in machine learning based classification literature where votes from different classifiers are used together to increase the robustness of final prediction. This name matching results in groups of similar sounding last names that have different spellings. Thus, for example, one last name group consists of *Goel* and *Goyal*, and another consists of *Rathore*, *Rathor* and *Rathour*.

We observe that the same last name systematically maps to different castes according to the geographical regions. We exploit the fact that different regions of India have different dominant languages. Therefore, the mapping from the last name to the caste is conditioned on the first language identified by the user. Following this approach, we count the total number of occurrences of each last name that is associated with the same language. In case of a last name group, each name in the group is assigned the sum of occurrences of all last names in the group, subject to the language being the same.[9] We take all religions/*jatis* reported for each last name over all its (within-language) occurrences and count their respective occurrences. Again, in case of a group, we identify the superset of all religions/*jatis* reported across all spelling variations of the same last name (within-language) and count their occurrences. Dividing the number of times a last name is associated with a particular religion/*jati* 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/*jati*. Doing so over all religions and

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

jatis associated with a last name gives us the probabilistic mapping of each last name to religions/*jatis*. In case of a name group, each name variant in the group is assigned the same probability distribution over religions/*jatis*.[10]

While the institution of caste originated in Hindu society, it is adopted de facto in many non-Hindu religions too (Dumont (1980)), although it may be weakened or incomplete. This is demonstrated in our data with some individuals only reporting their (non-Hindu) religion while others additionally reporting a *jati*. For our analysis, we assume that last names reported with a non-Hindu religion (7555 in number) do not have a *jati* associated with them. Therefore, in the last name to caste (*jati*) mapping, we simply assign the category of "not applicable (NA)" to these names.

We use the data generated mapping of last names to *jatis* to also map them to *varnas*. We rely on government of India reports and Wikipedia entries to identify this mapping. Thus, there is subjectivity and authors' judgment involved in the last name to *varna* mapping. This mapping yields seven *varna* categories – *Brahmin*, *Kshatriya*, *Vaishya*, *Shudra*, *Dalit*, others, and NA (not applicable). "Others" is the category given to names for whom we are unable to find a *jati* to *varna* mapping, or if a *jati* maps to multiple *varnas*. "NA" is assigned to non-Hindu and foreign sounding names.[11]

Table 1: Last Names Mapped to Religion and Caste – Examples

Last Name	**Total Occurrence**	**Religion**	***Varna***	***Jati***	
			Probabilities in Parentheses		
Wadhwani	335	Hindu (1)	Vaishya (1)	Sindhi (0.99)	Arora (0.01)
Dandriyal	11	Hindu (1)	Brahmin (1)	Brahmin Garhwali (0.82)	Brahmin Pandit (0.18)

Two examples of the resulting probabilistic mappings of last names to religion/*varna*/*jati* are provided in Table 1. Both last names are associated with a single religion and *varna* but have two possible *jatis*. More generally, the mappings have the following basic features. Each last name is associated with an average of 1.58 and a maximum of 6 religions, with 68% of last names being associated with a single religion. The probabilistic mass is

[10]In one region of South India (the state of Tamil Nadu), the last name of an individual is simply 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 *jati* instead of the last name. Although even first names vary by *jati*, we expect the mapping from first name to *jati* to be noisier when compared to mapping last names that remain invariant across generations. This suggests that the probabilistic mapping for Tamil names may be noisier than the mapping for non-Tamil names.

[11]There are no foreign sounding names in the matrimonial data. However, we do see such names in the sample of directors.

Table 2: Religion and Varna Composition of Matrimonial Data

Religion	Percent of Last Names	Varna	Percent of Last Names
Hindu	81.35	Brahmin	18.66
Muslim	6.67	Kshatriya	13.18
Christian	7.21	Vaishya	12.81
Jain	2.67	Shudra	31.84
Sikh	1.49	Dalit	1.90
Zoroastrian	0.55	Others	1.47
Buddhist	0.05	NA	20.15
Jew	0.01	–	–

concentrated in just the top two religions, with them jointly accounting for about 98% of the total likelihood, on average. As for *varnas (jatis)*, last names are associated with an average of 3.4 (9.8) and a max of 7 (138) *varnas (jatis)*. Table 2 shows the most likely religion and *varna* composition of names in the final matrimonial sample.[12] As expected, Hindus, at 81.35%, form the overwhelming majority of names. Christians and Muslims, 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. According to the 2011 census, Hindus constitute 80% of the populations, 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 *varna* or *jati*.

We use these concordances to assign religions and castes to corporate directors serving on firms in our sample, described next.

2.2 Firm Level Data

Data on Indian firms is obtained from the Prowess database provided by the Centre for Monitoring the Indian Economy (CMIE). This data source covers publicly (listed and unlisted) and privately held, and government owned firms over the period 1989-2012. This data source includes detailed information on the identity (entity type, ownership, Industry, and age etc), and governance and financial aspects of these firms. We choose

[12]For space considerations, we do not report the *jati* composition of matrimonial data but the information is available upon request.

to use data for the period 1999-2012, as the number of firms covered by Prowess is much smaller prior to 1999.

Prowess covers firms for which it can obtain publicly available information. The data are sourced mainly from annual reports, quarterly financial statements, and profit and loss accounts of firms. Thus, information on all listed companies that are reasonably active on the major stock exchanges of India is available in the database. Though the database includes mostly publicly listed firms, a smaller number of unlisted (public and private) firms are also included. The reason for smaller coverage of these firms is that they are not required to publicly disclose their financial statements.[13] Prowess does not cover informal firms since there is little publicly available information on them. Thus, the database does not cover the universe of all firms in India. Nonetheless, the firms included account for a substantial proportion of the economic activity in India, contributing 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. While the database follows firms longitudinally, most firms appear in the data only for a few years. Thus, for our analyses, we treat these annual data as repeated cross sections of firms, instead of a panel of firms.

Securities Exchange Board of India (SEBI) regulations stipulate (articulated in Clause 49 of the Listing Agreement) that the composition of the Board of Directors, remuneration of directors, and other information be disclosed by all listed companies. Prowess provides information on the composition of the board of directors – members' names, designations, the number of positions held by the directors in other companies, and the remuneration of each director. Prowess also classifies directors as independent and non-independent and separately as executive and non-executive directors. It also includes dates on which a director was appointed, retired, or resigned. Over the sample period, we have a total of 31,833 firms, of which 22,000 firms have board information available.

To identify the religions and castes of company 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. We retain only those firm-year observations for which we can (probabilistically)

[13] As of 31 March, 2009 7,86,774 companies registered with the Registrar of Companies, an administrative arm of the Ministry of Company 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.

Table 3: Features of Probabilistic Cultural Mapping of Corporate Directors

Number of cultural classifications	% last names with associated classifications	Average probability associated with top classifications
	Religion	
1	68.6	95.3
2	92.2	99.5
3	98.0	99.9
4	99.6	100
5	100	100
	Varna	
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
	Jati	
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

identify the religion/caste of all board members.[14] We also consider only those firm-years that have at least two directors serving on their boards.[15] In our matched sample, we have 20,209 unique firms with a total of 631,221 directorships. [16]

Table 3 profiles the religion and caste mapping for directors. For religion, we see that 68.6% of directors' last names are associated with a single religion, and 99.5% names are associated with up to two religions. As expected, the mapping for caste is noisier,

[14]There are also a few directors whose names suggest that they are not Indian nationals. We are unable to match these names with those in the matrimonial database. These directors are simply assigned a religion category of "Unknown" and a *varna* category of "NA."

[15]This drops a total of 1,635 firm-year observations, covering 1,199 firms.

[16]As a percentage of firms and directors in the Prowess database over 1999-2012, we are able to match 73.7% of firm-year observations and 79.6% of all directorships. We are unable to identify unique individuals serving as directors since we do not have unique director identification numbers assigned to them by the Ministry of Company Affairs. Since two individuals can have the same name, we do not rely on names to identify unique directors. This is especially the case in our data as we only have the initials instead of first names for many directors.

Table 4: Religion and Varna Composition of Corporate Directors

	1999		
Religion	Percent of Last Names	Varna	Percent of Last Names
Hindu	82.31	Brahmin	20.56
Muslim	1.98	Kshatriya	14.61
Christian	2.83	Vaishya	25.01
Jain	8.50	Shudra	16.83
Sikh	0.69	Dalit	0.89
Zoroastrian	1.05	Others	3.02
Unknown	2.68	NA	19.07
	2012		
Hindu	83.89	Brahmin	22.25
Muslim	2.12	Kshatriya	15.13
Christian	1.67	Vaishya	25.27
Jain	9.22	Shudra	16.44
Sikh	0.74	Dalit	0.79
Zoroastrian	0.56	Others	2.55
Buddhist	0.01	NA	17.57
Unknown	1.80	–	–

with 45% (40%) last names associated with a single *varna*(*jati*) and 85% (65.5%) last names associated with up to five *varnas*(*jatis*). Although the mapping is probabilistic, the probabilities are front loaded. Thus, the top two most likely religions account for 99.5% of the total likelihood, on average. Similarly, the top five most likely *varnas*(*jatis*) account for 99.3% and 89% of the total likelihood, on average. Table 4 provides the composition of directors for the beginning and end years (1999 and 2012) of our sample, as measured according to their most likely religion and *varna*.[17]

We also use information on several firm characteristics, including sales, total assets, annual profits, export status, industry, location of headquarters, etc. A few key characteristics of firms in our sample are provided in Table 4.

Homophily Index:

[17]For space considerations, we do not report the *jati* composition of directors but the information is available upon request.

Table 5: Summary Statistics

% Firms in Sector			Firm Characteristics			Board Characteristics		
Sector	1999	2012	Firm Characteristics	1999	2012	Board Characteristics	1999	2012
Mining, Construction & Utilities	6.25	8.49	Total firms	2048	4195	% with CEO Duality	14.45	19.05
Manufacturing	44.82	30.89	% Public	93.7	79.81	Mean % Independent Directors	8.89	3.3
Trade	11.87	11.7	% Exporters	37.6	29.61	Mean Board Size	5.6	5.78
Transport, Accommodation & Real Estate	3.61	4.22	Median Sales (Rupees millions)	2.56	2.04			
Information & Communication	5.18	11.32	Median Assets (Rupees millions)	2.6	2.03			
Finance & Insurance	20.8	22.12	Median Profits (Rupees millions)	0.19	0.09			
Professional, Technical & Administrative Services	3.66	4.08						
Education & Health	0.59	0.95						
Arts, Recreation & Others	1.22	5.01						
Diversified	2	1.22						

To measure the degree of cultural concentration or homophily (based on religion or caste), we develop a novel homophily index, the fuzzy Blau Index. The index is based on the Gini-Simpson index which is also known as the Blau or Hirschman-Herfindahl Index (Hirschman (1945), Herfindahl (1950)).[18]

Consider a board j of size N with k unique social categories for every board member.

[18]Diversity indices are popular measures used by ecologists, economists, and information scientists to measure distribution of a population over different types or subgroups. The two most widely used measures of diversity of a group are theSimpson Index and Shannon Index.

Simpson Index (Simpson (1949)):

$$I_S = \sum_{i=1}^{n} p_i^2$$

Each board member l has a probability distribution over k unique social categories.

$$p_l = (p_{lm})_{m=1}^{k}$$

p_{lm} is the probability that board member l is of social category m. A combination of a social category in a board is a count of directors that belong to different social catgeories. We denote a combination of a social category by $c = (n_m)_{m=1}^{k}$, where n_m is the number of directors that belong to category m. A board of size N and k unique social categories has N^k such distinct possible combinations of categories over its board members. Fuzzy Blau Index for a board j is defined as:

$$FB_j = \sum_{c=1}^{N^k} p_c B_c$$

where, p_c is the probability of cth combination of category in a board. We assume that the probability distribution of any board member is independent of that of others. Thus,

$$p_c = \Pi_{l=1}^{N} \Pi_{m=1}^{k} p_{lm}$$

B_c is the Blau Index of cth combination of category over board members defined as:

$$B_c = \sum_{m=1}^{k} (n_m/N)^2$$

2.3 Indian Boards Database and Other Data Sources

Unique identification of individuals serving as directors is not possible in Prowess. This prevents us from examining the degree and nature of interlocks across boards, i.e., which individuals hold multiple directorships and the firms these positions are in. To do this, we instead use the Indian Boards Database, a database maintained by the Prime database group, which provides a unique identification code for each individual serving as a director for about a thousand firms during 2006-2015, along with demographic information such as age, gender, nationality, educational qualifications, experience. Additional information

Shannon Index (Shannon (1948)):

$$I_{Sh} = (-) \sum_{i=1}^{n} p_i ln(p_i)$$

where n is number of types (or subgroups) of a population and p_i is relative share of type (or subgroup) i in the population.

Shannon index and Simpson index are proven to have the most desirable properties of all diversity indices (Routledge (1979)).

on their directorial position is also available, including independent/non-independent status, remuneration, date of appointment, cessation date, and reason of cessation. The unique identification code for directors and information about all the boards they serve on allows us to measure the degree of interlocks. We use this information to examine the relevance of our instrumental variable strategy. The information on other board membership in Indian Boards Database is only available starting 2012. Since the time period of our study is 1999-2012, we only examine the data for 2012 to calculate different measures of board interlocks within the same broad industries. For the 2012 sample, we have information on 13,453 unique directors across 1,352 firms.

We use two other data sources for state GDP and sectoral output. Data on real per capital state domestic product are obtained from the *Handbook of Statistics on the Indian Economy* published by Reserve Bank of India for the years 1999-2012. The annual report of the Ministry of Agriculture is used to obtain data on the real domestic product for broad sectors of the economy for the years 2004-2012. All nominal data are deflated by the all-India CPI (2001=100).

