Journal of Development Economics 159 (2022) 102973

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular article

Community origins of industrial entrepreneurship in colonial India[☆]

Bishnupriya Gupta[a,*,1], Dilip Mookherjee[b,1], Kaivan Munshi[c,d,1], Mario Sanclemente[e,1]

[a] Department of Economics, University of Warwick, Coventry, CV4 7AL, UK
[b] Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, USA
[c] Department of Economics, Yale University, New Haven, CT 06511, USA
[d] Toulouse School of Economics, Toulouse, France
[e] QuantumX, Santiago, 7580206, Chile

A R T I C L E I N F O

Keywords:
Industrialization
Social networks

A B S T R A C T

We provide evidence of the role of community networks in emergence of Indian entrepreneurship in early stages of cotton and jute textile industries in the late 19th and early 20th century respectively, overcoming lack of market institutions and government support. From business registers, we construct a yearly panel dataset of entrepreneurs in these two industries. We find no evidence that entry was related to prior upstream trading experience or price shocks. Firm directors exhibited a high degree of clustering of entrepreneurs by community. Consistent with a model of network-based dynamics, the stock of incumbent entrepreneurs of different communities diverged non-linearly, controlling for year and community fixed effects.

1. Introduction

Differences in the timing and the determinants of early industrialization across countries have constituted an important area of research in economic history. Most of this research has focused on countries that are currently developed. While some authors have argued that property rights and well functioning goods and financial markets created conditions for the first industrial revolution in Britain (North and Weingast, 1989), others have been skeptical of the empirical evidence on this and emphasized development of on-the-job learning and dissemination of technological knowhow (Mokyr, 2005, 2009). It is widely accepted that the state played a more important role in follower countries that industrialized during the late 19th century, such as Germany and the United States.[2] In Russia, tariffs were raised across the board in 1890 to increase revenue as well as to protect industry (Markevich and Nafziger, 2017); the government subsidized railway construction and established policies to attract foreign capital (Kahan, 1967). In Japan after the Meiji restoration of 1868, the state coordinated interactions between the financial sector and industry, facilitated imports and diffusion of machinery and technological know-how (Perkins and Tang, 2017).

Countries that are currently less developed were slower to industrialize, in the absence of well-functioning markets, supportive institutions and interventionist states. Despite this, a cotton textile industry developed and gained market share relative to imports in the late 19th century in colonial India. In this paper, we explore the extent to which the vacuum created by the absence of supportive institutions and policies was filled by ethnic networks that exchanged intermediate inputs, shared know-how, connections and capital amongst themselves, and overcame contractual moral hazards via informal community enforcement mechanisms. The setting is colonial India, where market institutions were weak and the colonial state had a limited interest in industrial development, suggesting a possible role of community networks to act as a substitute. In the Indian context, social networks are defined by distinct castes, religions and countries of origin. These groups engaged in the same occupation and married within the group.

☆ We thank the Economic Development and Institutions (EDI) network, the International Growth Centre (IGC) and Centre for Competitive Advantage in the Global Economy (CAGE) for financial support. The paper has benefited from comments of two anonymous referees, the Co-Editor Oeindrila Dube, Sam Asher, James Fenske, Martin Fiszbein, Walker Hanlon, Debraj Ray and others attending presentations at the 2019 Royal Economic Society conference, 2018 World Economic History Congress, Delhi School of Economics, Paris School of Economics India-China conference, CEPR conference at NYU Abu Dhabi, EDI network meetings at the University of Namur, and seminars in Cambridge, Oxford, Hitotsubashi, Bonn, King's College London, LSE and Warwick.

* Corresponding author.
E-mail address: B.Gupta@warwick.ac.uk (B. Gupta).

1 Bishnupriya Gupta, Kaivan Munshi, Dilip Mookherjee and Mario Sanclemente have made equal contributions in the development and writing of the paper.

2 Effective sectoral tariffs of 30% were widely used in both these countries to protect infant industries (Irwin, 2007; Webb, 1980). Investment banks in Germany supported the development of large and technologically complex industries (Burhop, 2006).

https://doi.org/10.1016/j.jdeveco.2022.102973
Received 19 December 2020; Received in revised form 19 July 2022; Accepted 1 September 2022
Available online 26 September 2022

Intra-group relationships manifested high levels of trust, mutual help and assisted economic transactions and sharing of information.

The role of community networks in enforcing contracts in medieval trade has been eloquently described in the work of Greif (2006). Social networks are highly visible in economic activity in Africa (Fafchamps, 2003) and in Chinese trade (Rauch and Trindade, 2002). Even in contemporary India, community networks continue to play an important role in credit and insurance (Banerjee and Munshi, 2004; Munshi, 2011), hiring and referrals in labor markets (Beaman and Magruder, 2012) and migration flows (Munshi and Rosenzweig, 2016). However, the role of community networks in early industrialization of less developed countries has not received comparable emphasis in the literature.

Business historians have accumulated case study based evidence on entry into industry from trade in India starting from the second half of the 19th century (Goswami, 1985; Rudner, 1994; Timberg, 1978; Tripathi and Mehta, 1990; Tripathi, 2004), focusing on particular entrepreneurs and their communities. A limitation of these rich historical accounts is a lack of quantification and systematic evidence. In this paper, we build a new data set of entrepreneurs and estimate the effect of social networks in industrial entrepreneurship.

We focus mainly on the cotton textile industry in this paper; a later section shows that similar results obtain for the jute industry as well. Cotton and jute played an important role in India's industrial development during the second half of the 19th century. As shown in Table 1, the share of the textile sector rose from 12% in 1851–65 to 47% in 1900, and the vast majority of this sector consisted of cotton and jute mills during this period (Rungta, 1970). For each industry we study the downstream sector producing yarn and cloth or fabric that required large initial investments in machinery, and at locations and periods when Indian entrepreneurship began to emerge in these sectors: the cotton industry in the Bombay region between 1865–90 and the jute industry in the Calcutta region between 1914–30. We draw upon business directories to gather names of directors of listed firms in upstream and downstream activities of the concerned industry,[3] and code their respective community identities from their names. This enables us to construct a yearly panel data set of active entrepreneurs and their investments by community, track their backgrounds prior to entering the downstream industry and examine patterns of community homophily in the composition of firms. We use this evidence to understand the role of community networks in the process of entry into the downstream industry, while controlling for the role of price and other industry-wide shocks and relevant community characteristics such as prior experience, literacy, population size or outside options.[4]

We use two pieces of quantitative evidence to assess the role of communities. First, we examine the extent to which entrepreneurs clustered by community within firms. Given the high degree of interdependence among entrepreneurs within the same firms, this is a natural way to assess the extent to which problems of trust and cooperation among principal shareholders and executive officers were overcome by partnering with members of the same community. Second, we use data on entry flows and investments to test a model of entry dynamics based on productivity-enhancing help provided by incumbent entrepreneurs to new entrants from their own community. The model is appropriate for early stages of industrialization with stable market growth and a given set of communities with stocks of potential entrepreneurs with stationary outside options. Such early stages are typically characterized by growing stocks of incumbents from each community. The model generates network-based dynamics of community-specific incumbent stocks, in which differences in initial presence of different communities exponentially amplify over time. The network effect can be identified by the presence of a non-linear divergence effect for incumbent stocks, while controlling for year dummies (which include the effect of price and other industry-wide shocks) and community dummies (which capture differences in levels of community-specific unobserved characteristics such as education, ability, wealth and outside options).

The dynamics of entry into downstream cotton in the Bombay region after the US Civil War turn out to be consistent with predictions of the network-based model. The evolution of active entrepreneurs from different communities during early stages exhibited the nonlinear amplification of early community presence predicted by the model. We use yearly data for stocks of active entrepreneurs at the community level from 1866 until 1890. For reasons explained within the paper, given the identification assumption described above we are able to infer the existence of a network effect, but not obtain an unbiased estimate of its magnitude. Under the stronger assumption of exogeneity of initial community presence at the end of the US Civil War, we estimate the effect of an additional active entrepreneur from a given community in 1866 to be 2.85 additional entrepreneurs from the same community in 1880, and 5.75 additional entrepreneurs in 1890.

A potential alternative explanation for these findings could reside in positive selection with regard to unobserved community characteristics such as entrepreneurial ability, wealth or access to capital. For instance, the exceptionally high levels of education, expertise and connection with the Western world of the Parsi community have noted by many historians (Anstey, 1949; Buchanan, 1934; Desai, 1968; Tripathi, 2004). If education or wealth of the Parsi community were growing faster than the other communities, it could generate a growing divergence between their respective incumbent stocks over time. However, the historical evidence reviewed in Section 2 reveals no such pattern of divergent trends in literacy or population shares; the differences across communities remained stationary in the late 19th century. Historical evidence also shows that Parsis were similar to other communities in terms of their pre-industrial trading opportunities, as well as connections with the colonial government.

To further address this concern, we use evidence based on entry-level investments of different communities. Under the alternative hypothesis of positive selection, investments of Parsi entrants (and more generally, of communities with higher initial presence) should grow faster. However, the evidence shows the opposite pattern. Indeed, trends in entrant investments were consistent with a pattern of misallocation predicted by the network model: *marginal entrants from communities with larger initial presence are less productive, and thus progressively invest less.*[5] Our theoretical analysis shows the same pattern can appear for average entrants as well, under suitable conditions on the distribution of (unobserved) ability. Both average and marginal (10th percentile) investments of new entrants from communities with higher initial presence exhibited weak evidence of lower time trends. Moreover, the divergence between communities with high and low presence accelerated significantly over time, contrary to the predictions of a positive selection model.