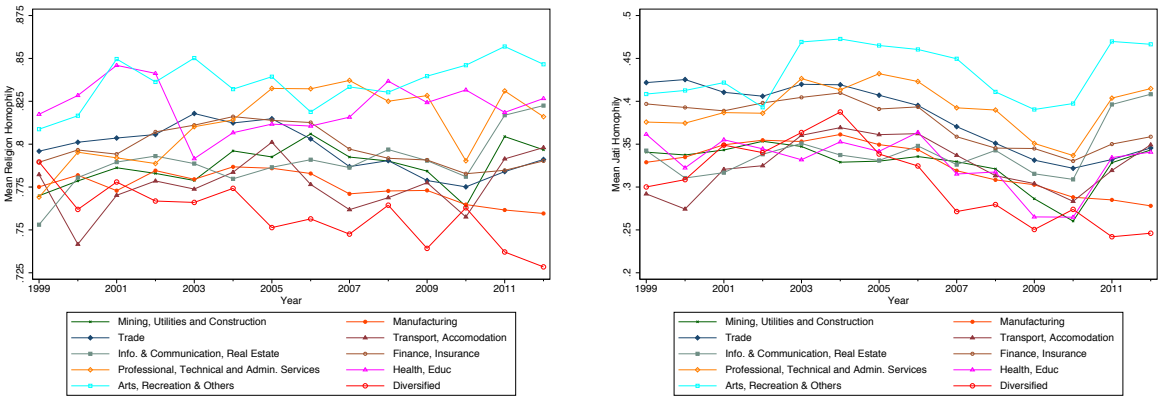
3 Patterns of Cultural Homophily in India's Corporate Boards

In this section, we document a rich set of trends in boards' religion and caste composition over time, across sectors and firms, and across states.

3.1 Homophily is High and Persistent

The average fuzzy Blau index in Indian firms is high for all measures of culture, at 0.8 for religion, 0.49 for *varna* and 0.37 for *jati*. To assess how high these averages are, we take a hundred random sub-samples of six directors (since the average board size is six), and calculate the average religion/ *varna*/*jati* homophilies across these samples. These averages constitute our benchmark levels of cultural homophily in corporate boards. These benchmark values are 0.73 for religion, 0.33 for *varna* and 0.2 for *jati*. Thus, the actual cultural homophily, for all three groupings, is considerably higher than the benchmark.

Further, homophily is persistent over time overall and across sectors, geography, and groups of firms with shared characteristics. Figure 2 shows that average religion and *jati* homophilies on boards are high and persistent in all broad sectors of the economy. Across all years and sectors, the average religion (*jati*) homophily varies in the range of 0.73-0.85



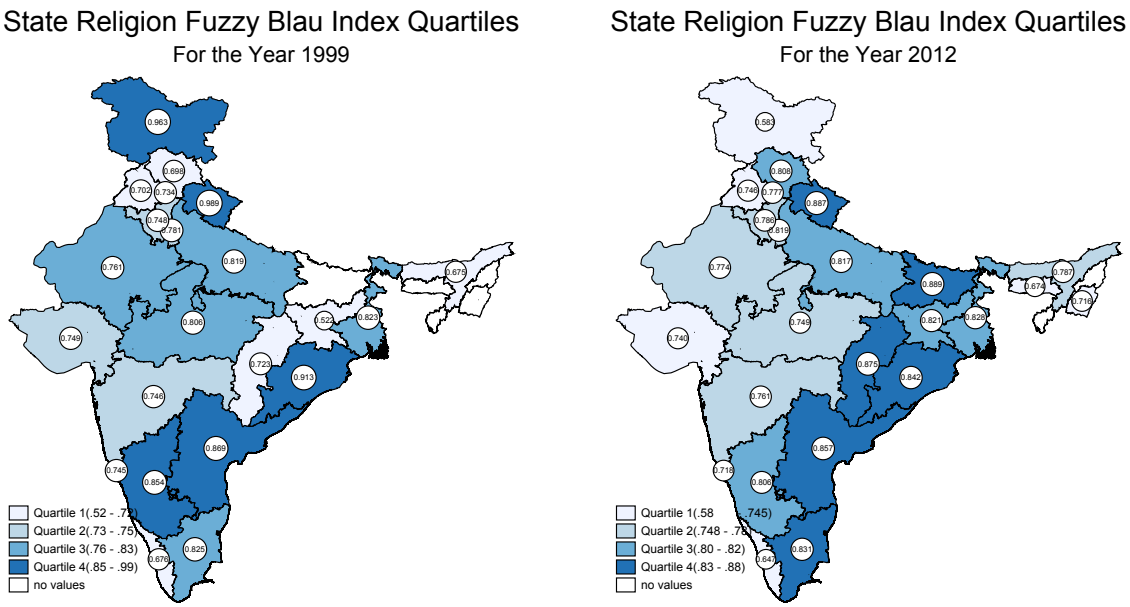
(a) (b)

Figure 2: Average Cultural Homophily in Broad Sectors[a]

[a]Source: Prowess, authors' mapping of last names to religion and *jati* using matrimonial data. Figure 1(a) (1(b)) shows the proportion of directors that belong to the religion (*jati*) with the highest representation on a board, as well as the overall average across boards, for each religion (*jati*). Both figures use data for the year 1999.

(0.25-0.47). The sectors of professional, technical, and administrative services, and arts and recreation have the highest homophily levels, while the manufacturing, diversified, and transport and accomodation sectors have the least homophily. There is no uniform secular trend in homophily across sectors; while some sectors witness a small decline, others see a modest rise.

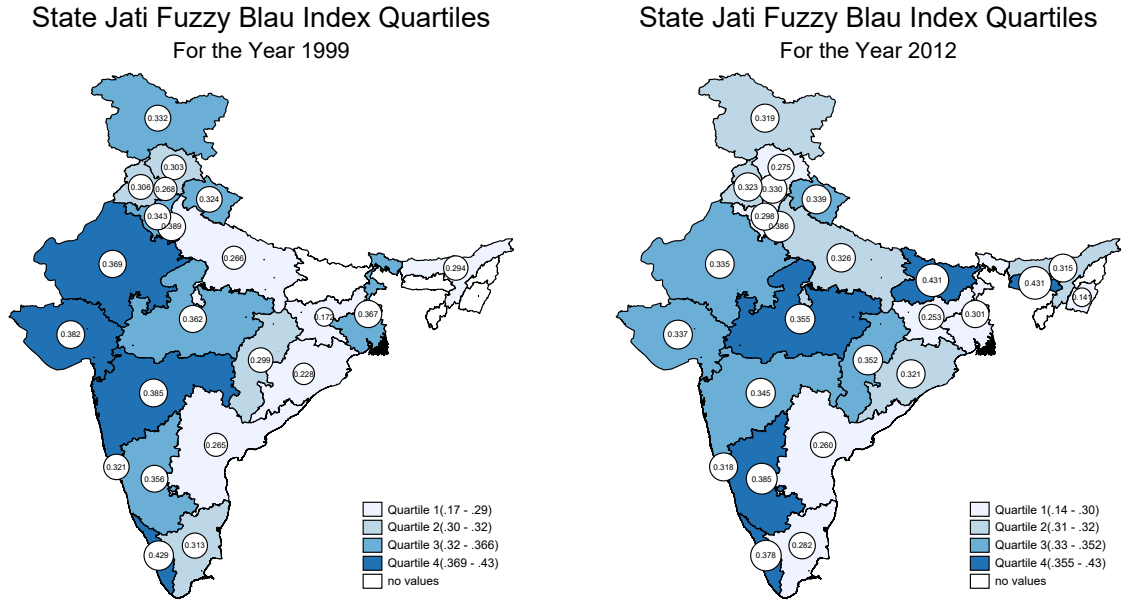
Figures 3 and 4 show that while the average homophily stayed high in most states over the sample period, states did change their relative positions in the overall distribution.



(a) (b)

Figure 3: Average Religion Homophily Across States[a]

[a]Source: Prowess, authors' mapping of last names to religion using matrimonial data.



(a) (b)

Figure 4: Average Jati Homophily Across States[a]

[a]Source: Prowess, authors' mapping of last names to jati using matrimonial data.

3.2 Homophily is Lower in Higher Quality Firms

Next, we document secular trends in homophily across different firm groups. Figures 5 and 6 show that larger firms, as measured by real assets, sales, and profits, respectively for panels (a), (b), and (c), have more diverse boards. The three panels in Figures 5 and 6 show the evolution over time of average religion and *jati* homophily in firms of different assets, sales, and profits quartiles. We see a systematic pattern of lower homophily as we move from lower to higher quartiles of assets and sales. For profits, there is considerable overlap in average homophily between the first and second quatiles. However, both have a higher homophily than the average of firms in the third quartile which, in turn, have higher homophily on average than firms in the fourth 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. We also find that within each year, the lower the firm's standardized homophily (measured as the standard deviation distance from the mean homophily), the higher its position in the standardized distributions of sales, assets, and profits. In all years, these correlations are negative and statistically significant at the 1% level.

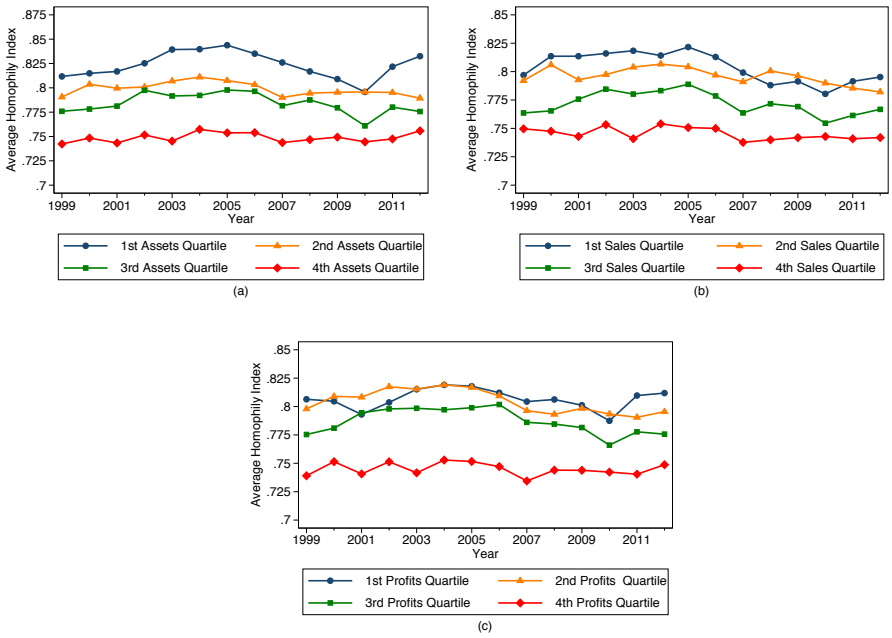


Figure 5: Average Religion Homophily by Firm Size[a]

[a]Notes: Source: Prowess, authors' last name to religion mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.

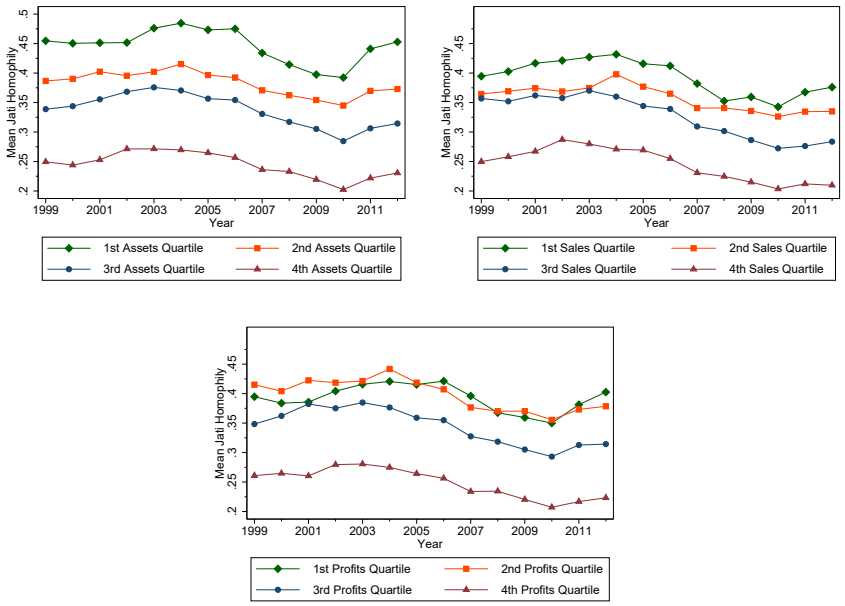
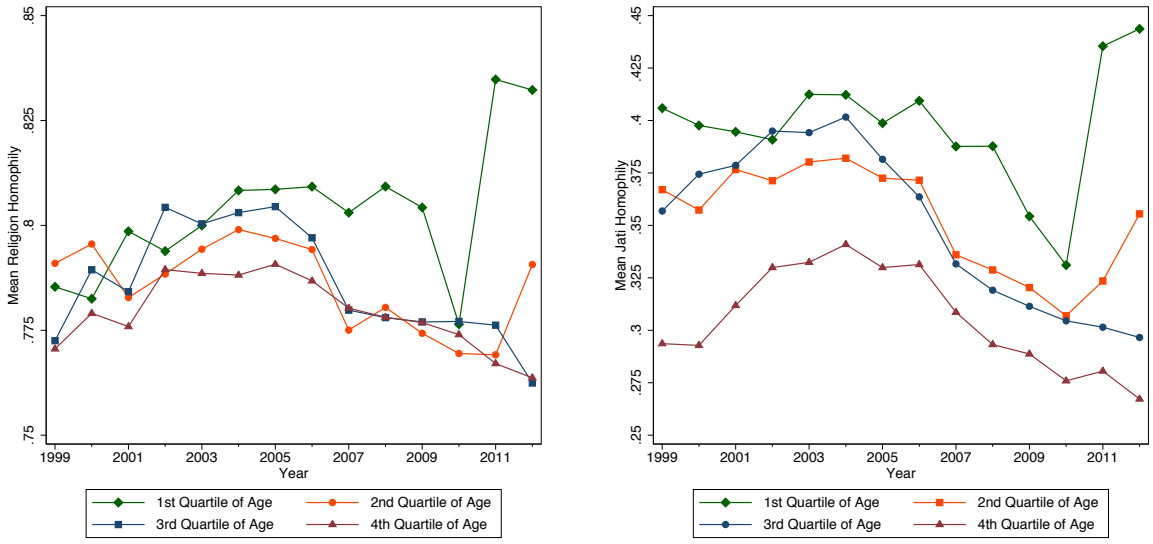


Figure 6: Average Jati Homophily by Firm Size[a]

[a]Notes: Source: Prowess, authors' last name to jati mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.