[3] There were two segments of the industry: an upstream component which used light machinery to process the raw material and bale it for export, and a downstream segment which invested in heavy machinery to produce yarn and fabric or cloth mainly for the domestic market. We follow convention and use the term 'textiles' to refer only to the downstream segment.

[4] The data pertains to the stock of active entrepreneurs in different years, which includes the effects of entry as well as exit. However, entry flows dominate exits, so changes in the stock of active agents primarily reflect entry forces. We also show that the results are robust to adjusting for exits.

[5] The reason is as follows. A larger network facilitates entry of entrepreneurs who are of lower individual ability, who have lower outside options. As the marginal entrant must be indifferent between entering and not, they must attain their outside option payoff. Hence they must be less productive: the larger benefit they obtain from their network must be outweighed by their lower ability. Being less productive, they endogenously select a smaller size of investment. Moreover, as network sizes diverge across communities, the same is true for investment levels. In contrast, under the alternative positive selection hypothesis, the ability threshold does not vary across communities; hence marginal entrants from different communities should have the same productivity and invest the same.

Table 1
Share of modern sectors in paid up capital raised in India (%).
Source: Rungta (1970) (Appendix 8, Appendix 17).

	Banking & insurance	Transport	Mining & plantation	Textiles	Food industries
1851–65	35	7	20	12	0.6
1881	14	5	23	37	2
1890	13	8	19	46	1.8
1900	13	9	14	47	1.6

Table 2
Capital per firm in Rupees.
Source: Own data.

Period	Upstream		Downstream	
	Mean	*Median*	*Mean*	*Median*
1860–1890	578,990	400,000	1,206,293	1,000,000
1891–1910	345,283	225,000	1,106,615	943,500

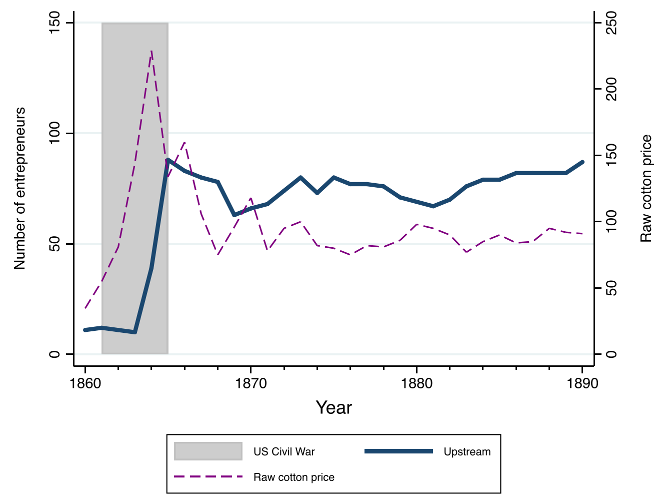


Fig. 1. Stock of Entrepreneurs in the Upstream Cotton Sector (1860–1890)
Source: Entrepreneur stocks: own data; prices: see Appendix A.

Besides community networks, we also examine the role of (and control for) other factors often emphasized in the literature as important determinants of early industrial entrepreneurship, such as pre-industrial accumulation of wealth (Marx, 1887; Banerjee and Newman, 1993) or experience in related upstream trading sectors in certain African countries (Sutton and Kellow, 2010; Sutton and Kpentey, 2012). In the context of the Bombay cotton industry, we find no evidence of a significant association of entry with prior upstream presence of the community during the US Civil War, a period during which Indian cotton traders achieved high export volumes and profits. 66% of entrepreneurs active during 1860–70 had no prior upstream experience in baling and trading raw cotton, a proportion which rose to 79% and 91% in the subsequent two decades. Hence accumulation of experience or wealth in the upstream sector was not a pre-condition for entry into the downstream industry. Nor were yearly cotton price movements an important determinant of entry patterns.

In summary the evidence is consistent with the predictions of a network based model. Moreover, it is inconsistent with a number of alternative explanations. Of course we cannot definitively rule out the possibility of some non-network explanation, but we are unaware of any such specific hypothesis.

The remainder of the paper discusses the role of community networks in subsequent evolution of the cotton textile industry after 1890, as well as the jute industry after the First World War. In the latter case we provide similar quantitative evidence: firms exhibited high degrees of clustering by community for newly entering Indian entrepreneurs, and the dynamic network model successfully predicts the evolution of community presence. Section 2 describes the historical background of the emergence of industrial entrepreneurship in India during the 19th century, and the role of various communities in Western India that played a role in the development of the cotton textile industry. The theoretical model of networks is presented in Section 3, followed by data and descriptive statistics in Section 4 and empirical results in Section 5 for the cotton industry. Section 6 provides corresponding evidence from the jute industry after the First World War, while Section 7 concludes.

2. Emergence of Indian entrepreneurship in the cotton industry

Industry, period and location

In the mid-1850s, the first cotton textile firm was set up by an Indian entrepreneur in Bombay. However, further entry was delayed for at least another decade, until the mid-1860s when the US Civil War ended. As shown in Fig. 1, during the US Civil War period the sharp increase in raw cotton prices on the world market was accompanied by a corresponding increase in entry into the upstream cotton sector (mostly devoted to exporting raw cotton bales).[6] The collapse of raw cotton prices when the war ended was followed by a stagnation in the number of entrepreneurs in the upstream industry. Fig. 2 shows a steady increase in entry into the downstream sector (producing cotton yarn and textiles from raw cotton, mainly aimed at selling mainly in the domestic market). In 1870, there were ten cotton mills in Bombay employing over 8,000 workers, which rose to 70 mills employing nearly 60,000 workers by 1890 (Morris, 1965).

The capital required to set up a textile mill was substantial. Table 2 shows that capital per firm in downstream industry was more than twice as large as the capital in upstream industry during 1860–90. The upstream industry comprised of mainly ginning and baling of raw cotton, and was technologically far less sophisticated than the downstream mills. For these reasons, we will focus on entry into the downstream industry, and treat the upstream industry as a related pre-industrial trading activity. The substantial capital required to set up a firm in the downstream may explain why growth in entry was gradual and took almost a decade after the end of the Civil War to accelerate further.

As the use of machinery and capital requirements were substantially larger in the downstream sector, we shall restrict our focus to this sector in this paper.

The downstream cotton industry was mostly located in the Bombay region (consisting of Bombay city and surrounding areas that belong to the current-day state of Maharashtra). Fig. 3 shows that in 1888, Bombay Presidency (which included the Bombay region in addition to areas near Ahmedabad, lying in current-day Gujarat) contained over 80 cotton mills, of which 60 were located in Bombay city alone. The rest of India altogether accounted for less than 40 mills, with the other two Presidencies (Calcutta and Madras) accounting for less than 10 each. To avoid confounding issues from difference in locations, we shall therefore restrict attention to the Bombay region, which excludes areas such as Ahmedabad located in present-day Gujarat. Fig. 4 shows that Bombay city accounted for nearly 90% of workers and 75% of firms in Bombay Presidency between 1865 until 1890. Thereafter its share declined steadily to 65% of workers and less than 50% of firms by 1910, owing mainly to increased entry in the Ahmedabad region accompanied by relative stagnation in the Bombay region (see Fig. 9 below). Hence 1865–1890 can be viewed as the first 'growing' phase of the Bombay cotton textile industry. We shall provide additional reasons for restricting attention to this period later in the paper.

[6] The stock incorporates both entry and exit of entrepreneurs. We explain later in the paper how the data on the stock of entrepreneurs and prices were constructed, with further details provided in the Appendix A.

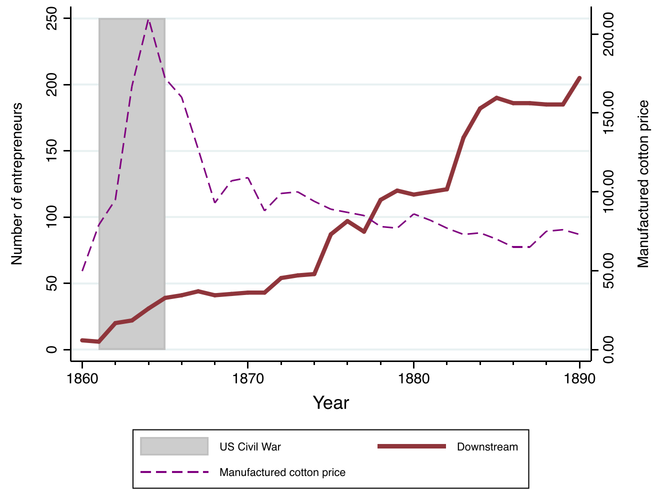


Fig. 2. Stock of Entrepreneurs in the Downstream sector (1860–1890)
Source: Entrepreneur stocks: own data; prices: see Appendix A.

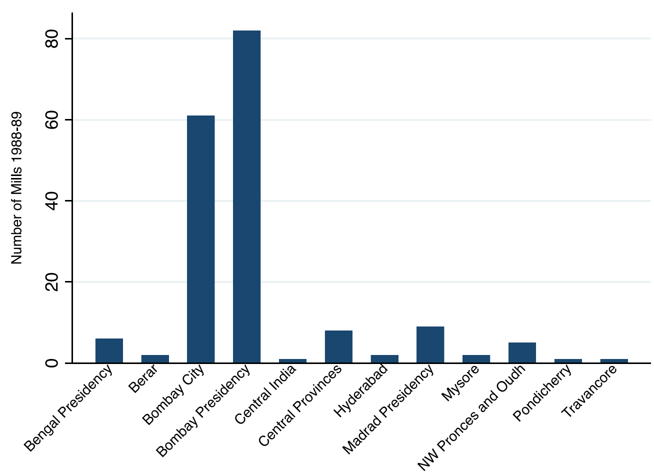


Fig. 3. Location of Textile Mills in India, 1888
Source: Statistical Tables Relating to Indian Cotton: Indian Spinning and Weaving Mills. Times of India, Bombay 1889.