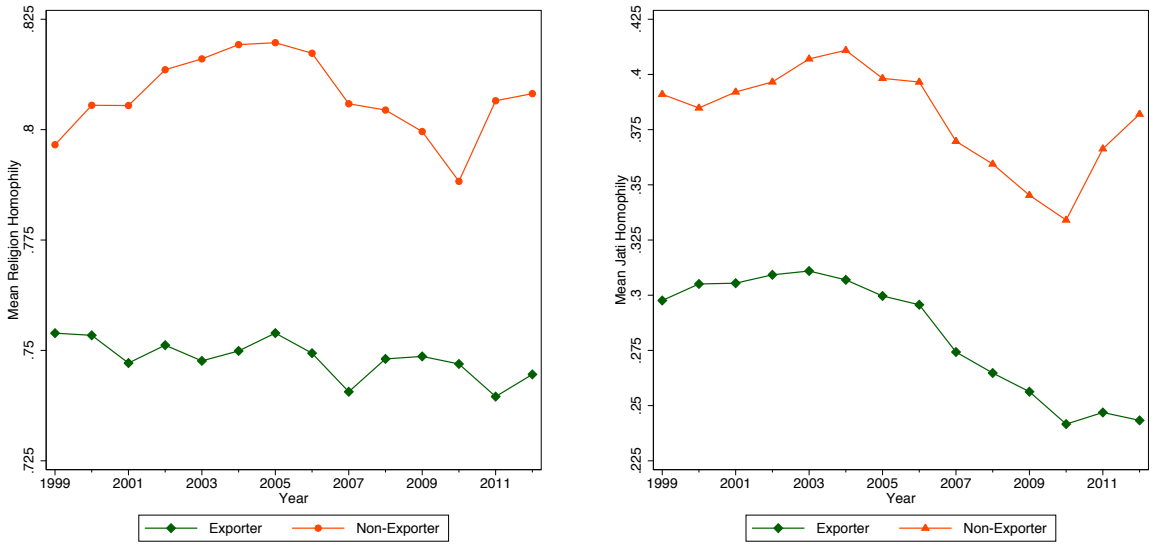
Figure 7 demonstrates that older firms have board members from more diverse cultural backgrounds. As we go from lower to higher quartiles of firm age (measured s years since incorporation), the average homophily falls. Moreover, firms in all age groups, except the youngest, witnessed a small decline in their boards' religion and *jati* homophily. Figure 8 shows that exporting firms have significantly more diverse boards, on average, than non-exporting firms. Moreover, while exporters saw a slight increase in their board diversity, non-exporters' board diversity fluctuated over the years. We see little difference in average board homophily across other firm characteristics such as public versus private firms, government versus non-government firms, and between firms owned by business groups versus others.



(a) (b)

Figure 7: Average Cultural Homophily by Firm age[a]

[a]Notes: Source: Prowess, authors' last name to religion*(jati)* mapping using matrimonial data. Homophily is averaged over all firms in a given age quartile.



(a) (b)

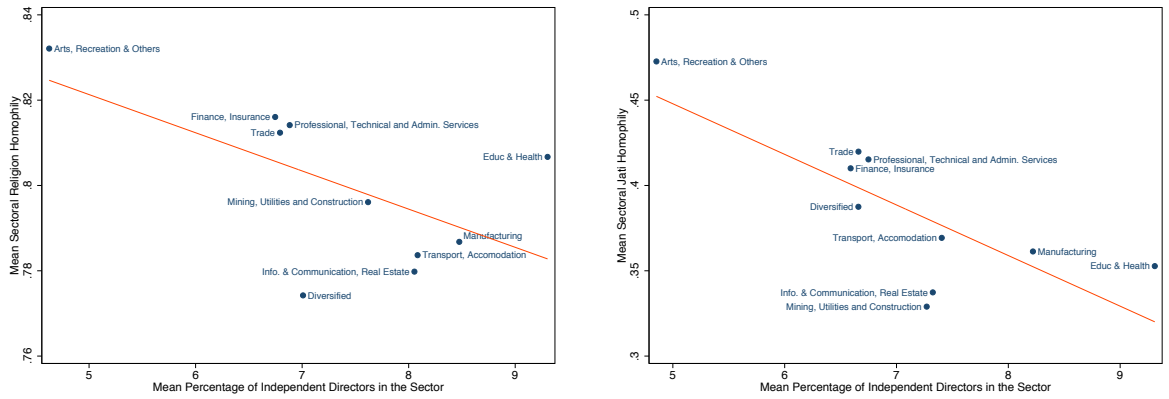
Figure 8: Average Cultural Homophily by Exporting Status[a]

[a]Notes: Source: Prowess, authors' last name to religion*(jati)* mapping using matrimonial data. Homophily is averaged over all exporting and non-exporting firms separately.

3.3 Better Corporate Governance Accompanies Lower Homophily

We also find that boards that display lower levels of cultural homophily also have features indicative of superior corporate governance. Figures 9, 10, and 11 show the association between religion and *jati* homophily and three features of corporate governance: proportion of independent directors on the board, CEO duality, and size of the board. Figure 9 shows the association between average board cultural homophily and the average proportion of independent directors across firms in broad one-digit sectors. Panels (a) and (b) show remarkable negative correlation between mean religion and *jati* homophily and the proportion of independent directors across sectors. Sectors with the highest homophily levels such as arts and recreation and finance and insurance also have the lowest proportion of independent directors in their corporate boards. The figures also show that all sectors have significantly lower percentages of independent directors compared to international standards.

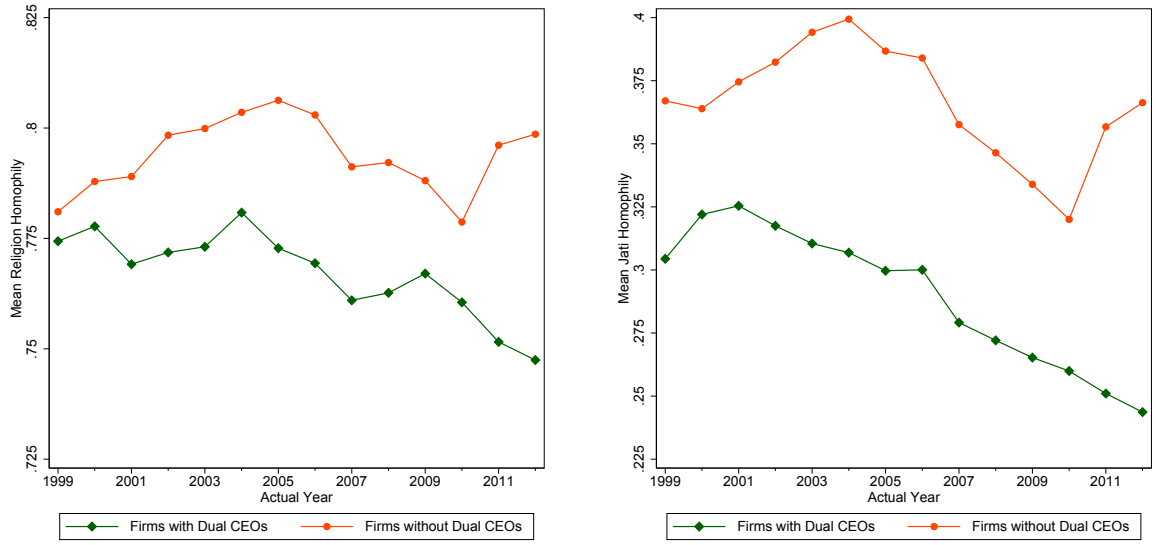
CEO duality refers to situations where the CEO of a firm also serves as the chair of the board of directors. This clearly constitutes poor governance practice since agency problems of the CEO cannot be effectively monitored by the directors. Figure 10 shows that firms with dual CEOs have significantly higher cultural homophily than those without dual CEOs. Finally, Figure 11 presents average cultural homophily for firms with different board sizes, grouped into four quartiles. We take larger board sizes as indicative of better governance. Yet again, we see that firms with larger corporate boards have lower cultural homophily among their directors.



(a) (b)

Figure 9: Cultural Homophily and Percentage of Independent Directors[a]

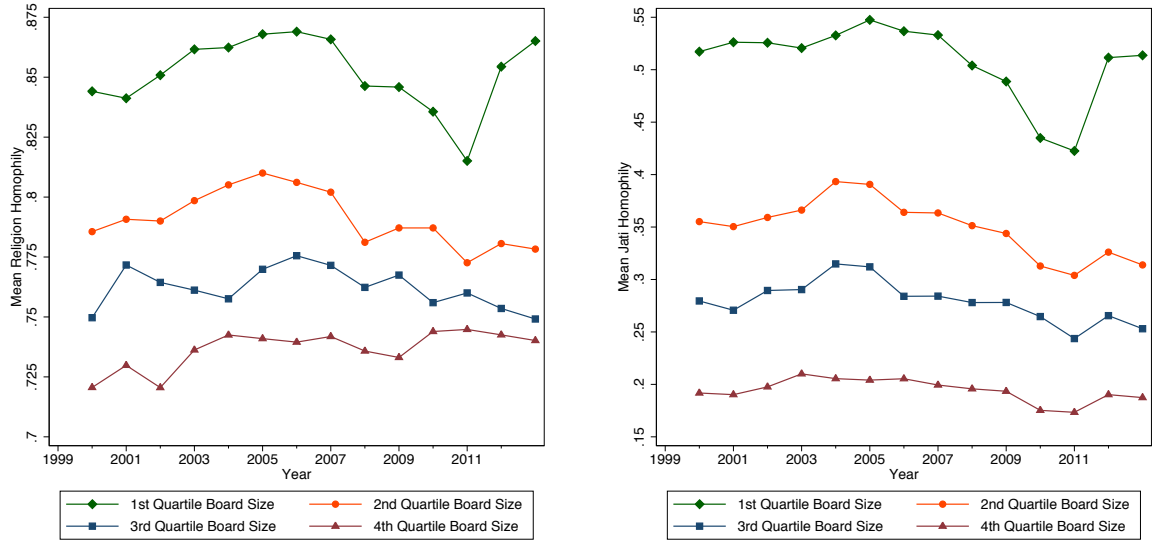
[a]Notes: Source: Prowess, authors' last name to religion*(jati)* mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.



(a) (b)

Figure 10: Average Cultural Homophily by CEO Duality[a]

[a]Notes: Source: Prowess, authors' last name to religion*(jati)* mapping using matrimonial data. Homophily is averaged over all firms in a given age quartile.



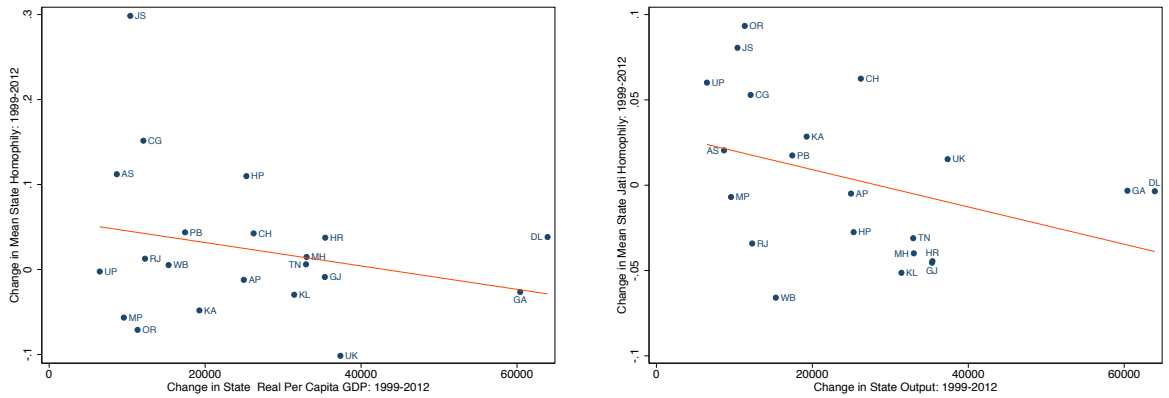
(a) (b)

Figure 11: Average Cultural Homophily by Board Size[a]

[a]Notes: Source: Prowess, authors' last name to religion*(jati)* mapping using matrimonial data. Homophily is averaged over all exporting and non-exporting firms separately.

3.4 Homophily Negatively Associated With Aggregate Economic Outcomes

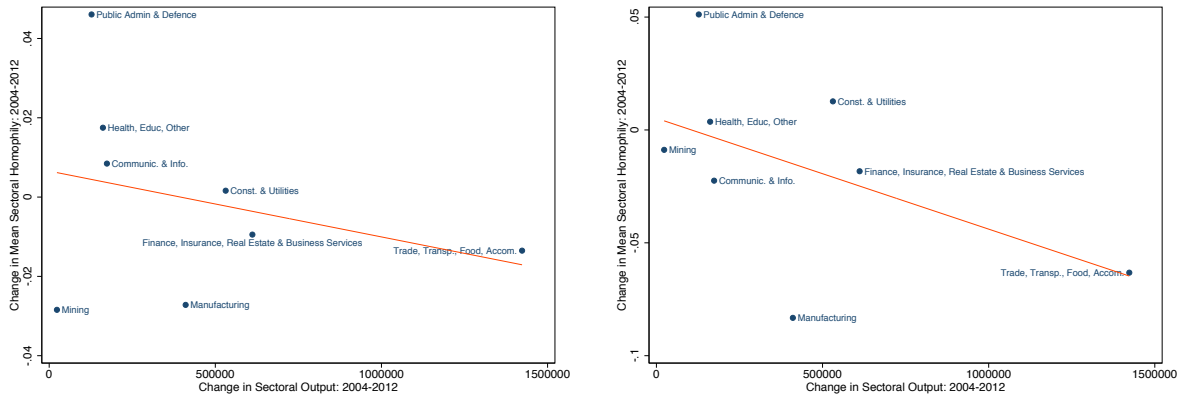
Remarkably, we find that cultural homophily of boards of directors is also associated strongly with other aggregate economic indicators. Figures 12 and 13 show that an increases in sectoral and state output are associated with declines in the religion and *(*(jati)) homophily of the sectors and states showing clearly that an increase in diversity over time leads to positive economic outcomes. In figure 14 we see that sectors with higher proportion of older employees have higher religion and jati homophily, this means that a diverse number of cultural groups started having better employement opportunities over time.



(a) (b)

Figure 12: Cultural Homophily and State Output[a]

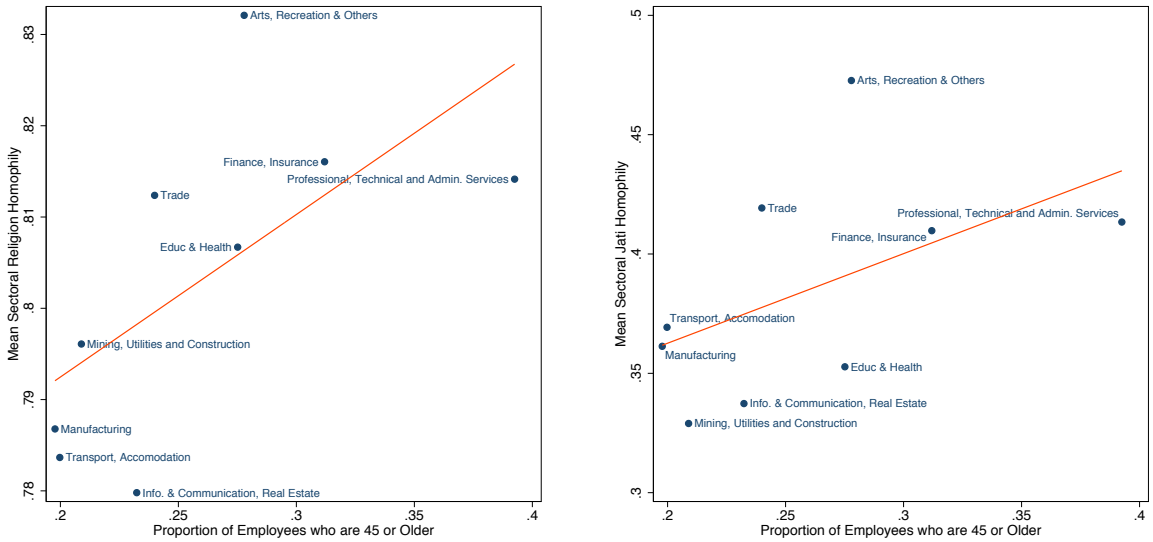
[a]Notes: Source: Prowess, authors' last name to religion*(jati)* mapping using matrimonial data. Homophily is averaged over all firms in a state in each year.



(a) (b)

Figure 13: Cultural Homophily and Sectoral Output[a]

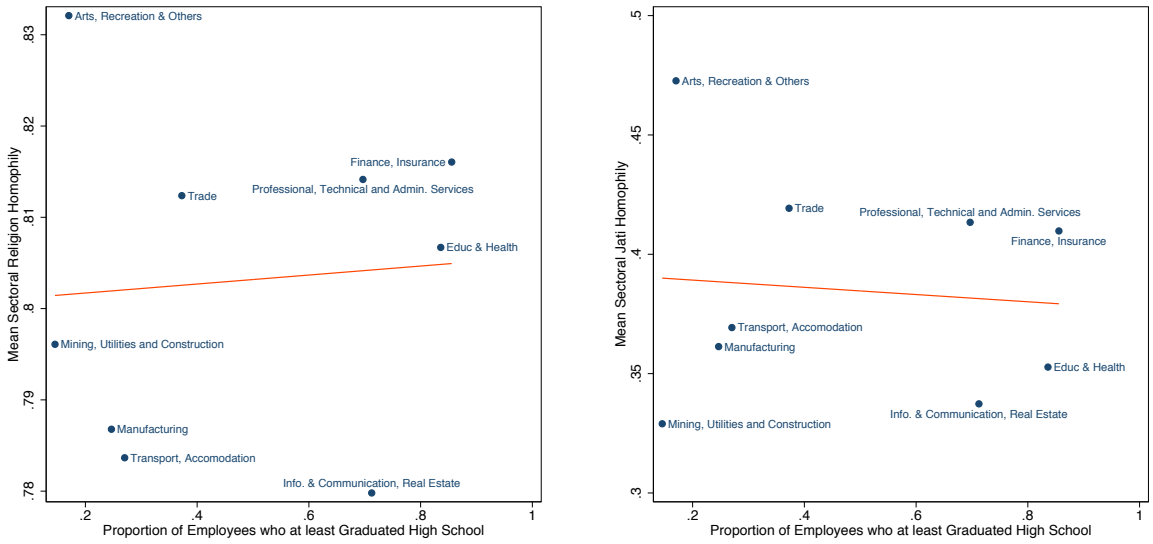
[a]Notes: Source: Prowess, authors' last name to religion*(jati)* mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.



(a) (b)

Figure 14: Cultural Homophily and Worker Age[a]

[a]Notes: Source: Prowess, authors' last name to religion*(jati)* mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.



(a) (b)

Figure 15: Cultural Homophily and Worker Education[a]

[a]Notes: Source: Prowess, authors' last name to religion*(jati)* mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.

Finally, we show in Table 6 that not all firms are dominated by Hindus, even though they are the overwhelming majority of directors. Indeed, in both the beginning and end years of the sample, a non-negligible proportion of firms are dominated by directors belonging to other religions. Further, a large fraction of directors in these firms belong to the dominant religion. For example, in both years, Muslims constituted about 65% of the board in Muslim-dominated firms, while Jains constituted 71% and 65% of the board in firms that were dominated by them in 1999 and 2012, respectively. We also observe that firms dominated by non-Hindu religions are characterized by relatively lower levels of homophily. This is both because non-Hindus account for smaller shares of boards dominated by them and because there is greater religion diversity among the remaining directors in these boards compared to Hindu dominated boards. Overall, this table shows that the high observed religion homophily seen in Indian firms is not entirely driven mechanically by most directors (and most of India's population) being Hindu. Instead, directors belonging to other religious groups also appear to prefer serving on boards that have high percentages of directors of their own religions. Hence, the desire for cultural proximity is pervasive across all religious groups.

Table 6: Dominant Religions in Firm Boards

Dominant Religion	Percent of Firms	Average Share of Directors of Dominant Religion	Average Homophily
		1999	
Hindu	86.38	0.91	0.81
Muslim	1.32	0.66	0.61
Christian	1.76	0.65	0.61
Jain	8.01	0.71	0.60
Sikh	0.2	0.66	0.59
Zoroastrian	0.78	0.56	0.48
Unknown	1.56	0.55	0.51
		2012	
Hindu	88.84	0.90	0.81
Muslim	1.36	0.65	0.68
Christian	0.86	0.57	0.57
Jain	7.65	0.65	0.59
Sikh	0.24	0.57	0.58
Zoroastrian	0.36	0.51	0.53
Unknown	0.69	0.62	0.59

4 Empirical Strategy

Thus far, we have seen that although religion concentration in corporate boards has remained high throughout the sample period, it systematically varies across firms such that it is lower in larger and higher performing firms, sectors, and states. Next, we examine whether this systematic pattern is causal. In particular, we investigate whether and to what extent higher religion and caste concentration on firm boards negatively impact firm performance along several dimensions. In this section, we describe our empirical strategy for this analysis.

Consider the following regression equation:

$$P_{it} = \beta_0 + \beta_1 H_{it} + \beta_2 X_{it} + \beta_3 B_{it} + \delta_1 I_j + \delta_2 S_k + \delta_3 T_t + \epsilon_{it} \tag{4.1}$$

where P_{it} denotes firm performance, measured by return on assets (ROA), operating cash flow (OCF), and firm value (market to book ratio) of firm i in year t, H_{it} is religion, varna, or jati 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 board characteristics, I_j denotes a vector of industry fixed effects, S_k is a vector of state fixed effects, and T_t is a vector of year fixed effects. We cluster the standard errors by year and industry and correct them for arbitrary heteroskedasticity.

As discussed in the introduction and section 3, we can rule out the presence of homophily in firm boards as caused by coincidence or by members of certain cultural groups being systematically more talented than others. This leaves us with two potential underlying mechanisms – either incumbent directors rely on their social networks to hire a new board member, and their networks tend to be homophilous, or incumbents exercise favoritism and systematically prefer hiring members of their own communities than of others even if the latter may be more suited to the job. In the first scenario, the effect of the board's cultural homophily on firm performance can be positive if it improves corporate governance by increasing trust and coordination among directors and negative if it prevents directors from expressing their independent views. In the second scenario, high cultural homophily is likely to be detrimental to firm performance. Since we cannot rule out either of these mechanisms, we interpret the regression results as reflecting the net impact of board cultural homophily on firm outcomes via both mechanisms. In the above regression, this impact is captured by β_1. 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 a firm performance measure such as return on assets. Moreover, 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 the board composition and, hence, homophily.

To overcome this endogeneity, we employ an instrumental variable strategy, and use the following four variables as instruments: (1) the religion/caste concentration among all directors in the industry that the firm belongs to, (2) the religion/caste concentration among all directors in the state where the firm is located, (3) the Euclidean distance of its board's religion/caste composition from that of the full set of directors in the corresponding industry, and (4) the Euclidean distance of its board's religion/caste composition from that of the full set of directors in the corresponding state. We measure industry at the 3 digit level following the National Industrial Classification (2008).[19]

[19] A more disaggregated classification is unsuitable for two reasons. First, directors may not serve on closely competing firms' boards due to conflicts 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

Further, in calculating the homophily indices for the firm, state, or industry, we only consider the most likely religion/caste of each director. Therefore, this homophily index is simply the Blau or the Hirschman-Herfindahl index (HHI). The reason for this choice is that it is computationally extremely time consuming to calculate the fuzzy Blau index taking all possible religions or castes of all directors serving in an industry or state.[20]

The intuition for using these measures is that a firm's corporate culture may be similar to that of other firms in the same industry or geographical area. This suggests that if the group of all directors in the industry/region displays high levels of cultural homophily, the firms in that industry/region may also be more likely to have highly concentrated boards. Moreover, directors may be chosen from its local geographical region and/or the broad industry it belongs to. In other words, directors serving on other firms in the same industry or in the same region provide the supply pool of potential directors for a firm. 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 broad industry and (b) the religion and caste composition of directors on firm boards is very similar to that in the entire industry (even if directors on these boards have few or no members with additional directorships in the same industry) or in the entire region (identified by state). Table 4 documents within-industry board interlocks for one-digit industries for the year 2012. 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 as a director on a firm that 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. This gives us a firm's degree of within-industry interlock. Key moments of these interlocks for firms in each 1-digit industry in the year 2012 are presented in Table 4. We observe that the average interlock ranges from 1% to 30% across these broad industries. But the maximum degree of interlocks can be as high as 83%. Looking at two-digit and three digit industries, we see that even at these narrower levels, there are interlocks, albeit to a smaller degree. The mean interlock in 2012 is 5.2% (3.4% in three-digit), although the maximum interlock is over 80% in many industries.

disaggregated classification level, on the other hand, is undesirable as it will not yield enough variation in the industry level homophily index.

[20]In some states and industries, calculating the fuzzy Blau index for all directors while taking all their possible religions/castes into account can take over one year.

Table 7: Within-Industry Board Interlocks

NIC	Within-Industry Board Interlocks			
	Mean	Median	Minimum	Maximum
1	0.16	0.10	0	1
2	0.30	0.25	0	1
3	0.15	0.10	0	0.83
4	0.12	0.03	0	1
5	0.18	0.11	0	0.75
6	0.25	0.20	0	1
7	0.01	0	0	0.17
8	0.02	0	0	0.33
9	0.01	0	0	0.11

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 religious composition of directors is similar to that of directors in the entire industry. We show this by comparing the distribution of directors in firms that have below (and above) median interlocks to that of the entire industry using the Kolmogorov-Smirnov test (K-S test).[21] The test statistic requires sample sizes of the two samples that it compares. Thus, it does not allow us to compare the distribution of the top religion of directors of an average firm with that of the entire industry since the sample size of directors for this distribution is not defined. Thus, we compare the distribution of the top religion of unique directors in the set of firms that have below-median interlocks to that of the entire industry. We repeat this for firms with above-median interlocks. To look at this differently, we peform another K-S test where the distribution of firms' dominant religion for the set of below (and above) median firms is compared with that of firms in the entire industry. Results from these tests are presented in Table 5. The table shows that for each 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 religion.

Table 6 presents results from K-S test that compares the distribution of top religions of directors in each of the four size (sales) quartiles of firms in a state to that in all firms in that state in the year 2012. It also presents results from the analogous test

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

Table 8: K-S Test Results Comparing Religion Distributions of Directors in Firms to Industries

	NIC	Above Median			Below Median		
		Test Stat	Critical Value	Decision	Test Stat	Critical Value	Decision
Directors	1	0.02	0.07	Not Rejected	0.02	0.06	Not Rejected
	2	0.01	0.04	Not Rejected	0.01	0.04	Not Rejected
	3	0.02	0.08	Not Rejected	0.018	0.08	Not Rejected
	4	0.02	0.07	Not Rejected	0.02	0.08	Not Rejected
	5	0.03	0.11	Not Rejected	0.02	0.11	Not Rejected
	6	0.01	0.06	Not Rejected	0.01	0.06	Not Rejected
	7	0.13	0.57	Not Rejected	0.01	0.19	Not Rejected
	8	0.22	0.80	Not Rejected	0.01	0.18	Not Rejected
	9	0.11	0.48	Not Rejected	0.02	0.22	Not Rejected
Firms	1	0.03	0.18	Not Rejected	0.03	0.18	Not Rejected
	2	0.03	0.11	Not Rejected	0.03	0.11	Not Rejected
	3	0.04	0.22	Not Rejected	0.04	0.22	Not Rejected
	4	0.03	0.21	Not Rejected	0.03	0.21	Not Rejected
	5	0.07	0.31	Not Rejected	0.03	0.31	Not Rejected
	6	0.03	0.16	Not Rejected	0.03	0.15	Not Rejected
	7	0.08	1.41	Not Rejected	0.01	0.54	Not Rejected
	8	0.13	1.41	Not Rejected	0.01	0.51	Not Rejected
	9	0.08	1.42	Not Rejected	0.01	0.57	Not Rejected

that compares firms' dominant religion distribution to that of the entire state. For this case, however, we use Prowess instead of the Indian Boards database. The reason is that the small sample of firms in the latter database means that in several states we have very few firms to make any meaningful comparisons. Using Prowess also means that we cannot identify unique directors to measure interlocks. Table 6 presents these results for the top ten states with the most number of firms in the 2012 sample. We observe that in every state, the religious composition of directors in firms in all four sales quartiles is statistically indistinguishable from that of the set of directors across all firms in that state. The same holds true of the dominant religious composition of firms. Results are similar for other states and years.