Role of the state

In 1858, India came under Crown rule. This ended a 100-year rule of the East India Company and formally integrated India into the free trade regime of the British Empire. Under Company rule, the share of revenue from trade taxes had been small. British exports to India incurred tariffs of 5% on cotton piece goods and 3.5% on yarn. In 1859, this increased to 10% on piece goods and 5% on yarn as revenue concerns overrode the free trade argument (Harnetty, 1972, p. 7–10) and was strongly resisted by the Manchester Chamber of Commerce. Under Crown rule, the British state in India was receptive to political lobbying by British interest groups. By 1862, tariffs on cotton textiles had been lowered to the previous level (Harnetty, 1972, p.26). When Manchester demanded compensating excise on Indian products on the ground that the 5% tariff was contrary to the principles of free trade, the government conceded. Hence for the post-1865 period that we study, tariffs for revenue purposes were combined with a countervailing excise tax on import competing goods, effectively equivalent to free trade. The early development of the Indian cotton textile industry was therefore not aided by protective tariffs as in the case of early industrialization experiences of USA, Germany and Russia.

The colonial government was frequently lobbied by Lancashire not only on matters of tariffs, but also factory legislation. Lancashire

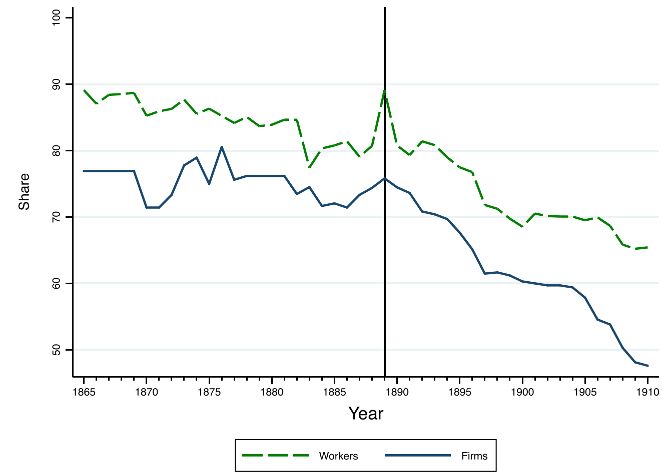


Fig. 4. Cotton Textile Industry Share of Bombay City in Bombay Presidency, 1865-1910.
Source: Report of Bombay Millowners' Association, 1925.

wanted adoption of similar regulations and labor laws as the British Factory Acts for the industry in India. On policies that affected the industry adversely, the Bombay Mill Owners' Association (BMOA) which was set up in 1875, was united in its opposition. The colonial government ended up passing the Indian Factories Acts of 1881 and 1891 which regulated child labor, work hours and working conditions (Upadhyay, 1990).

Weaknesses in market institutions

Industrial entrepreneurs and traders faced problems of contract enforcement, as legal institutions in India in the middle of the 19th century did not have a civil code for enforcing contracts. During Company rule, problems related to contract enforcement appeared in different contexts: from the procurement of textiles from weavers to cultivation of opium and indigo, the East India Company adopted ad hoc solutions using customary systems and in some cases coercion. The Indian Contract Act was passed in 1872, but as with other legal interventions, it did not create an institutional setting similar to that in Britain. Consequently, Indian traders, creditors, and manufacturers largely continued to rely on pre-existing community norms and institutions that were outside the structure of formal British-Indian law (Roy and Swamy (2016), Chapter 7).

Equally if not more significant were weaknesses of the capital market. Although there was a formal banking sector and the Indian trading communities had shares in banks, the latter did not provide long term capital to industry. Bank lending was limited to short term working capital (Bagchi, 1987, p. 110). Hence the capital market was the most important source of finance for investing in fixed capital. The colonial administration undertook a series of legislations to facilitate the development of a capital market. An 1850 Act enabled registration of unincorporated partnerships, which was extended in 1857 and 1860 to recognize limited liability joint stock companies, and subsequently amended in 1887 to incorporate priority of debts during insolvency proceedings. Moreover, the Bombay stock exchange was established in 1875, and high dividends of 10%–15% were offered by textile firms (Rungta, 1970, p.158).

Yet, despite these developments, the capital market failed to achieve any significant diversification of ownership. The business historian Tripathi (2004) writes:

> "The tight hold of the mercantile interests on the sources of industrial finance could have been loosened, had there been an efficiently functioning stock market and an alternative source of credit, such as banks organized on modern lines. Such institutions were woefully underdeveloped at the turn of the century".

Part of the problem stemmed from weaknesses in company law (Roy and Swamy (2016), Chapter 8). The key lacunae arose with respect to its failure to incorporate 'managing agency' contracts which was mainly a South Asian innovation. The managing agency contract was drawn up between a company and a 'managing agent' firm, wherein the latter was appointed to manage the company in return for a fixed or variable commission. This was not recognized as a legal document until the passage of the 1913 Companies Act. In practice, Bombay area firms which were mostly owned by Indian families used the managing agency contract as a means of maintaining close control of the company, while raising capital from the stock market under provisions of limited liability. As the law could not prevent such practices during the 19th century, it heightened problems of moral hazard associated with separation of ownership and management when these firms sought to raise capital from the stock market:

> "Possibly because the managing agency was an Indo-European system, English law did not supply ready guidance on the contract, and because it was poorly regulated, distortions arose and persisted for a long time. The contract could be one-sided. The agent could gain almost complete and near-perpetual control (some contracts were valid for sixty years) over a cluster of companies with which it had management contracts. This was not a problem for the shareholders when ownership and management were vested in one family, but it was a problem after public shareholding expanded.... The upshot was that the shareholders could do little to stop the mismanagement of the firm they owned when the agency firm was in the hands of a bad manager... In Bombay... (T)he owner had a controlling stake in the company and appointed a board of directors from friends and relations. The board appointed an agent, which was a firm belonging to the owner. The owner, in short, exercised control over a company by three means, shareholding, a compliant board, and the agency contract....(T)he landmark case ...(*Nusserwanji Merwanji Pande v. Gordon*, 1882) made the managing agent's authority more absolute than that of the board of directors. The judge ruled that in matters of direct conflict, where service to the company's interest was not in serious question, the agent's wish would prevail". (Roy and Swamy, 2016, p. 152–4)

It is therefore not surprising that under these circumstances, owners had a hard time convincing the general public to purchase shares in their company:

> "The typical industrial concern around 1900 was a company with shareholding of family and friends as well as the public, and managed by another firm, which was either a partnership of a company closely held by the same family or, on rarer occasions, a trust. The company had a legal identity as a public body, but it was managed like personal property.... Shareholding was usually confined within small groups of people known to each other". (Roy and Swamy, 2016, p. 146–7)

Community-based segmentation

Owing to these problems, the scope of widening ownership was restricted beyond immediate family and friends to the community to which the entrepreneur belonged. India has been historically characterized by a high degree of social segmentation between different castes, sub-castes, religions, ethnic groups and nationalities. Membership of these groups is defined by birth, with strict social norms restricting marriage and social interaction only within groups, and also featuring clustering into distinct locations and occupations (Ghurye, 1961). Occupational clustering arose as a result of transmission of information and skills within these groups:

> "...one important feature of caste: membership of a caste makes a person part of a person-based social network which controls insiders' information about economic opportunities, transmits skills and provides varied types of human and material support". (Panini, 1996, p. 39)

The British colonial authorities officially recognized these social groupings: e.g., the 1881 and 1891 Censuses used *jati* or sub-caste as a basis for classification of the population, and used it to determine eligibility of different groups to jobs in the colonial administration. (Beteille, 1996) argues that the term *jati* being restricted to different groups within the Hindu population is too narrow, and that a better representation of the Indian concept of *jati* in the English language is 'ethnic group', a term broad enough to include different tribes, sects, religious and linguistic minorities, and nationalities. In what follows, we shall use the term 'community' as a shorter version of 'ethnic group'.

The early cotton entrepreneurs were segmented into a number of different communities. The five principal communities, included the Parsis, a small group of Zoroastrians who had migrated from Persia from the 8th century, Hindu and Jain Vanias, the Muslim merchant communities of Bohras and Khojas, Jews who had migrated from Baghdad in the 18th century and the Bhatias, a small Hindu sub-caste. The Hindu and Jain Vanias had a high degree of economic and social interactions (including inter-marriage), and will thus be treated as a single community ('Vanias') in the rest of the paper.

Social interactions were highly segmented across these different communities, with corresponding implications for segmentation of economic relationships. For instance, Davar who started the first textile firm raised capital mainly from his own Parsi community. When Ranchhodlal (who did not belong to a trading caste) tried to set up a textile firm in Ahmedabad, local merchants were not willing to invest in his firm (Mehta, 1991, pp 182–3). Similarly, when Tata, a Parsi, offered shares of the first cotton mills outside Bombay city to a prominent Marwari trader, the response was negative (Tripathi, 2004, p. 121). Even for a late entrant in 1897 such as Lalbhai from the Jain community, investors from this community were the main source of capital (Tripathi and Mehta, 1990, p. 90). The capital market thus remained highly segmented despite the presence of a stock exchange. Later in the paper we will provide further evidence of the high degree of concentration by community in the composition of principal shareholders of different cotton mills.

The main theme of our paper is that these community-based means of support formed a substitute for weaknesses in underlying market and state institutions, wherein entrepreneurs were forced to rely largely on connections within their own communities for necessary capital, managerial skills, information and technical knowhow. The model that we develop in the next section to depict the nature of such community-network based entrepreneurship, is based on the assumption that these different communities were broadly similar in terms of necessary pre-conditions: pre-industrial trading opportunities, experience, wealth and connections with the colonial authorities. We now turn to a description of the historical evidence relating to this.

Inter-community differences in wealth, literacy and size

A challenge to identify the role of community networks in facilitating entry into industry is the problem of distinguishing trends resulting from a dynamic network effect from trends in underlying community characteristics. If some communities had privileged access to capital or enjoyed certain benefits due to their connections with the British in India, and these were growing at rates positively correlated with initial entry patterns, it would confound the effect of the dynamics generated by the social network itself.

It is undeniable that the communities differed in size and education. In particular, the Parsis were exceptional in their educational attainment. In 1881, over 70% of the Parsis were literate and 50% of those literate had secondary education, comparable to the literacy of the Europeans in India. 5% of the literate Parsis and 4% of the literate European Christians had a college education. Other communities did not have the same literacy rate. The Vania trading groups had primary education, but a very small proportion went to college. Table 3 shows the differences in literacy among the different communities.

Table 3
Literacy by community in Bombay Presidency (%).
Source: Censuses of India.

	1881	1891	1901
Hindu[a]	10.9	11.8	11.2
Parsi	72.9	76.3	75.2
Jain	51.5	53.4	48.9
Jewish	52.0	54.2	–
Muslim[b]	6.8	9.2	7.3
Christian	38.9	44.4	36.5
Total Bombay Presidency	11.1	12.6	11.67

[a]Hindus include Bhatias and Vanias. The 1911 Census shows male literacy for Bhatias was 56% and Vanias 60%.

[b]Within Muslims, the 1911 Census shows male literacy rate of 41% among the trading communities of Muslim Bohras, Khojas and Memons.

Table 4
Population share in Bombay city by community (%).
Source: 1872–1891: The Gazetteer Bombay City and Island 1909. 1901 data comes from the Census.

	1864	1872	1881	1891	1901
Hindu[a]	64.1	63.4	65	–	65
Bhatia	2.7	1.5	1.2	–	1.2
Parsi	6	6.8	6.3	5.8	6
Jain[b]	1	2.3	2.2	–	1.8
Jewish	–	0.4	0.4	–	0.7
Muslim[c]	17.8	21.4	20.4	18.9	20.1
European Christian	1	1.1	1.4	–	1.5

[a]Includes Bhatias.

[b]Includes Buddhists.

[c]Includes Bohras and Khojas.

However, *the key point to note is that literacy rates did not materially change over time for any community.* Nor do we see differential growth in population in Bombay city (Table 4) where most of the cotton textile industry was located until the 1890s. Accordingly, in our empirical analysis we will allow for level differences in relevant community characteristics with community fixed effects, and rely on the identification assumption that there were no significant differences in time trends of these characteristics.

Pre-industrial trading activity of various communities

All five communities were active in trading during the 18th century, starting as agents of the East India Company: Hindu Vania traders were the main contact of the Company in Ahmedabad, while the Parsis, the Bhatia and Bohra and Khoja Muslims were more important in Bombay (Das Gupta, 2001). Table 5 shows their respective shares of non-European trading capital in the Western town of Surat, a major location for trade and local industry. Some of these groups migrated overseas to pursue trading opportunities, with the exception of the Vanias who were restricted by their religious code from traveling abroad. The Parsis were involved in shipping and external trade, engaging particularly in the opium and cotton trade to China. Muslim trading groups of the Khojas and the Bohras operated in the Indian Ocean trade between East Africa, India and East Asia, and later in the opium trade to China. For the Vanias, the main channels of wealth accumulation was internal trade, informal banking and moneylending. They financed the East India Company's interregional trade within India, especially between Bengal and the Bombay regions.

The opium trade was a main source of wealth accumulation for Indian traders and bankers (Farooqui, 2005, Chapter 8). Farooqui shows that out of 120 trading companies involved in the opium trade 1800–1830, 49 belonged to the Parsis. They were the largest group, but Marwaris, Gujarati Vanias and Konkani Muslims were also involved at various stages of the opium trade. A few Parsi firms had partnerships with members from other communities. By 1844, a quarter of the merchants were Parsis and there is evidence of cooperation rather than conflict among the trading companies in the opium trade (Farooqui, 2005, Appendix III).

Table 5
Non-European trading capital 1746 in Surat: Shares of communities (%).
Source: Guha (1984).

	European protection	Independent	Total
Jew	1.2	0	1.2
Parsi	9.9	0	9.9
Hindu[a]	18.8	30.5	49.3
Muslim	2	31.7	33.7
Total	31.9	68.1	100

[a]Refers primarily to Bhatias and Vanias.

Overseas trading opportunities however greatly shrunk owing to a number of external shocks in the first half of the 19th century. The Opium War in China and the decline in the monopoly of the East India Company over trade with China resulted in Parsi merchants looking for suitable alternatives to invest their wealth. The second major shock resulted from rising competition from European ships as a result of improvements in shipping technology, resulted in reduced opportunities for Indian traders in the Indian Ocean and China trade. These communities turned to exports of raw cotton, which was produced in abundance in Western India as the hinterland of Bombay city was India's cotton producing region. Raw cotton from the region was sold in the regional markets and exported to Europe in periods of high demand. The US civil war was one such period, which witnessed a big rise in export of raw cotton from India to Britain, as India replaced the United States as a source of raw cotton for the British textile industry. This provided a big impetus for investment in cotton pressing and baling, preparing raw cotton for export. The end of the civil war marked yet another shock to the cotton traders as the export trade in raw cotton became less profitable. Moreover, a pre-existing ban on export of British textile machinery was lifted. This created the conditions for emergence of a new wave of entrepreneurs in the Bombay area into the downstream cotton industry after the end of the US Civil War, which aimed to produce cotton yarn and textiles for the domestic Indian market in competition with foreign imports (mainly from Britain).

The experience of the Vania traders was markedly different, as they had not participated in overseas trade previously: their principal outside option was trading and money lending within India, where returns remained high until the late 1870s when new legal changes affected moneylending. The peasant riots against moneylenders in 1875 in the Deccan region of Bombay Presidency, were targeted to destroy debt records held by money lenders (Kranton and Swamy, 1999). Following the riots, the government became increasingly concerned about agrarian distress from indebtedness. The Deccan Agricultural Relief Act of 1879 which imposed restrictions on lending to farmers was the first of a series of resulting legal reforms. Consequently, industry became more attractive as a contender for investment for Vanias only after 1880, and they entered in large numbers from the 1890s onwards in both Bombay and Ahmedabad regions. Hence the Vanias were relative latecomers to the (downstream) cotton textile industry. As the size of their population was substantially larger than the other communities active in cotton in Bombay, they came to dominate the industry. The assumption of parallel trends in outside options would therefore be violated if we incorporated the post-1890 period. This constitutes the main reason for truncating the period we study between 1865–90. Further discussion of the post-1890 period is provided in a later section.

Connections with the colonial government and banks

As already explained above, the British colonial authorities pursued a laissez faire policy with respect to trade from 1862 onwards, thus removing the scope for preferential tariffs. However, this still left scope for communities to have different levels of connections with the colonial authorities, and obtain selective treatment from policymakers. The historical evidence shows, however, that all the five main communities were actively involved in lobbying activities and in access to bank credit.

The Bombay Mill Owners Association was set up in 1875 to coordinate the interests of the textile mills. The annual meetings discussed market conditions and their stance on various policies of the British government ranging from industrial regulation, tariffs and other taxes that impacted the industry. The association held annual meetings and its deliberations were published in annual reports, which were reported in the Times of India newspaper. These records reveal that members of different communities attended these meetings, and were members of various committees. For instance, Dinshaw Petit, a Parsi, was the founding member of the BMOA and chairman of the association in the early phase. British entrepreneurs George Cotton and James Greaves, who had been involved in the cotton trade, also held this position. Among the attendees in 1877 were Tapidas Vurjidas and Morarjee Golculdas (Vanias), Manchorjee Banajee and Nanabhoy Jeejeebhoy (Parsis), Damodar Thackersey (Bhatia), David Sassoon (Baghdadi Jew) and Fazulbhoy Curimbhoy from the Khoja community. All of them were active members, proposing motions and raising complaints. The re-election of Dinshaw Petit as Chairman in 1887 was proposed by a Parsi, Nusserwanjee Manockjee Petit and seconded by a Hindu Vania, Damodar Tapidas. Sassoon, Curimbhoy and Manmohandas Ramjee went on to become chairmen after 1900.

The British Factory Commission, which was appointed in 1875 to look into the working conditions of the cotton mills, held meetings with factory owners in Bombay that year. Evidence on existing working conditions and wages were given by entrepreneurs, managers, engineers and workers across the industry and all communities.[7]

There is also no historical evidence suggesting that any particular community had better access to bank credit. The Bank of Bombay financed cotton traders in the early stages and later provided short term capital to the cotton mills (Bagchi, 1987, p.110). Members of all communities were involved in the management of the bank. Parsis, Bhatias, Baghdadi Jews, Hindu and Muslim merchants were on the board of directors between 1876 and 1920 (Bagchi, 1987, Appendix). Several cotton textile firms belonging to different social groups enjoyed a high credit limit with the bank discounting bills or borrowing against goods (Bagchi, 1987, p. 337–8). They all relied on credit offered by British machinery exporting firms.