Table 9: K-S Test Results Comparing Religion Distributions of Directors in Firms to States

	State	1st Quartile			2nd Quartile			3rd Quartile			4th Quartile		
		Test Stat	Critical Value	Decision	Test Stat	Critical Value	Decision	Test Stat	Critical Value	Decision	Test Stat	Critical Value	Decision
Directors	Maharashtra	0.02	0.05	Not Rejected	0.01	0.04	Not Rejected	0.01	0.04	Not Rejected	0.02	0.04	Not Rejected
	NCT of Delhi	0.01	0.08	Not Rejected	0.02	0.07	Not Rejected	0.04	0.07	Not Rejected	0.02	0.06	Not Rejected
	Gujarat	0.06	0.08	Not Rejected	0.04	0.08	Not Rejected	0.02	0.07	Not Rejected	0.03	0.06	Not Rejected
	West Bengal	0.01	0.09	Not Rejected	0.02	0.08	Not Rejected	0.01	0.08	Not Rejected	0.03	0.07	Not Rejected
	Tamil Nadu	0.07	0.09	Not Rejected	0.03	0.08	Not Rejected	0.03	0.08	Not Rejected	0.03	0.07	Not Rejected
	Andhra Pradesh	0.01	0.09	Not Rejected	0.01	0.09	Not Rejected	0.02	0.08	Not Rejected	0.02	0.07	Not Rejected
	Karnataka	0.05	0.13	Not Rejected	0.06	0.12	Not Rejected	0.04	0.12	Not Rejected	0.02	0.11	Not Rejected
	Uttar Pradesh	0.07	0.16	Not Rejected	0.02	0.13	Not Rejected	0.03	0.13	Not Rejected	0.01	0.11	Not Rejected
	Haryana	0.15	0.19	Not Rejected	0.03	0.14	Not Rejected	0.03	0.14	Not Rejected	0.05	0.13	Not Rejected
	Punjab	0.01	0.18	Not Rejected	0.05	0.16	Not Rejected	0.02	0.14	Not Rejected	0.02	0.13	Not Rejected
Firms	Maharashtra	0.03	0.10	Not Rejected	0.03	0.10	Not Rejected	0.01	0.10	Not Rejected	0.05	0.10	Not Rejected
	NCT of Delhi	0.02	0.17	Not Rejected	0.01	0.17	Not Rejected	0.01	0.17	Not Rejected	0.01	0.17	Not Rejected
	Gujarat	0.11	0.18	Not Rejected	0.03	0.18	Not Rejected	0.00	0.18	Not Rejected	0.08	0.18	Not Rejected
	West Bengal	0.05	0.21	Not Rejected	0.02	0.21	Not Rejected	0.01	0.21	Not Rejected	0.02	0.21	Not Rejected
	Tamil Nadu	0.05	0.20	Not Rejected	0.05	0.20	Not Rejected	0.04	0.20	Not Rejected	0.06	0.20	Not Rejected
	Andhra Pradesh	0.01	0.21	Not Rejected	0.01	0.22	Not Rejected	0.05	0.22	Not Rejected	0.06	0.22	Not Rejected
	Karnataka	0.08	0.29	Not Rejected	0.05	0.29	Not Rejected	0.03	0.29	Not Rejected	0.06	0.29	Not Rejected
	Uttar Pradesh	0.04	0.34	Not Rejected	0.01	0.35	Not Rejected	0.01	0.35	Not Rejected	0.04	0.35	Not Rejected
	Haryana	0.09	0.39	Not Rejected	0.02	0.39	Not Rejected	0.05	0.39	Not Rejected	0.05	0.40	Not Rejected
	Punjab	0.02	0.38	Not Rejected	0.09	0.39	Not Rejected	0.02	0.39	Not Rejected	0.12	0.39	Not Rejected

Table 10: Panel A: Distance Between Firm and Industry/State Director Composition

Year	10th Percentile	50th Percentile	90th Percentile	Mean	Standard Deviation
	Distance Between Firm and Industry Director Composition				
1999	0.09	0.19	0.53	0.26	0.24
2004	0.11	0.19	0.59	0.27	0.26
2009	0.09	0.19	0.50	0.24	0.22
2012	0.09	0.18	0.51	0.24	0.21
	Panel B: Distance Between Firm and State Director Composition				
1999	0.09	0.17	0.58	0.26	0.25
2004	0.08	0.18	0.59	0.26	0.25
2009	0.07	0.16	0.48	0.23	0.22
2012	0.08	0.19	0.48	0.23	0.21

We additionally use the distance between the firm and industry/region with regard to their cultural composition. Note that several different religion/caste compositions can yield the same homophily index. So whether a firm is representative of its industry/region can be determined not only by comparing its overall homophily index with that of the industry/region but also its underlying religion/caste composition. The larger this distance, the less representative the firm is of 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 7 demonstrates that Euclidean distances between firms and industry/state level religion composition of directors vary considerably, but are generally quite small. Panel A of the table shows key moments of these 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. However, even beyond the 50th percentile, the distances remain small, so that at the 90th percentile, the distance ranges between 0.5-0.6 across years. 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 validity of our set of instruments is also 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. Second, as explained above, we define the industry broadly at the three-digit level. The number of firms in a three-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 state-level homophily index. Third, to further ensure against this possibility, we also include several firm and board characteristics besides homophily in our regression. This accounts for mechanisms through which any one firm may influence the state or industry level homophily.

Finally, 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 three reasons described above for the validity of the state and industry-level homophily indices also apply to the distance measures. Further, there is an additional reason why this distance of a board's composition from that of industry or state composition is valid. Consider the following example. Suppose an industry's directors belong to three different religions – 50% are Hindus, 25% are Muslims, and another 25% are Christians. A firm in this industry may have ten directors, eight of whom are Hindu and two are Christians. Now, if this firm replaces two of its Hindu directors with two Muslims, then the firm's director composition becomes closer to the industry-level composition. The only mechanism that this distance can affect firm outcomes is through the endogenous regressor (board homophily). There is no reason to expect, ceteris paribus, that simply replacing two Hindus on the board with two Muslims would have any independent effect on firm performance, i.e, there is no reason to expect a pure religion effect.

The performance measures that we use as dependent variables in our regressions are Return on assets, Profits over Assets and Tobin's Q. Return on assets (ROA) is defined as the ratio of total operating income and total assets. ROA tells us what earnings were generated from invested capital (assets). Profits over assets is defined as the ratio of Profits Before Depreciation, Interest, Tax and Amortization (PBDITA) and total assets. PBITDA is essentially net income with interest, taxes, depreciation and amortization added back to it. It can be used to compare profitability among companies and industries as it eliminates the effects of financing and accounting decisions. Tobin's Q is calculated as the market value of a company divided by the replacement value of the firm's assets. A low Q (between 0 and 1) means that the cost to replace a firm's assets is greater than the value of its stock. This implies that the stock is undervalued. Conversely, a high Q

(greater than 1) implies that a firm's stock is more expensive than the replacement cost of its assets, which implies that the stock is overvalued.

The control variables we use in our regressions are firm size (measured by real assets), Tangible Asset Intensity, Firm Volatility and Book Leverage. Tangible Asset Intensity is defined as the fraction of tangible assets in the total assets of a company. A tangible asset is an asset that has a physical form. Tangible assets include both fixed assets, such as machinery, buildings and land, and current assets, such as inventory. Firm Volatility is a statistical measure of the dispersion of returns for a given security or market index. Volatility is measured by using the standard deviation between returns from same security over a year. Volatility is interpreted as the amount of uncertainty or risk about the size of changes in a security's value. A higher volatility means that a security's value can potentially be spread out over a larger range of values. This means that the price of the security can change dramatically over a short time period in either direction. A lower volatility means that a security's value does not fluctuate dramatically, but changes in value at a steady pace over a period of time. Book Leverage is calculated as the ratio of the total debt of a company and the total assets. Book Leverage is a measure of the company's assets that are financed by debt, rather than equity. This leverage ratio shows how a company has grown and acquired its assets over time. Investors use the ratio to evaluate whether the company has enough funds to meet its current debt obligations, and also assess whether the company can pay a return on their investment. Creditors use the ratio to see how much debt the company already has and if the company has the ability to repay its debt, which will determine whether additional loans will be extended to the firm.

We additionally include controls for board characteristics other than homophily – CEO duality and percent of independent directors. Other control include indicators for whether the firm is public, listed on the stock market, year, state, and industry fixed effects. Standard errors are corrected for arbitrary heteroskedasticity and are clustered at the level of two-digit industries.

5 Homophily and Firm Performance

5.1 Homophily Is Negatively Associated with Firm Performance

We first present fixed effects results for the association between firm performance measures and religion and caste homophily of boards of directors. Results are presented in Table 11. Columns (1)-(3) present results for the association between *varna* homophily of

boards and ROA, Tobin's Q, and profits / assets, respectively. Columns (4)-(6) present analogous results for *jati* homophily and columns (7)-(9) for religion homophily. We observe that for all measures of cultural homophily, the association is negative with firm performance measures, albeit statistically insignificant in some cases. Firms within industries, states, and years, that have higher homophily perform worse. It is interesting to note that board homophily has a much stronger association with performance measures than other board characteristics. We do not include board size since it is highly correlated with our homophily measure which by construction already accounts for board size.

Table 11: OLS Results

	Varna			Jati			Religion		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	ROA	Tobin's Q	Profits to Assets	ROA	Tobin's Q	Profits to Assets	ROA	Tobin's Q	Profits to Assets
Board Homophily	-1.137	-0.207	-0.038***	-1.959**	-0.322**	-0.060***	-27.560***	-0.860	-0.260***
	(0.746)	(0.151)	(0.011)	(0.751)	(0.154)	(0.018)	(4.188)	(0.728)	(0.080)
Total Assets	0.001***	0.000	0.000	0.000**	-0.000	-0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pct Independent Dirs	-0.061**	0.003	-0.000	-0.065**	-0.002	-0.001	-0.112***	0.000	-0.001
	(0.029)	(0.006)	(0.001)	(0.031)	(0.006)	(0.001)	(0.027)	(0.006)	(0.000)
Dual CEO	0.202	-0.008	0.010	-0.024	-0.013	0.008	0.094	-0.025	0.008
	(0.311)	(0.044)	(0.006)	(0.325)	(0.045)	(0.006)	(0.270)	(0.043)	(0.006)
Public	-5.275***	0.997***	-0.031	-6.313***	0.784**	-0.092***	-4.789***	0.922***	-0.042
	(1.084)	(0.217)	(0.043)	(1.400)	(0.298)	(0.022)	(1.057)	(0.195)	(0.038)
Listed	-0.393	0.014	-0.014	0.463	0.061	-0.010	-0.297	0.017	-0.009
	(0.646)	(0.069)	(0.012)	(0.591)	(0.074)	(0.013)	(0.560)	(0.070)	(0.011)
Tangible Asset Intensity	3.871	-0.535***	0.041	3.283	-0.397***	0.040	3.718	-0.474***	0.037
	(2.860)	(0.155)	(0.048)	(2.983)	(0.135)	(0.049)	(2.750)	(0.145)	(0.044)
Firm Volatility	-1.413	-2.221**	-0.003	-1.218	-1.901**	-0.012	-1.368	-1.932***	-0.004
	(1.362)	(0.965)	(0.011)	(1.187)	(0.739)	(0.009)	(1.254)	(0.687)	(0.010)
Book Leverage	-2.527*	0.978***	-0.009	-2.361*	0.993***	-0.006	-2.337*	0.978***	-0.006
	(1.377)	(0.027)	(0.011)	(1.353)	(0.021)	(0.010)	(1.345)	(0.028)	(0.010)
Constant	13.782***	-0.388	0.139**	16.699***	-0.188	0.233***	18.975***	0.437	0.205***
	(3.372)	(0.295)	(0.057)	(2.869)	(0.302)	(0.036)	(3.132)	(0.632)	(0.053)
Observations	13,383	8,519	14,198	12,508	7,919	13,361	14,695	9,357	15,602
R-squared	0.025	0.417	0.006	0.022	0.420	0.010	0.035	0.380	0.010
Number of 2 digit industries	63	62	63	63	61	63	64	62	64
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

5.2 Homophily Negatively Affects Firm Performance

As explained in section 4, cultural homophily of the board is an endogenous regressor. To examine the causal effects of homophily on firm performance, we instrument for it using the homophily among directors of all firms in the same state and industry of the firm. In another specification, we also include the Euclidean distance between the cultural composition of firm boards and of all directors in the state and industry.

Table 12 presents the first stage results. Columns (1)-(3) include as instruments the homophily among directors of all firms in the same state and industry of the firm. Columns (4)-(6) additionally include the Euclidean distance between the cultural composition of firm boards and of all directors in the state and industry. As we go through the columns we present results for the three measures of cultural homophily – varna, jati, and religion, respectively. We see that firm board homophily is higher in industries and states with greater director homogeneity. In the second specification, we see that the greater the distance between the firm and the state or industry, the more culturally concentrated the firm board tends to be. In all cases, the F-statistic is well above 10 indicating that the instruments are highly relevant to the endogenous regressor.