Absence of inter-community conflict

A remaining assumption of the model is that the dynamics of entry in each community followed an independent trajectory, apart from exposure to common time-varying shocks to the industry or factors common to the Bombay city region. This rules out the possibility that the entry of entrepreneurs from some communities was actively deterred by other communities, owing to the existence of inter-community tensions or conflict. The historical evidence does not show any indication of such conflict. For instance, members of different communities attended BMOA meetings and participated in relevant committees. Minutes of the meetings indicate cooperation rather than conflict, where issues of disagreement were openly debated and resolved. Members of different communities were united in their opposition to the excise duty imposed by government in response to tariffs on British imports, and to the Factory Acts regulating working hours and the participation of women and children. There were positive interventions too in the interest of the industry — e.g., in the annual meeting of 1887, the proposal to set up a vocational institute to train artisans for the industry received support across all the communities.

Finally, while our data on the composition of principal shareholders of firms shows significant clustering by community (Fig. 7 below), it also shows (Fig. 8) that no community was systematically excluded from membership of firms dominated by other communities.

3. Network-based model of entry dynamics

The model extends (Munshi, 2011) and is a special case of (Dai et al., 2018) with a single destination sector.[8] There are successive cohorts of potential (infinitely lived) entrepreneurs (or agents, in short) $t = 1, 2, \ldots$ who have a once-and-for-all opportunity to enter the downstream industry at date t. Their outside option is to continue pursuing their current occupation, e.g., trading or the upstream stage of the same industry. The agents belong to different communities $c = 1, \ldots, C$. Individuals within a community vary in their entrepreneurial ability ω, drawn from a log-uniform distribution whose mean varies across communities. Specifically, $\log\,\omega$ is distributed uniformly with mean $\log\,\omega_c$ on support $[\log \omega_c - \frac{1}{2}, \log \omega_c + \frac{1}{2}]$. The profit of an agent with individual ability ω in their outside option is $\Pi_c \omega^{\sigma}$ where $\sigma \in (0, 1)$.

Communities can differ in levels of underlying characteristics such as access to capital (borrowing cost r_c), mean ability (ω_c) and outside options (Π_c), which are assumed to be fixed over time.[9] The historical evidence described in previous sections suggests this is a reasonable assumption for the early stage of the cotton textile industry in the Bombay city region.

Entry decisions are irreversible: once a cohort t agent enters the downstream industry he stays there at all dates $t' > t$. It will be in the interest of every entrant to stay every period thereafter, so this assumption is not restrictive. Entrepreneurs are myopic and selfish: each cohort t agent decides whether to enter based on a comparison of his own profit upon entering the industry at date t, with his outside option. Extending the model to accommodate foresight makes it more complicated, while reinforcing further the network effects. Hence myopia is a useful simplifying assumption. Similarly, extension to incorporate some altruism towards fellow community members will also reinforce the network effects at the cost of complicating the analysis considerably.

Network effects in the downstream industry are represented by dependence of the total factor productivity (TFP) of any cohort t entrepreneur from community c on the presence of incumbents from the same community from all preceding cohorts $t-1$, denoted by n^c_{t-1}. A community c entrepreneur of ability ω will have a production function at date t:

$$y^c_t = A^c_t \omega^{1-\alpha} K^{\alpha} \quad (1)$$

where K denotes capital size chosen by the entrepreneur, $\alpha \in (0, 1)$ represents diminishing returns to capital, and the 'community-based TFP' (CTFP) is represented by

$$A^c_t = A_0 \exp\{\theta n^c_{t-1}\} \quad (2)$$

The underlying assumption is that incumbents from the same community share knowhow and help one another, to overcome the absence of government support and markets for knowhow. Mutual cooperation

[7] They included Parsis (Dinshaw Petit, Shavakshaw Dhunjeenroy, Nusserwanjee Dadabhoy, Muncherjee Nowrojee Banaji), Vanias (Mothiram Bhagubhoy, Rao Bahadur Becherdass, Tapidass Vurjdass, Morarjee Goculdass Assur Virjee, Munguldas Nathogbhoy) and Europeans (Joseph Sharpe, Henderson C. S. I. *T* Blaney, J. Helm, and F. Arbuthnot).

[8] The latter paper tests a more detailed version of the same model in the context of growth of privately owned firms in China over the period 1990–2009, using comprehensive firm registration data covering all industries. The richness of the Chinese data allows Dai et al. (2018) to test implications of community networks on entry, size, and concentration across sectors and locations, using a different econometric strategy relating firm dynamics to a measure of quality of the community networks. The relative sparseness of the Indian historical data rules out the use of a similar estimation strategy, so we rely instead on inferences based on variation in initial presence of different communities with subsequent entry flows.

[9] This can be further weakened to the assumption of parallel time trends across communities.

is supported by social sanctions involving suspension of social interactions by the community with deviants. Fresh entrants can benefit from these forms of intra-community cooperation; they do not contribute in the period that they enter but do contribute thereafter. Exchanges of different community members complement each other, giving rise to the exponential form (2) involving the size of the network. For instance, each entrepreneur relies on a given level of informal help $h > 0$ provided by other members to raise productivity or market access: $A_t^c = A_0(1+\gamma h)^{n_{t-1}^c}$, which reduces to expression (2) if we define $\theta \equiv \log(1+\gamma h)$. θ represents the strength of the network effect, manifested by a productivity spillover among network members. This formulation of spillovers differs from standard specifications of agglomeration spillovers in the economic geography literature, as the domain of the spillovers at a given location is restricted to entrepreneurs from the same community. In contrast, agglomeration effects are location-specific but do not vary by community origins of the entrepreneurs involved.

It could be argued that active collaboration is actually limited to social groups (based on family ties, physical proximity or other sources of social association) which are smaller than the entire community. In that case the definition of network size for any given agent should be the number of incumbents from the same social group. If there are k social groups within the community, $m_{t-1}^c \equiv \frac{n_{t-1}^c}{k}$ denotes the (average) number of incumbents from each group, and θ' denotes the strength of ties within the group, formulation (2) would be replaced by

$$A_t^c = A_0 \exp\{\theta' m_{t-1}^c\} \tag{3}$$

for a representative agent of community c, which reduces to (2) if we define $\theta = \frac{\theta'}{k}$.

The product price q_t at date t is the same across all communities. Each entrepreneur and community takes this price as given; we therefore abstract from collusive behavior. Conditional on entering, an entrepreneur of ability ω will enter with capital size $K_t^c(\omega)$ which maximizes profit $\pi_t \equiv q_t A_t^c \omega^{1-\alpha} K - r_c K$. Hence realized profits and capital size in the downstream industry, conditional on entry at t by a potential cohort-t entrepreneur with individual ability ω and belonging to a community with an incumbent network size of n_{t-1}^c, are given by

$$\log \pi_t^c(\omega) = \log \omega + \frac{1}{1-\alpha}[\theta n_{t-1}^c + \log q_t] - \frac{\alpha}{1-\alpha}\log r_c + \log \mu + \frac{1}{1-\alpha}\log A_0 \tag{4}$$

$$\log K_t^c(\omega) = \log \omega + \frac{1}{1-\alpha}[\theta n_{t-1}^c + \log q_t] - \frac{1}{1-\alpha}\log r_c + \frac{1}{1-\alpha}[\log A_0 + \log \alpha] \tag{5}$$

where μ denotes $[\alpha^{\frac{\alpha}{1-\alpha}} - \alpha^{\frac{1}{1-\alpha}}]$.

Entry is endogenous and incorporates a pattern of selection (analogous to the standard Roy model) based on entrepreneurial ability and network size as follows. A cohort t agent will decide to enter if $\log \pi_{t-1}^c(\omega) > \log \Pi_c + \sigma \log \omega$, i.e. if his individual ability ω exceeds the threshold $\underline{\omega}_t^c$ given by

$$\log \underline{\omega}_t^c = \frac{1}{1-\sigma}[\log \Gamma_c - \frac{1}{1-\alpha}\{\log q_t + \theta n_{t-1}^c\}] \tag{6}$$

where $\log \Gamma_c \equiv \Pi_c + \frac{\alpha}{1-\alpha}\log r_c - \log \mu - \frac{1}{1-\alpha}\log A_0$.

Owing to yearly inflows of fresh cohorts, network size n_{t-1}^c grows over time, while other community characteristics are fixed: hence the ability threshold (6) declines over time (assuming the product price is stationary). The implied dynamics of entry flows will then depend on whether the ability threshold is in the interior of the support of the ability distribution, which will be the case in early stages of the development of the industry. During the early stage, the threshold will be falling in t, implying that the entry flows will be rising. Once this early stage is over, there is no further scope for entry flows to rise: the threshold drops below the lower endpoint of the support, and *all* entrepreneurs in subsequent cohorts will enter.

We thus define the *early stage* of the entry process to be when the entry threshold is interior, i.e. the network size is still small:

$$n_{t-1}^c < \frac{1}{\theta}[(1-\alpha)\log \Gamma_c - \log q_t] \tag{7}$$

while the *mature stage* is characterized by the reversal of this inequality. The actual date of transition will of course depend on the realization of the price shocks, and is thus random. Moreover, the process can flip back and forth between the two stages. These complications will not arise if the product price does not fluctuate over time, which we abstract from to simplify the exposition.

The resulting entry flows are as follows. During the early stage, the proportion of each cohort that enters is:

$$e_t^c = \frac{1}{(1-\alpha)(1-\sigma)}[\theta n_{t-1}^c + \log q_t] + \delta_c < 1 \tag{8}$$

while during the late stage:

$$e_t^c = 1 \tag{9}$$

where $\delta_c \equiv \frac{1}{1-\sigma}\Gamma_c + \log \omega_c + \frac{1}{2}$ is a 'composite' community characteristic, which depends on mean ability, outside options and capital access of community c members.