Tables 13, 14, and 15 present the second stage results for the three firm performance measures: ROA, profits/aseets, and Tobin's Q, respectively. In each table, columns (1)-(3) include as instruments the homophily among directors of all firms in the same state and industry of the firm. Columns (4)-(6) additionally include the Euclidean distance between the cultural composition of firm boards and of all directors in the state and industry. As we go through the columns we present results for the three measures of cultural homophily – varna, jati, and religion, respectively. For all three performance measures, and for each measure of cultural homophily, we see that higher cultural homogeneity on the board negatively affects firms' ROA, profits/assets, and Tobin's Q. In some cases, the coefficients are not statistically significant due to large standard errors. However, the magnitude of the coefficients is economically large.

6 Conclusion

This paper shows that boards of directors of large public and private firms in India are characterized by high and persistent levels of cultural homophily. In general, this has a detrimental effect on firms. Results show that firms with better economic performance have more diverse boards. Similarly, sectors and states that witnessed the greatest increases in real output also saw their firms' board homophilies decline. We show that these

Table 12: First Stage IV Results

	Instruments: Homophily in State and Industry			Instruments: Homophily in State and Industry, Distance between caste composition of firm boards and state or industry		
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: ROA					
	Varna	Jati	Religion	Varna	Jati	Religion
Varna Homophily for 2-digit industries	0.307*			0.388**		
	(0.184)			(0.153)		
Varna Homophily for States	0.415***			0.129***		
	(0.036)			(0.024)		
Varna Euclideean Distance for 2-digit Industries				0.679***		
				(0.009)		
Varna Euclideean Distance for States				0.136***		
				(0.009)		
Jati Homophily for 1-digit industries		1.844***			0.718**	
		(0.654)			(0.316)	
Jati Homophily for State		0.487***			0.620***	
		(0.064)			(0.038)	
Jati Euclideean Distance for 1-digit Industries					0.825***	
					-0.013	
Jati Euclideean Distance for States					0.233***	
					(0.013)	
Religion Homophily for 2-digit industries			0.313***			0.321***
			(0.095)			(0.095)
Religion Homophily for State			-0.032***			-0.025***
			(0.004)			(0.005)
Religion Euclideean Distance for 2-digit Industries						0.045***
						(0.005)
Religion Euclideean Distance for States						-0.006
						(0.005)
Total Assets	-0.000***	-0.000***	-0.000***	-0.000***	0.000*	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pct Independent Dirs	-0.004***	-0.007***	-0.002***	-0.000	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Dual CEO	0.004	-0.002	0.000	-0.001	0.000	-0.000
	(0.003)	(0.003)	(0.001)	(0.002)	(0.001)	(0.001)
Public	-0.026**	0.066***	0.023***	0.026***	0.146***	0.022***
	(0.011)	(0.011)	(0.004)	(0.010)	(0.005)	(0.004)
Listed	0.020***	0.020***	-0.002	0.011*	0.013***	-0.003
	(0.007)	(0.008)	(0.002)	(0.006)	(0.005)	(0.002)
Tangible Asset Intensity	-0.042***	-0.047***	-0.007**	-0.031***	-0.006	-0.007*
	(0.010)	(0.011)	(0.004)	(0.006)	(0.006)	(0.003)
Firm Volatility	0.012	0.022	0.010	0.004	0.003	0.009
	(0.021)	(0.028)	(0.012)	(0.006)	(0.004)	(0.011)
Book Leverage	0.003	0.005**	0.010***	-0.002*	-0.003***	0.010***
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.003)
Observations	13,390	12,516	14,701	13,390	12,516	14,701
F-Static	68.65	32.93	31.11	2891.13	5179.09	76.26
Number of 2 digit Industries	62		62	62		62
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Number of 1 digit Industries		10			10	

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Table 13: Second Stage IV Results: Return on Assets

	Instruments: Homophily in State and Industry			Instruments: Homophily in State and Industry, Distance between caste composition of firm boards and state or industry		
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: ROA					
	Varna	Jati	Religion	Varna	Jati	Religion
Varna Homophily	-9.470			-3.060***		
	(8.453)			(0.783)		
Jati Homophily		-18.027*			-2.827***	
		(10.378)			(0.674)	
Religion Homophily			-12.316			-27.931**
			(42.392)			(12.669)
Total Assets	0.001*	0.000	0.001	0.001***	0.001***	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Pct Independent Dirs	-0.098*	-0.169**	-0.085	-0.072***	-0.070**	-0.118***
	(0.054)	(0.084)	(0.083)	(0.027)	(0.031)	(0.035)
Dual CEO	0.267	-0.028	0.139	0.251	0.007	0.141
	(0.268)	(0.274)	(0.236)	(0.256)	(0.276)	(0.238)
Public	-5.714*	-5.215	-5.423	-5.627*	-6.283*	-5.091
	(3.289)	(3.399)	(3.421)	(3.292)	(3.276)	(3.293)
Listed	-0.133	0.362	-0.207	-0.277	0.043	-0.252
	(0.526)	(0.594)	(0.454)	(0.465)	(0.532)	(0.427)
Tangible Asset Intensity	3.578	4.056*	3.837	3.813	4.748**	3.711
	(2.344)	(2.254)	(2.563)	(2.553)	(2.369)	(2.371)
Firm Volatility	-1.324	-1.048	-1.494	-1.392	-1.379	-1.340
	(1.229)	(0.919)	(1.515)	(1.358)	(1.318)	(1.284)
Book Leverage	-2.519**	-2.388**	-2.510*	-2.535**	-2.476**	-2.350**
	(1.168)	(1.149)	(1.311)	(1.181)	(1.175)	(1.148)
Observations	13,390	12,516	14,701	13,390	12,516	14,701
R-squared	0.012	-0.015	0.030	0.022	0.022	0.033
Number of 2 digit Industrie	62		62	62		62
Industry FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES			
Year FE	YES	YES	YES	YES	YES	YES
Number of 1 digit Industries		10			10	

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Table 14: Second Stage IV Results: Profits / Assets

	Instruments: Homophily in State and Industry			Instruments: Homophily in State and Industry, Distance between caste composition of firm boards and state or industry		
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Profits / Assets					
	Varna	Jati	Religion	Varna	Jati	Religion
Varna Homophily	-0.341*			-0.038***		
	(0.176)			(0.013)		
Jati Homophily		0.240			-0.058***	
		(0.355)			(0.012)	
Religion Homophily			1.233*			-0.066
			(0.711)			(0.164)
Total Assets	-0.000	0.000	0.000**	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pct Independent Dirs	-0.001	0.001	0.003*	-0.000	-0.001	-0.000
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dual CEO	0.010*	0.007	0.008	0.010*	0.006	0.007
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Public	-0.077***	-0.098***	-0.106***	-0.072***	-0.078***	-0.079***
	(0.019)	(0.032)	(0.024)	(0.018)	(0.018)	(0.018)
Listed	-0.006	-0.012	-0.006	-0.014	-0.005	-0.008
	(0.011)	(0.012)	(0.010)	(0.009)	(0.011)	(0.009)
Tangible Asset Intensity	0.028	0.051	0.040	0.040	0.039	0.038
	(0.040)	(0.040)	(0.041)	(0.044)	(0.039)	(0.041)
Firm Volatility	0.002	-0.022	-0.021	-0.002	-0.014	-0.006
	(0.014)	(0.023)	(0.025)	(0.011)	(0.011)	(0.011)
Book Leverage	-0.008	-0.006	-0.013	-0.009	-0.006	-0.007
	(0.007)	(0.007)	(0.009)	(0.007)	(0.007)	(0.007)
Observations	14,205	13,369	15,608	14,205	13,369	15,608
R-squared	-0.026	-0.019	-0.043	0.004	0.003	0.003
Number of 2 digit Industries	62		62	62		62
Industry FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES			
Year FE	YES	YES	YES	YES	YES	YES
Number of 1 digit Industries		10			10	

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Table 15: Second Stage IV Results: Tobin's Q

	Instruments: Homophily in State and Industry			Instruments: Homophily in State and Industry, Distance between caste composition of firm boards and state or industry		
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Tobin's Q					
	Varna	Jati	Religion	Varna	Jati	Religion
Varna Homophily	-4.389***			-0.353***		
	(0.802)			(0.121)		
Jati Homophily		-2.631***			-0.354***	
		(0.887)			(0.123)	
Religion Homophily			16.984***			-0.243
			(4.800)			(1.788)
Total Assets	-0.000**	-0.000*	0.000***	0.000*	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pct Independent Dirs	-0.014***	-0.018**	0.035***	0.001	-0.004	0.000
	(0.005)	(0.007)	(0.011)	(0.005)	(0.005)	(0.006)
Dual CEO	0.007	-0.024	-0.011	-0.002	-0.020	-0.018
	(0.030)	(0.027)	(0.030)	(0.026)	(0.026)	(0.025)
Public	0.518***	0.664***	0.325**	0.623***	0.538***	0.607***
	(0.120)	(0.130)	(0.152)	(0.099)	(0.098)	(0.091)
Listed	0.069	0.077	0.016	-0.026	0.043	-0.028
	(0.071)	(0.063)	(0.078)	(0.057)	(0.059)	(0.054)
Tangible Asset Intensity	-0.699***	-0.484***	-0.515***	-0.577***	-0.422***	-0.513***
	(0.123)	(0.107)	(0.112)	(0.095)	(0.093)	(0.096)
Firm Volatility	-0.657	-0.806	-4.762**	-2.112*	-2.026*	-2.001**
	(0.607)	(0.859)	(1.967)	(1.086)	(1.099)	(0.946)
Book Leverage	1.000***	0.991***	0.913***	0.979***	0.990***	0.975***
	(0.033)	(0.021)	(0.061)	(0.026)	(0.022)	(0.028)
Observations	8,522	7,925	9,361	8,522	7,925	9,361
R-squared	0.231	0.363	0.140	0.410	0.412	0.373
Number of 2 digit Industrie	59		60	59		60
Industry FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES			
Year FE	YES	YES	YES	YES	YES	YES
Number of 1 digit Industries		10			10	

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

relationships are causal. Causal regression analysis shows that high cultural homophily negatively affects firm performance as measured by firm profitability and firm value.

References

Ajit, D., H. Donker, and R. Saxena, 2012, "Corporate Boards in India Blocked by Caste?," *Economic & Political Weekly*, 47(32), August 11, 2012, 39-43.

Alesina, A., and P. Giuliano, 2010, "The power of the family," *Journal of Economic Growth*, Springer, 15(2), 93-125, June.

Alesina, A., and P. Giuliano, 2011, "Family Ties and Political Participation," *Journal of the European Economic Association* 9(5), 817-839.

Alesina, A., and P. Giuliano, 2014, "Family Ties," *Handbook of Economic Growth, edited by Philippe Aghion and Steven N Durlauf*, 2A, 177-215, The Netherlands: North Holland.

Alesina, A., B. Reich, and A. Riboni, 2017, "Nation-Building, Nationalism and Wars," No. w23435. *National Bureau of Economic Research*.

Alesina, A., P. Giuliano, and N. Nunn, 2013, "On the Origin of Gender Roles: Women and the Plough," *Quarterly Journal of Economics* 128(2), 469-530.

Alesina, A., R. Baqir, and W. Easterly, 1999, "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, Vol. 114, 1234-84.

Alesina, A., S. Michalopoulos, and E. Papaioannou, 2016, "Ethnic Inequality," *Journal of Political Economy* 124(2), 428-488.

Anderson, S., 2011, "Caste as an Impediment to Trade," *American Economic Journal: Applied Economics*, 3w(1): 239-63. DOI: 10.1257/app.3.1.239

Ashraf, Q. H., and O. Galor, 2007, "Cultural assimilation, cultural diffusion and the origin of the wealth of nations."

Aslund, O., L. Hensvik, and O. N. Skans, 2014, "Seeking Similarity: How Immigrants and Natives Manage in the Labor Market," *Journal of Labor Economics*, 32(3), 405-441.

Atkin, D., 2016, "The caloric costs of culture: Evidence from Indian migrants," *American Economic Review*, 106(4), 1144-81.

Banerjee, A., E. Duflo, M. Ghatak, J. Lafortune , 2013, "Marry for What? Caste and Mate Selection in Modern India," *American Economic Journal: Microeconomics*, 5(2), 1561-1605.

Barro, R. J., and R. M. McCleary, 2003, "Religion and Economic Growth across Countries," *American Sociological Review*, 68(5), 760-81. http://www.jstor.org/stable/1519761.

Bengtsson, O. and D. H. Hsu, 2015, "Ethnic matching in the U.S. venture capital market," *Journal of Business Venturing*, 30(2), 338-354.

Benjamin, D. J., J. J. Choi, and G. Fisher, 2016, "Religious Identity and Economic Behavior," *The Review of Economics and Statistics*, 98(4), 617-637

Bernile, G., V. Bhagwat, S. Yonkers, 2017, "Board Diversity, Firm Risk, and Corporate Policies", *Journal of Financial Economics*.

Bertrand, M., S. E. Black, S. Jensen, and A. Lleras-Muney, 2014, "Breaking the glass ceiling? The effect of board quotas on female labor market outcomes in Norway," No. w20256. *National Bureau of Economic Research*.

Bloom, N., J. Liang, J. Roberts and Z. J. Ying, 2015, "Does Working from Home Work? Evidence from a Chinese Experiment," *The Quarterly Journal of Economics*, 130(1), 1 February 2015, 165218. https://doi.org/10.1093/qje/qju032

Bloom, N., R. Sadun, J. V. Reenen, 2012, "The Organization of Firms Across Countries," *The Quarterly Journal of Economics*, 127(4), 1 November 2012, 16631705. https://doi.org/10.1093/qje/qje029

Campante, F., and D. Yanagizawa-Drott, 2015, "Does Religion Affect Economic Growth and Happiness? Evidence from Ramadan," *The Quarterly Journal of Economics*, 130(2), 1 May 2015, 615658, https://doi.org/10.1093/qje/qjv002

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

Currarini, S., M.O. Jackson, and P. Pin, 2009, "An Economic Model of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, Vol. 77, 1003-1045. doi:10.3982/ECTA7528

Damodaran, H., 2011, "Indias New Capitalists: Caste, Business, and Industry in a Modern Nation," *South Asian Journal of Management*, p.141.

Dass, N., O. Kini, V. Nanda, B. Onal, and J. Wang, 2014, "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., F. Hoffmann, and P. Oreopoulos, 2014, "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom," *American Economic Review*, 104(8), 2567-2591.

Fernandez, R., 2007, "Alfred Marshall lecture women, work, and culture," *Journal of the European Economic Association*, 5(2-3), 305-332.

Fernndez, R., 2011, "Does culture matter?," In *Handbook of social economics*, Vol. 1, 481-510. North-Holland.

Fernndez, R., 2013, "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century," *American Economic Review*, 103w(1), 472-500. doi: 10.1257/aer.103.1.472

Fernndez, R., and A. Fogli, 2006, "Fertility: The Role of Culture and Family Experience," *Journal of the European Economic Association* 4(2-3), 552-61. http://www.jstor.org/stable/40005121.

Fernndez, R., and A. Fogli, 2009, "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 1w(1), 146-77. doi: 10.1257/mac.1.1.146

Fisman, R., D. Paravisini, and V. Vig, 2017, "Cultural Proximity and Loan Outcomes," *American Economic Review*, 107(2), 457-492.

Freeman, R. B., and W. Huang, 2015, "Collaborating with People Like Me: Ethnic Coauthorship Within the United States," *Journal of Labor Economics* 33(S1) (July): S289S318. doi:10.1086/678973.

Ghani, E., W. R. Kerr, and S. O'Connell, 2013, "Spatial Determinants of Entrepreneurship in India," *Special Issue on Entrepreneurship in a Regional Context . Regional Studies* 48(6), December 2013.

Giuliano, L., and M. Ransom, 2013, "MANAGER ETHNICITY AND EMPLOYMENT SEGREGATION," *ILR Review*, 66(2), 346-379. http://www.jstor.org/stable/24369537

Giuliano, L., D. I. Levine, and J. Leonard, "Manager Race and the Race of New Hires," *Journal of Labor Economics* 27(4) (October 2009): 589-631.

Gorodnichenko, Y., and G. Roland, 2015, "Culture, institutions and democratization," No. w21117. *National Bureau of Economic Research*.

Guiso, L., P. Sapienza and L. Zingales, 2003, "People's opium? Religion and economic attitudes," *Journal of Monetary Economics*, Elsevier, 50(1), pages 225-282, January.

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.* http://dx.doi.org/10.2139/ssrn.1939587.

Herfindal, O. C., 1950, "Concentration in the U.S. Steel Industry," *PhD Dissertation, Columbia University*.

Hirschman, A. O., 1980, "National power and the structure of foreign trade," Vol. 105, *Univ of California Press*.

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

Hnatkovska, V., A. Lahiri, and S. Paul, 2012, "Castes and Labor Mobility," *American Economic Journal: Applied Economics*, 4w(2), 274-307. DOI: 10.1257/app.4.2.274

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

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

Iyer, L., T. Khanna, and A. Varshney, 2013, "Caste and Entrepreneurship in India," *Economic & Political Weekly*, 48(6), February 9, 2013, 52-60.

Jayaraman, R. , 2005, "Personal Identity in a Globalized World: Cultural Roots of Hindu Personal Names and Surnames," *The Journal of Popular Culture*, Vol. 38, 476490.

Jodhka, S. S., 2010, "Dalits in business: Self-employed scheduled castes in North-West India," *Economic & Political Weekly*, 45(11), 41-48.

Joshi, P. , 2005, "Slash Surname to Kill Caste," *The Telegraph*, October 15, 2005

Kenneth R. A., A. K. Dittmar, 2012, "The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation," *The Quarterly Journal of Economics*, 127:1, 1 February 2012, 137197, https://doi.org/10.1093/qje/qjr049

Knyazeva, A., M. Knyazeva, and R. W. Masulis, 2013, "The Supply of Corporate Directors and Board Independence," *The Review of Financial Studies*, 26(6), 1561-1605.

Kosheleva, O., C. Tweedie, and V. Kreinovich, 2015, "Which Bio-Diversity Indices are Most Adequate," *Fuzzy Systems (FUZZ-IEEE)*, 2015 IEEE International Conference on. IEEE.

Kramarz, F. and D. Thesmar, 2013, "SOCIAL NETWORKS IN THE BOARDROOM," *Journal of the European Economic Association*, Vol. 11, 780-807. doi:10.1111/jeea.12021

Kramarz, Francis, and D. Thesmar, 2013, "Social networks in the boardroom," *Journal of the European Economic Association*, 11(4), 780-807. doi:10.1111/jeea.12021

M. N. SRINIVAS, 1963, "Social Change in Modern India," (The Rabindranath Tagore Memorial Lectureship,delivered at the University of California,Berkeley, May 1963.) Berkeley and Los Angeles: *University of California Press*,1966. 15(194) pp.,index, notes, reading list.

Macdonell, A. A., 1914, "The Early History of Caste," *The American Historical Review*, 19(2), 230-44. doi:10.2307/1862285.

McCleary, R. M., and R. J. Barro, 2006, "Religion and Economy," *Journal of Economic Perspectives*, 20w(2): 49-72. doi: 10.1257/jep.20.2.49

McPherson, M., L. Smith-Lovin, and J. M. Cook, 2001, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, Vol. 27, 415-444.

Munshi, K., 2011, "Strength in Numbers: Networks as a Solution to Occupational Traps," *Review of Economic Studies*, Vol. 78, 1069-1101.

Munshi, K., and M. Rosenzweig, 2006, "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy," *American Economic Review*, 96(4), 122552.

Munshi, K., and M. Rosenzweig, 2013, "Networks, commitment, and competence: caste in Indian local politics," No. w19197. ,*National Bureau of Economic Research*.

Munshi, K., and M. Rosenzweig, 2016, "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *American Economic Review*, 106w(1), 46-98, doi: 10.1257/aer.20131365

Noland, M., 2005, "Religion, and Economic Performance," *World Development*, 33(8), 12151232, doi:10.1016/j.worlddev.2005.03.006

Petersen, T., I. Saporta, and M.-D. L. Seidel, "Offering a Job: Meritocracy and Social Networks," *American Journal of Sociology* 106(3) (November 2000): 763-816.

Routledge, R. D., 1979, "Diversity Indices: Which Ones are Admissible?," *Journal of theoretical Biology*, 76(4), 503-515.

Shannon, C. E. "A Mathematical Theory of Communication," *The Bell Systems Technical Journal*, 1948, Vol. 27, 379423 and 623656.

Simpson, E. H. "Measurement of Diversity," *Nature*, 1949, Vol. 163, p. 688

Tabellini, G., 2010, "CULTURE AND INSTITUTIONS: ECONOMIC DEVELOPMENT IN THE REGIONS OF EUROPE," *Journal of the European Economic Association*, 8(4), 677-716. http://www.jstor.org/stable/25700901.

Terjesen, S., R. Sealy, and V. Singh, 2009, "Women Directors on Corporate Boards: A Review and Research Agenda," *Corporate Governance: An International Review*, Vol. 17, 320-337. doi:10.1111/j.1467-8683.2009.00742.x

Thorat, S.K., D. Kundu and N. Sadana, 2010, "Caste and Ownership of Private Enterprises: Consequences of Denial of Property Rights", in S.K. Thorat and K.S. Newman (eds), "Blocked By Caste: Economic Discrimination in Modern India," New Delhi,*Oxford University Press*

Times of India. , 2016, "Ban Surnames to End Casteism: Narendra Jadhav," *Times of India*, August 8, 2016

Titzmann, F.-M., 2013, "Changing Patterns of Matchmaking: The Indian Online Matrimonial Market," *Asian Journal of Women's Studies*, Vol. 19, 64-94.

Varshney, A., 2012, "Two Banks of the Same River? Social Order and Entrepreneurialism in India," *Anxieties of Democracy: Tocquevillean Reflections on India and the United States*, 225-56.

Vissa, B., 2011, "A matching theory of entrepreneurs' tie formation intentions and initiation of economic exchange," *Academy of Management Journal*, 54(1), 137-158.

Appendices

Appendix A Stylized Facts for Varna Homophily

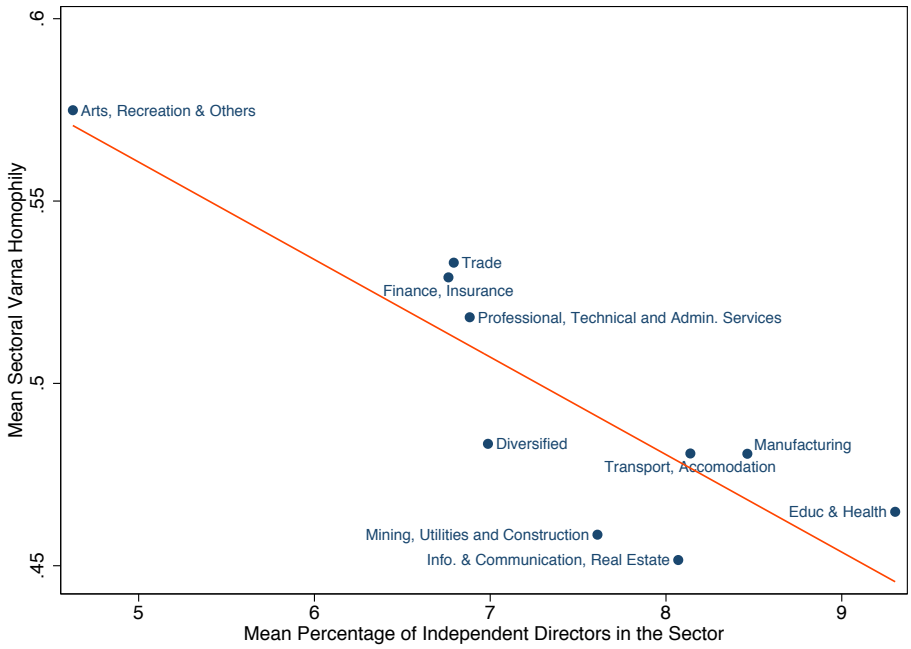


Figure Appendix A.1: Varna Homophily and Percentage of Independent Directors[a]

[a]Notes: Source: Prowess, authors' last name to varna mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.

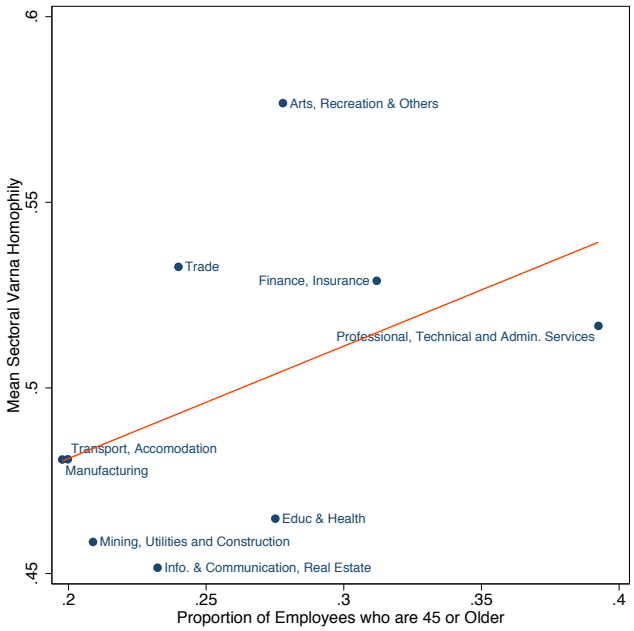


Figure Appendix A.2: Association between Varna Homophily and Worker Age[a]

[a]Notes: Source: Prowess, authors' last name to varna mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.

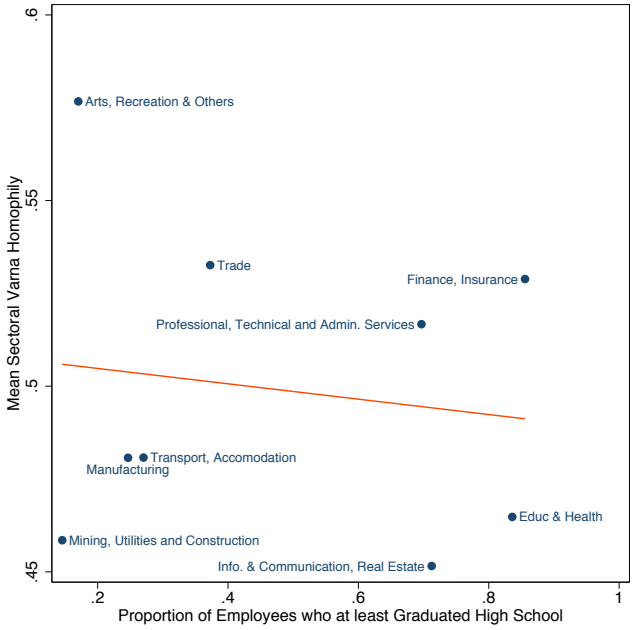


Figure Appendix A.3: Association between Varna Homophily and Worker Education[a]

[a]Notes: Source: Prowess, authors' last name to varna mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.