The evolution of network size during the early stage is thus given by

$$n_t^c = n_{t-1}^c + e_t^c = M n_{t-1}^c + P \log q_t + \delta_c \tag{10}$$

where $M \equiv 1 + \frac{\theta}{(1-\alpha)(1-\sigma)}$ and $P \equiv \frac{1}{(1-\alpha)(1-\sigma)}$. M is the *network multiplier parameter* which exceeds 1 in the presence of network effects, and equals 1 otherwise. This parameter represents the incremental effect of higher network size in any cohort on the network size in the succeeding cohort.

In the mature stage, the dynamic of network size is

$$n_t^c = n_{t-1}^c + 1 = n_T^c + t - T \tag{11}$$

if T denotes the last date of transition into the mature stage.

We summarize the predicted network dynamic in the following Proposition.

Proposition 1. *During the early stage (assuming (7) holds at dates 1, …, t), the evolution of network size of community c given initial cohort size n_0^c is given by*

$$n_t^c = M^t[n_0^c + \frac{\delta_c}{M-1}] - \frac{\delta_c}{M-1} + P\sum_{k=0}^{t-1} M^k \log q_{t-k} \tag{12}$$

in the presence of network effects ($M > 1$), and

$$n_t^c = n_0^c + \delta_c.t + P\sum_{k=0}^{t-1} \log q_{t-k} \tag{13}$$

when they are absent ($M = 1$).

In the mature stage, network size follows (11) instead.

Network effects result in greater sensitivity of network size in later cohorts to initial network size during the early stage of the development of the industry: $\frac{\partial n_t^c}{\partial n_0^c} = M^t$ which is rising in t when $M > 1$. A network multiplier operates across successive cohorts: one more member raises the number of members at the next cohort by M, the one after that by M^2 and so on. Initial differences in network size across communities will be progressively magnified across time, a pattern of divergence that grows exponentially over time. In the absence of any network effect, this divergence can grow, if at all (i.e if the community characteristic δ_c is positively correlated with initial network size n_0^c), at a linear rate. In particular, the (community-specific) time trend for network size will be linear in the absence of network effects, and exponential in the presence of network effects.

During the mature stage, this difference disappears, as time trends become linear even in the presence of network effects.

Proposition 1 thus generates a strategy of testing for network effects, using the early stage of the development of the industry, under the assumption that relevant community characteristics (ability, access to capital or outside options) summarized by δ_c are fixed over time. Taking a quadratic approximation for $M^t = 1 + t\zeta + t^2\frac{\zeta^2}{2}$, (12) generates the following regression specification for dynamics of network size

$$n_t^c = [1+\zeta t+\frac{\zeta^2}{2}t^2]n_0^c + \frac{1}{M-1}[\zeta t+\frac{\zeta^2}{2}t^2]\delta_c + P\sum_{k=0}^{t-1}(1+\zeta k+\frac{\zeta^2}{2}k^2)q_{t-k} \quad (14)$$

in the presence of network effects ($\zeta > 0$). When network effects are absent, the regression is given by (13) instead, where community-specific quadratic time trends do not appear. Hence the significance of community-specific quadratic time trends, interacted either with initial network size n_0^c or community characteristic δ_c, indicates the presence of network effects. Since initial network size and δ_c are perfectly collinear, we cannot identify their effects separately. But this is unnecessary to show evidence of a network effect, i.e., that ζ is positive. The significance of quadratic time trends interacted with either n_0^c or δ_c is sufficient to infer the existence of a network effect. Intuitively, they represent a pattern of divergence between the size of different communities among incumbents that is growing faster than can be represented by differential linear time trends.

3.1. Misallocation, adverse selection and entering capital size

While community-based cooperation facilitates entry and thereby hastens the pace of industrialization, its scope is restricted to members of specific communities and excludes others. This would be reflected in growing divergence in the presence of different communities among incumbents during the early stage of development. This is in marked contrast to market or state-sponsored industrialization patterns whose scope is not restricted to particular social communities. As we now explain, the growing imbalance in community composition of entrepreneurs will be accompanied by a form of misallocation, where incumbents from favored communities with larger networks will tend to be of lower productivity. This reflects a pattern of adverse selection generated by the network-based mechanism, during the early stage of development. This pattern helps distinguish the network based model from competing explanations based on positive selection (i.e., positive correlation between initial levels and trends in community-specific ability or wealth).

Observe that the log of the productivity P_t^{mc} of a marginal entrant from cohort t of community c equals

$$\log P_t^{mc} = \log A_t^c + (1-\alpha)\log \underline{\omega}_t^c = \log A_0 - \frac{\theta\sigma}{1-\sigma}n_{t-1}^c - \frac{1}{1-\sigma}\log q_t + \frac{1-\alpha}{1-\sigma}\log \Gamma_c \quad (15)$$

during the early stage. This implies that marginal entrants (among any given cohort) from communities with a higher network size have lower productivity. The lower individual ability of the marginal entrant outweighs the larger community component of productivity. Intuitively, this is because the lower threshold for individual ability implies that the outside option (and hence equilibrium payoff) of the marginal entrant is lower. It follows that there is misallocation of talent in the economy: average productivity in the industry would rise if it were possible to replace the marginal entrant from a community with a larger network, by a marginal non-entrant from a different community with a smaller network (i.e., who did not actually enter but was almost indifferent between entering and not).

The misallocation appears even when comparing the average productivity of **all** entrants (rather than marginal entrants) across communities of varying network size if the parameter σ (elasticity of the outside option with respect to ability) exceeds $\frac{1}{2}$, since the log of the productivity P_t^{ac} of an entrant from cohort t of community c equals

$$\log P_t^{ac} = \log A_t^c + \frac{1-\alpha}{2}[\log \omega_c + \frac{1}{2} + \log \underline{\omega}_t^c]$$
$$= \log A_0 + \frac{1-2\sigma}{1-\sigma}\theta n_{t-1}^c + \frac{1-\alpha}{2}[\log \omega_c + \frac{1}{2} + \frac{1}{2(1-\sigma)}\log \Gamma_c] - \frac{1}{2(1-\sigma)}\log q_t$$

However, in the opposite case where $\sigma < \frac{1}{2}$, entrants into larger network communities are of higher average productivity. Hence adverse selection for average entrants is a stronger form of misallocation which may or may not arise.

How can we test for such forms of adverse selection? Note the close connection between entering capital size and productivity P of an entrant:

$$K(P;r_c) = [\frac{\alpha P}{r_c}]^{\frac{1}{1-\alpha}} \quad (16)$$

If access to capital varies across communities, this makes it difficult to test for misallocation directly by comparing entering capital sizes in different communities. Instead we can test for adverse selection by comparing time trends in entrant capital size, since according to the model these reflect changes in network size while capital access is unchanging over time within any community. To elaborate on this, observe that among the entrants within a given community, those entering with a smaller capital size are of lower productivity. Hence we can identify the marginal entrant in a community-cohort pair, by the entrepreneur with the smallest entering capital size. The model predicts the smallest entering capital size from a given community c at any given date t, is decreasing in the network size n_{t-1}^c, and therefore in the determinants of this size (which has been provided in Proposition 1). In particular, it predicts that *the smallest entering capital size in a given community will be decreasing in the entry date t, that this declining time trend will be steeper for communities with higher initial presence* n_0^c, *and finally that this tendency will accelerate over time (owing to the accelerating growth of network size)*.

In contrast the opposite hypothesis of positive selection combined with absence of network effects will predict that communities with larger initial presence owing to superior ability or wealth will enter with larger average size, while marginal entrants will enter with the same size (since the minimum ability threshold (6) for entry is independent of the ability distribution in the absence of network effects). Moreover communities with incumbent stocks of entrepreneurs that grow faster must be experiencing faster growth of ability or wealth, hence entering average sizes of later cohorts must also grow faster. These contrasting predictions allow us to discriminate empirically between the two hypotheses.

4. Cotton industry: Data and descriptive statistics

Our data is compiled from business directories, which list all registered firms in the industry in the region and the names of directors and partners. The firm level data used in the paper comes from the Bombay Almanac, Times of India Calendar and Directory and Thacker's Indian Directory; see the Appendix for further details and an example of the firm level information. We combine the three sources to create a data set of industrial entrepreneurs, who can be linked across firms and over time from pre-industrial sectors to the cotton textile industry. All subsequent figures and tables are based on this data.

Each entrepreneur is either a director or a partner in a listed firm. We construct a data set at the level of the entrepreneur using names of partners and directors. The data include all entrepreneurs in registered firms in Bombay Presidency (Bombay and Ahmedabad regions). We observe them at the point of entry, and track their previous and subsequent trajectories between 1860 and 1910. The data includes some characteristics of the firm that the entrepreneur is listed in: the year of establishment, its location and in many cases paid-up capital. From the names of the entrepreneurs, we can identify the community they belong to. We construct a panel data set of entrepreneurs by community. The communities are Parsi, Bhatia, Vania, Baghdadi Jews

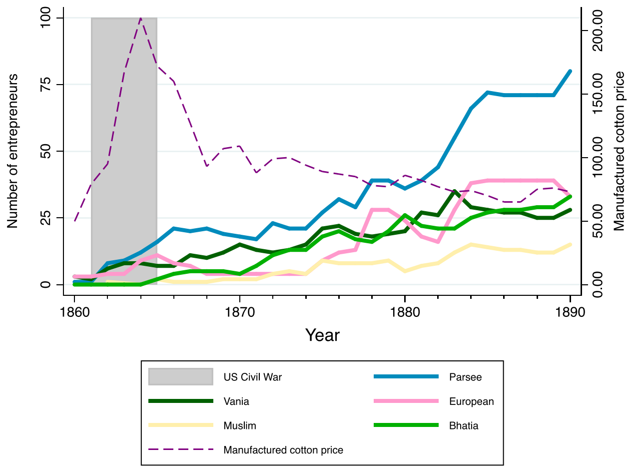


Fig. 5. Stock of Entrepreneurs by Community in Downstream Cotton, Bombay 1860-1890.
Source: Entrepreneur stocks: own data; prices: see Appendix A.

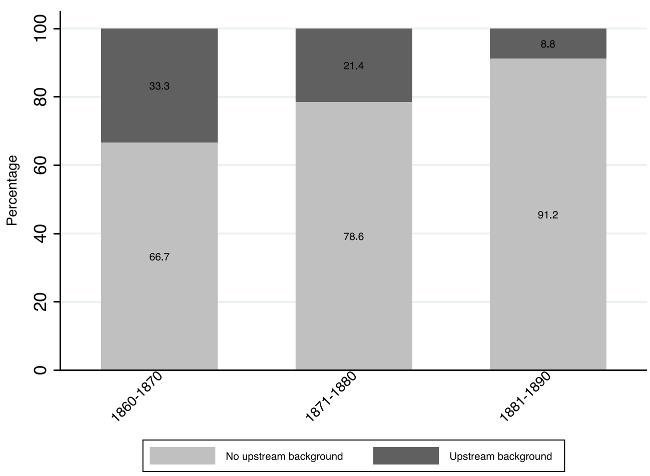


Fig. 6. Upstream experience of entrants into downstream Bombay cotton 1860–90.
Source: Own data. The figure shows the percentage of entrepreneurs entering the downstream industry in Bombay between 1860–90 who respectively were and were not listed in any previous year in an upstream cotton firm.

Fig. 7. Community clustering of entrants into downstream Bombay cotton 1860–90.
Source: Own data. 'Same community' share equals the proportion of newly entering entrepreneurs belonging to the community with the largest intra-firm share in that year, averaged across downstream firms in Bombay listed during each of the time periods.

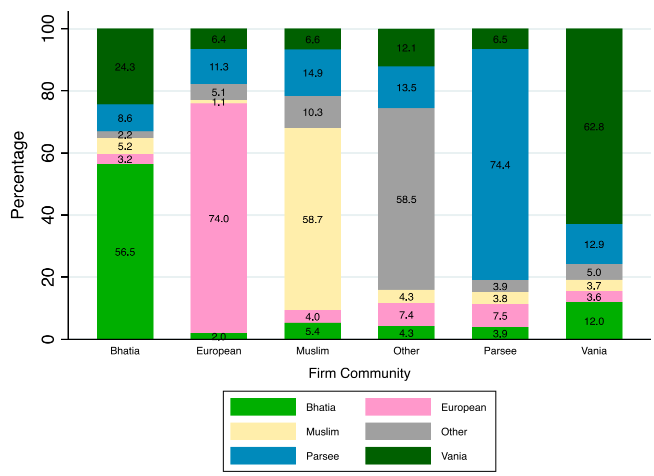


Fig. 8. Intrafirm Distribution of downstream Bombay cotton entrepreneurs by community, 1860–1890.
Source: Own data. Downstream Bombay cotton firms grouped by community with largest intra-firm share of entrepreneurs. For each group, the figure provides percentage of entrepreneurs belonging to different communities, averaged across all firms in the group.

and Khoja Muslim. The reference community is European, primarily British.

As explained earlier, we focus on entrepreneurial presence in the Bombay region between 1865–90. Fig. 5 shows the evolution of the stock of entrepreneurs in the Bombay downstream industry by community in a given year between 1860 and 1890. It reveals the dominance of the Parsis at the initial stage and a pattern of non-linear divergence in the presence of different communities. It also shows that rising entry was not accompanied by rising prices of manufactured cotton.

The role of prior experience in pre-industrial trade in entrepreneurship has been highlighted in the recent 'enterprise map' studies from some contemporary African countries, which suggest pre-existing upstream experience is an important determinant of entry into downstream industry (Sutton and Kellow 2010; Sutton and Kpentey 2012). To check the relevance of this in the current context, we calculate the proportion of downstream entrants who had earlier participated in the upstream industry. Fig. 6 shows that less than 35% of the downstream entrants between 1860–70 had prior upstream experience. This proportion declined further in the two subsequent decades. By 1890 less than 10% of the entrants had prior upstream experience.

Next, we examine clustering of entrepreneurs by community within firms. For each firm in any given year we compute the share of different communities in the stock of incumbent entrepreneurs. Then we calculate the 'same community' share of entrants, i.e., the fraction of entrepreneurs that newly entered the firm that year and belonged to the community with the largest share among incumbents. Fig. 7 shows the resulting average across firms and years in the three different decades between 1860–90. In each decade, over 60% of all entering entrepreneurs belonged to the community with the largest incumbent intra-firm share. Moreover, this proportion rose to 71% during the decade 1880–90, indicating continued importance of community links throughout successive decades, despite the growth of the stock market.

However, segregation of firms along community lines was partial rather than total. Fig. 8 provides the community-wise breakdown of entrants from other communities (i.e., different from the dominant community among incumbent entrepreneurs in any given firm-year). For each group of firms with a specific dominant incumbent community, the Figure provides the fraction of newly entering entrepreneurs

Table 6
Stock of cotton entrepreneurs in Bombay at the end of the US Civil War (1865).
Source: Own data.

Community	Upstream	Downstream
Bhatia	9	2
European	41	11
Vania	9	7
Muslim	2	2
Parsi	26	16

belonging to every other community, averaged over the entire period 1860–90. Every community had some representatives in firms dominated by other communities.

We construct a second panel of paid up capital per entrepreneur at the community level, using firms at entry point only. The variable capital per entrepreneur is constructed by using firm capital at the point of entry and dividing this by the total number of entrepreneurs listed as directors or partners in the firm. This is used in the next section to test for adverse selection. Unfortunately reporting of capital size during post-entry years was not systematic, so we could not construct a reliable data set on post-entry growth rates of firm size.

5. Regression results for cotton entry

5.1. The early phase 1866–1890

Network size regressions

The regression analysis focuses on the early period in the development of the downstream industry in Bombay city, following the US Civil War years. We test the model developed in Section 3 using data for the downstream industry over 1866–1890. The initial presence of each community (date 0 in the model) is represented by number of entrepreneurs already present in the downstream industry in 1865. The regression also tests the role of prior upstream experience and opportunity to accumulate profit during the Civil War years, measured by 1865 upstream community presence.

Table 6 shows the presence of different communities in the upstream and downstream industry during the Civil War years. There were 38 entrepreneurs downstream in 1865, dominated by Parsis, followed by European and Vania communities, and also featuring a small number of Khoja Muslims and Hindu Bhatias.

Column 1 of Table 7 shows results for the following regression specification based on (14), run for the period 1866–1890:

$$n_t^c = \phi_c * t^2 + \delta_c * t + \gamma_c + \nu_t + \epsilon_t^c \quad (17)$$

where $\phi_c, \delta_c, \gamma_c$ denote community c dummies, n_t^c denotes number of incumbent entrepreneurs from community c in year t, and $t = 1$ corresponds to year 1866. ν_t is a year dummy representing (current and lagged) price or other industry level shocks, and ϵ_t^c is an error term. This is a particular version of Eq. (14) predicted by the network model, in which the initial network size is subsumed within the community dummies interacted with linear and quadratic time trends. Given the small size of the data set and possibility of serially correlated errors, standard errors are clustered using the wild cluster bootstrap of Cameron et al. (2008). In Table 7, the Europeans constitute the reference community group; the online appendix shows the results are robust to changing the reference group to any other community.

In the absence of a network effect, the community specific quadratic time trend dummies ϕ_c should all equal zero. Since the estimated coefficients are not of intrinsic interest, we report only the F-statistic for this test. Column 1 shows the null hypothesis of a zero network effect is rejected with a p-value less than .01.

To confirm the pattern of divergence predicted by the network model, and obtain an estimate of the magnitude of the network effect, Column 2 replaces the community dummies by their levels of initial presence n_0^c in 1866, and also includes upstream presence n_0^{uc} in that year:

$$n_t^c = \chi n_0^c * t^2 + \delta n_0^c * t + \gamma n_0^c + \eta n_0^{uc} + \nu_t + \epsilon_t^c \quad (18)$$

This specification requires the stronger assumption that levels of initial presence n_0^c were uncorrelated with unobserved community characteristics. Since this assumption is strong, the estimates reported are purely illustrative. The estimate of χ is 0.008, significant at the 1% level, while that of δ is negative and statistically insignificant. The estimated coefficient η of initial upstream presence is negative and significant, reinforcing the evidence against hypotheses stressing the role of prior accumulation of skills or wealth in the upstream sector.

Columns 3,4 and 5 show that the estimated magnitude and significance of χ in Column 2 is robust under alternative specifications of (18). Column 3 replaces the initial downstream and upstream presence by community dummies; Column 4 additionally replaces the interaction of linear time trend with n_0^c by δ_c. Finally, Column 5 replaces year dummies in Column 4 by lagged ratio of manufactured goods price and raw cotton price. The two period lagged price ratio is significant at the 5% level.