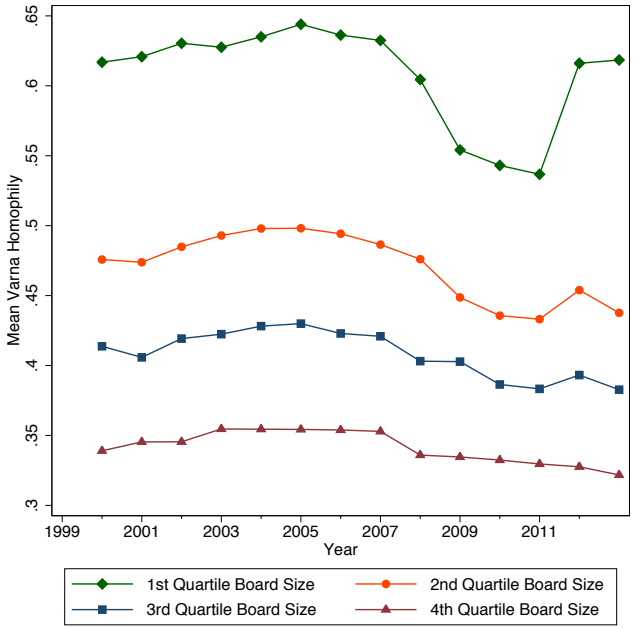


Figure Appendix A.4: Average Varna Homophily by Board Size

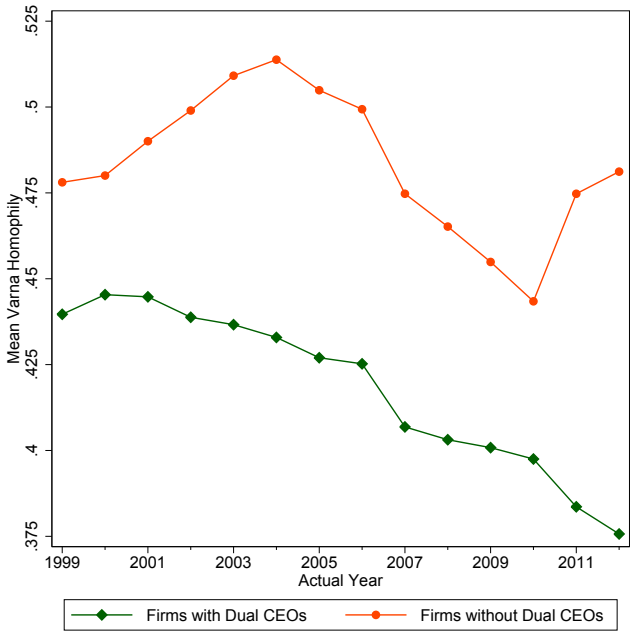


Figure Appendix A.5: Average Varna Homophily by CEO Duality

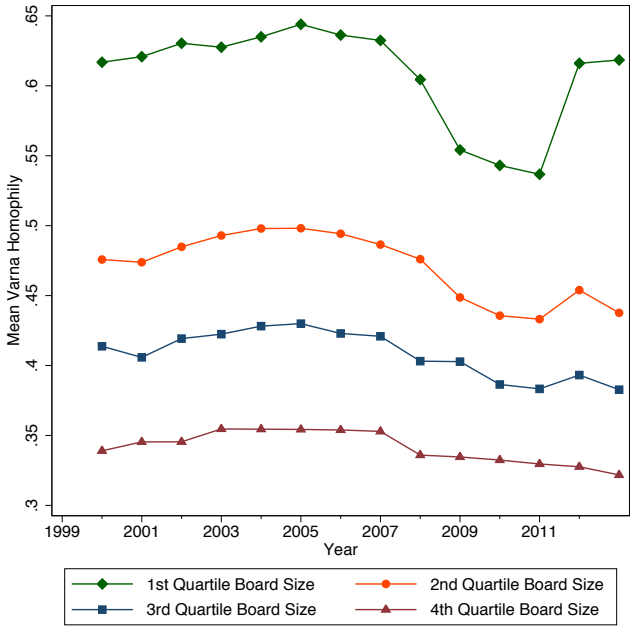


Figure Appendix A.6: Average Varna Homophily by Board Size

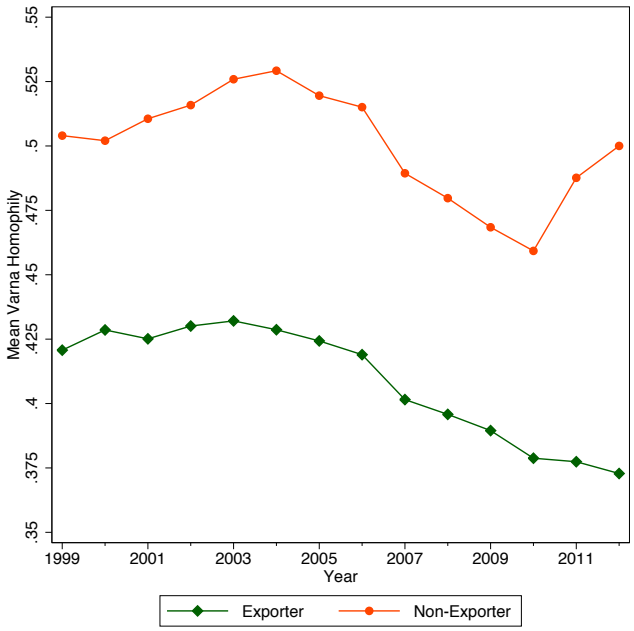


Figure Appendix A.7: Average Varna Homophily by Exporting Status

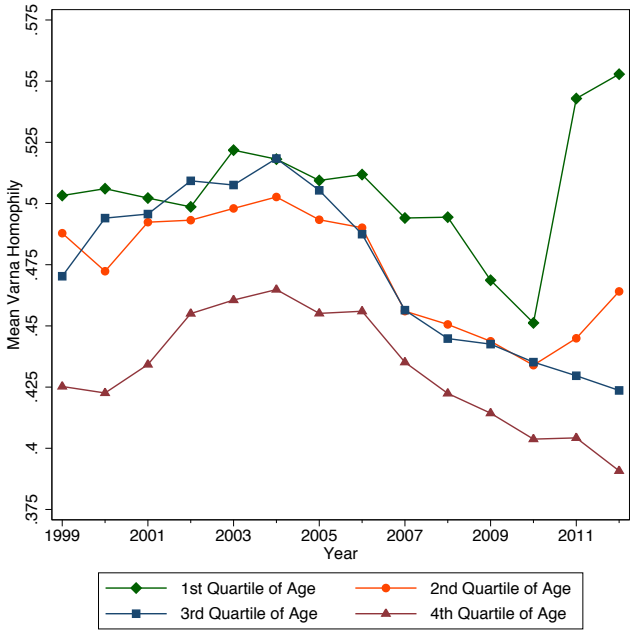


Figure Appendix A.8: Average Varna Homophily by Firm Age

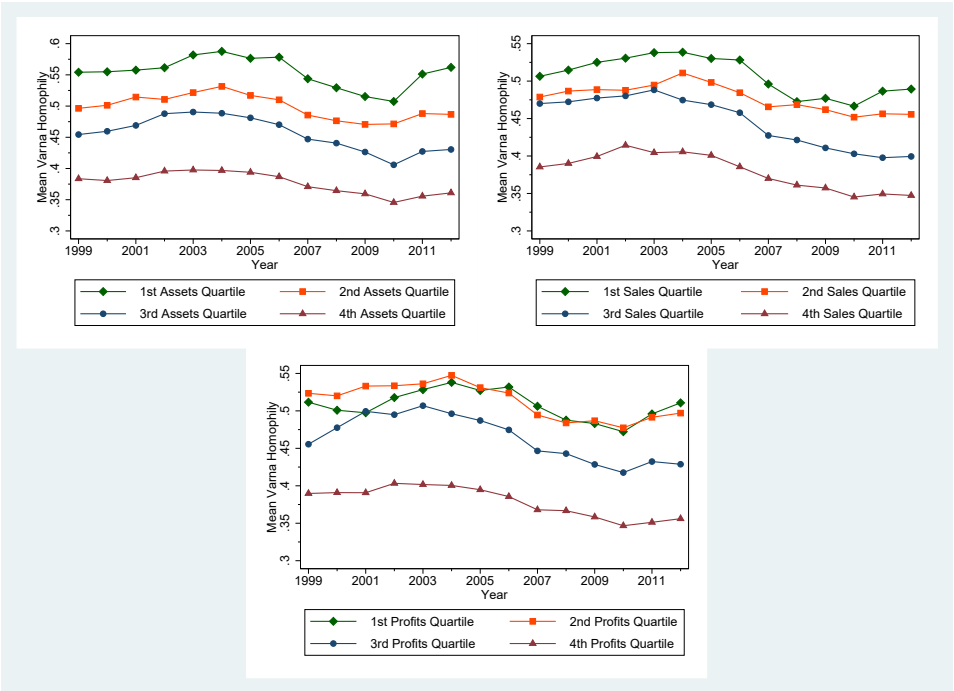
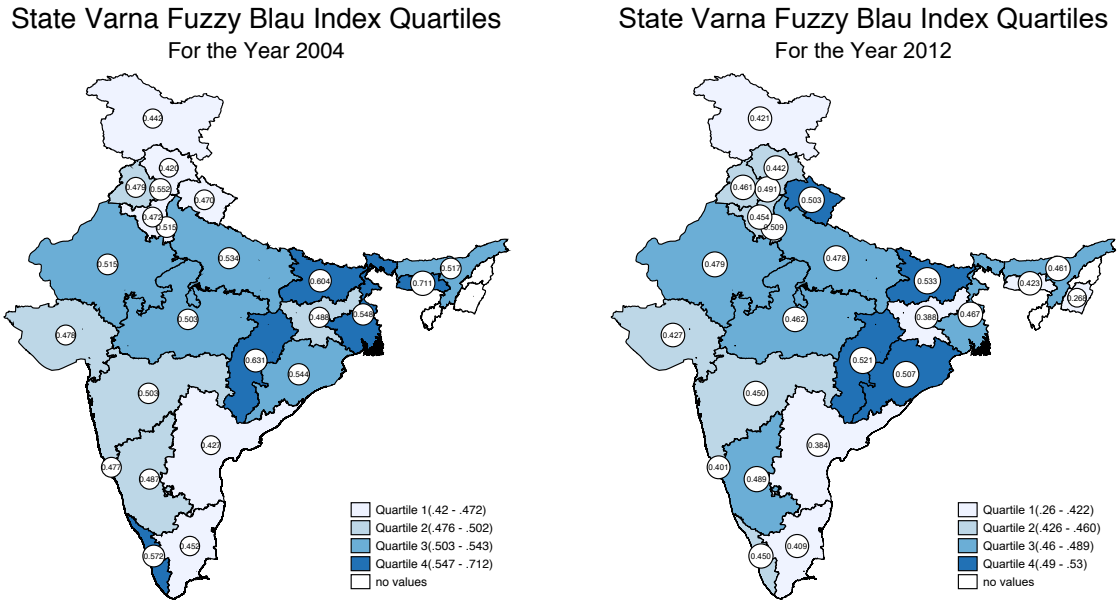


Figure Appendix A.9: Average Varna Homophily by Firm Size[a]

[a]Notes: Source: Prowess, authors' last name to religion mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.

Figure Appendix A.10: Average Varna Homophily in Broad Sectors



(a) (b)

Figure Appendix A.11: Average Varna Homophily Across States[a]

[a]Source: Prowess, authors' mapping of last names to religion using matrimonial data.

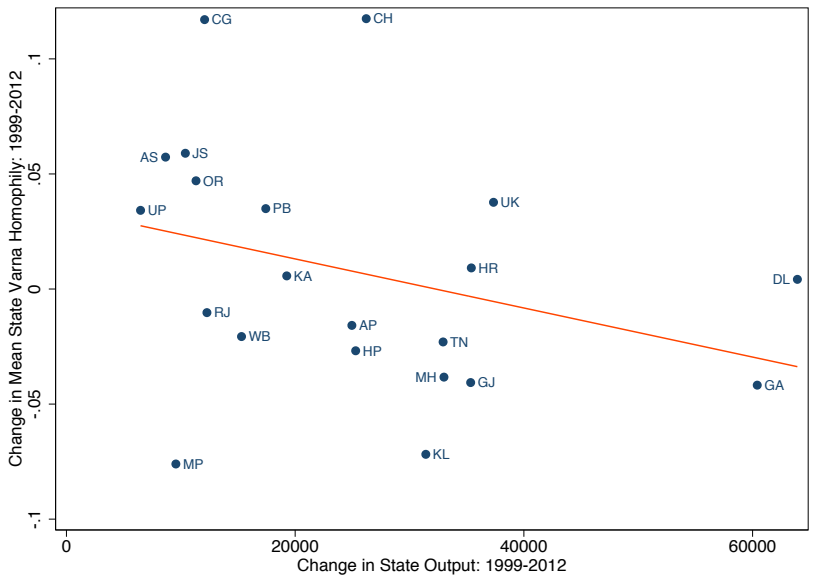


Figure Appendix A.12: Association between Varna Homophily and State Output[a]

[a]Notes: Source: Prowess, authors' last name to varna mapping using matrimonial data. Homophily is averaged over all firms in a state in each year.

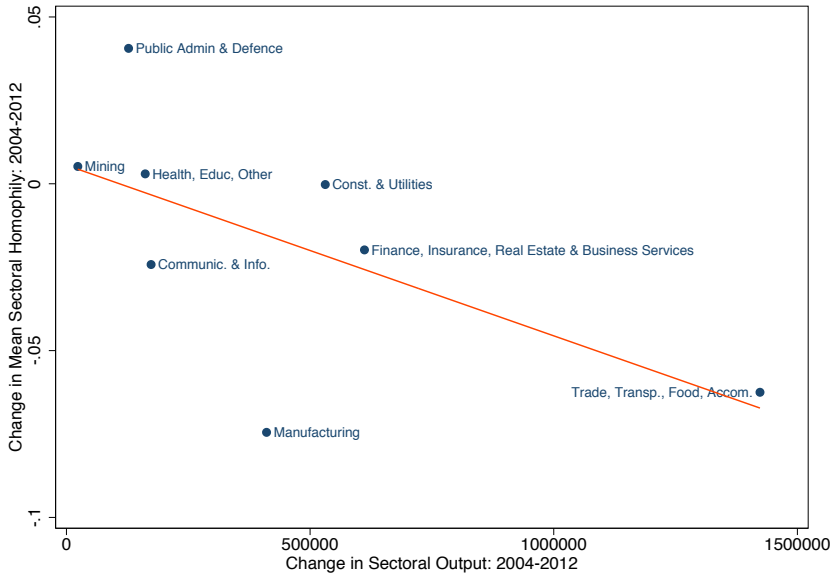


Figure Appendix A.13: Association between Varna Homophily and Sectoral Output[a]

[a]Notes: Source: Prowess, authors' last name to varna mapping using matrimonial data. Homophily is averaged over all firms in a broad sector in each year.