The common estimate of χ across Columns 2–4 corresponds to a value of the network effect ζ (in Eq. (14)) of approximately 0.13. In other words, one extra entrant from a community in any given year was associated with 0.13 more entrants from the same community in the following year. Using the quadratic approximation for M^t, this cumulates to 2.85 additional entrants fifteen years later, and 5.75 more entrants twenty five years later.[10]

Table 7 used data for the actual stock of incumbents from each community at different dates, which reflects the joint effect of entry and exits. In addition we sometimes observe entrepreneurs listed in some years, drop out for intervening years and then get listed again subsequently. This could reflect errors of omission in the listing of data for some years, or intermittent entry and exit by certain entrepreneurs. However our theory focused entirely on entry decisions, assuming that entrepreneurs that once enter do not ever exit. We now check the extent to which the preceding results continue to hold if we focus in the data on the entry process exclusively. To do this we replace the actual community-year incumbent stock data series by cumulating past flows of entrants from each community. This would have been the stock of incumbents if entrants never exited. Table 8 shows the results of the same regression as Table 7, using this alternative series as the dependent variable. It is evident that the results are almost entirely the same, and thereby unaffected by the pattern of exits or measurement error resulting from omission of some entrepreneurs in the business registers for some years.

The Online Appendix contains a number of robustness checks of Table 7: (a) when each of the communities is dropped from the dataset one at a time, and (b) when the community forming the reference group is altered from the Europeans to any one of the remaining communities. In (a), the coefficient of $n_0^c * t^2$ becomes substantially less precise and is no longer statistically significant when any of the communities (apart from Vanias) is dropped. This is not surprising given the small number of communities. It also indicates that the results are not driven by any single community. Part (b) shows that the results are robust to the choice of the reference group.

Entering capital size regressions: Testing adverse selection

Next, we examine evidence for misallocation as predicted by the network model: that entering capital size declines faster among marginal entrants from communities with higher initial presence. To avoid effect of outliers, we use as dependent variable K_t^{mc}, the 10th percentile of the distribution of capital per entrepreneur for firms from

[10] Recall the quadratic approximation is $M^t = 1 + \zeta * t + \frac{\zeta^2}{2} * t^2$.

Table 7
Regression: Stock of active downstream Bombay entrepreneurs at community level (1866–1890).

VARIABLES	(1)	(2)	(3)	(4)	(5)
n_0^c downstream $*$ t^2		0.008***	0.008***	0.008***	0.007***
		(0.000)	(0.000)	(0.000)	(0.000)
n_0^c downstream $*$ t		−0.082	−0.082		
		(0.220)	(0.220)		
n_0^c downstream		1.673*			
		(0.090)			
n_0^c upstream		−0.322			
		(0.200)			
Log cotton price ratio,					5.947
lagged one year					(0.100)
Log cotton price ratio,					6.609**
lagged two years					(0.020)
Year dummies	Y	Y	Y	Y	N
Community dummies	Y	N	Y	Y	Y
Community $*$ t	Y	N	N	Y	Y
Community $*$ t^2	Y	N	N	N	N
F statistic: Community $*$ t^2	14.09				
Prob >F	0.00				
Observations	125	125	125	125	125
R-squared	0.992	0.936	0.978	0.990	0.973

Unit of observation: community-year level. The dependent variable is the total number of active entrepreneurs for any given community in a given year in the downstream Bombay industry. Coefficients not reported in any column correspond to excluded regressors in that regression, owing to collinearity with controls.
Column 1, F-statistic refers to F-test for equality of coefficients of all community dummies interacted with t^2.
Owing to the small number of clusters, we correct standard errors using the wild bootstrap procedure of Cameron et al. (2008) and report bootstrapped p- values in parentheses. *** : $p < 0.01$, ** : $p < 0.05$, * : $p < 0.1$.

Table 8
Regression: Cumulative past entry of downstream Bombay entrepreneurs at community level (1866–1890)

VARIABLES	(1)	(2)	(3)	(4)	(5)
n_0^c downstream $*$ t^2		0.009***	0.009***	0.009***	0.014***
		(0.000)	(0.000)	(0.000)	(0.000)
n_0^c downstream $*$ t			−0.023	−0.023	
			(0.580)	(0.580)	
n_0^c downstream				1.926***	
				(0.000)	
n_0^c upstream				−0.145	
				(0.380)	
Log cotton price ratio,					−1.581
lagged one year					(0.760)
Log cotton price ratio,					9.833*
lagged two years					(0.080)
Year dummies	Y	Y	Y	Y	N
Community dummies	Y	Y	Y	N	Y
Community $*$ t	Y	Y	N	N	Y
Community $*$ t^2	Y	N	N	N	N
F statistic: Community $*$ t^2	14.09				
Prob >F	0.00				
Observations	125	125	125	125	125
R-squared	0.992	0.990	0.978	0.936	0.973

Unit of observation: community-year level. The dependent variable is the cumulative number of entrepreneurs for any given community in a given year who have entered at or in a prior year since 1865 in the downstream Bombay industry. Coefficients not reported in any column correspond to excluded regressors in that regression, owing to collinearity with controls.
Column 1, F-statistic refers to F-test for equality of coefficients of all community dummies interacted with t^2.
Owing to the small number of clusters, we correct standard errors using the wild bootstrap procedure of Cameron et al. (2008) and report bootstrapped p- values in parentheses. *** : $p < 0.01$, ** : $p < 0.05$, * : $p < 0.1$.

Table 9
Regression: Log of capital per entrepreneur (10th percentile): 1866–1890.

VARIABLES	(1)	(2)	(3)	(4)
n_0^c downstream $* t^2$		−0.004***	−0.002	
		(0.000)	(0.150)	
n_0^c downstream $* t$			0.062	−0.001
			(0.210)	(0.720)
Year dummies	Y	Y	Y	Y
Community dummies	Y	Y	Y	Y
Community $* t$	Y	Y	N	N
Community $* t^2$	Y	N	N	N
F test Community $* t^2$	8.66			
Prob >F	0.02			
Observations	31	31	31	31
R-squared	0.964	0.917	0.784	0.720
Intercept not reported.				

Unit of observation: community-year level. The dependent variable is the 10th percentile of the distribution of log capital per entrepreneur (firm capital divided by number of entrepreneurs) for new entrants from any community in that year. Coefficients not reported in any column correspond to excluded regressors in that regression, owing to collinearity with controls.
Column 1, F-statistic refers to F-test for equality of coefficients of all community dummies interacted with t^2.
Owing to the small number of clusters, we correct standard errors using the wild bootstrap procedure of Cameron et al. (2008) and report bootstrapped p- values in parentheses. $^{***}: p < 0.01$, $^{**}: p < 0.05$, $^{*}: p < 0.1$.

Table 10
Regression: Log of capital per entrepreneur (Average): 1866–1890.

VARIABLES	(1)	(2)	(3)	(4)
n_0^c downstream $* t^2$		−0.003***	−0.002*	
		(0.000)	(0.090)	
n_0^c downstream $* t$			0.063**	−0.001
			(0.020)	(0.720)
Year dummies	Y	Y	Y	Y
Community dummies	Y	Y	Y	Y
Community $* t$	Y	Y	N	N
Community $* t^2$	Y	N	N	N
F test Community $* t^2$	5.02			
Prob > F	0.05			
Observations	31	31	31	31
R-squared	0.958	0.929	0.737	0.658
Intercept not reported.				

Unit of observation: community-year level. The dependent variable is the average of log capital per entrepreneur (firm capital divided by number of entrepreneurs) for new entrants from any community in that year. Coefficients not reported in any column correspond to excluded regressors in that regression, owing to collinearity with controls.
Column 1, F-statistic refers to F-test for equality of coefficients of all community dummies interacted with t^2.
Owing to the small number of clusters, we correct standard errors using the wild bootstrap procedure of Cameron et al. (2008) and report bootstrapped p- values in parentheses. $^{***}: p < 0.01$, $^{**}: p < 0.05$, $^{*}: p < 0.1$.

community c (i.e., firms for which c constitutes the largest proportion of entrepreneurs) that enter in year t. Combining equations ((15), (16)) with (18), we obtain

$$\log K_t^{mc} = -\frac{\theta\sigma}{1-\sigma}[\chi n_0^c * (t-1)^2 + \delta n_0^c * t + \gamma n_0^c + \eta n_0^{uc} + v_t] - \frac{1}{1-\sigma}\log q_t + \frac{1-\alpha}{1-\sigma}\log \Gamma_c + \log A_0 \tag{19}$$

Hence the regression specification is (if initial network size n_0^c is exogenous):

$$\log K_t^{mc} = \mu_1 n_0^c * t^2 + \mu_2 n_0^c * t + \mu_3 \delta_c + v_t' + \epsilon_{ct}' \tag{20}$$

where $\mu_1 \equiv -\frac{\chi\theta\sigma}{1-\sigma} < 0, \mu_2 \equiv \frac{(2\chi-\delta)\theta\sigma}{1-\sigma}$. If we drop the exogeneity assumption of initial network size, we can replace n_0^c interactions with time trends with corresponding community dummy interactions:

$$\log K_t^{mc} = \phi_{1c} * t^2 + \phi_{2c} * t + \phi_{3c} + v_t' + \epsilon_{ct}' \tag{21}$$

where network effects imply the existence of community-specific quadratic and linear time trends.

Table 9 shows these regression results, restricting the sample to community-year pairs where there was positive entry of firms from the community in question. As in the entry regressions, Column 1 corresponds to specification (21), and shows that the community specific quadratic time trends are jointly significant. Column 3 corresponds to specification (20), while Column 2 shows an intermediate specification where the quadratic time trend is interacted with initial network size but the linear time trend is interacted with community dummies. In the latter we see a statistically significant negative interaction of the quadratic trend with initial network size. This is no longer the case in Column 3 (though in the latter case the signs of the estimated coefficients are consistent with the network theory).

Table 10 shows corresponding regressions for the log of capital per entrepreneur K_t^{ac} for the average (rather than marginal) entrant from community c in year t. The results in Column 2 are very similar to those of marginal entrants in the previous table, while in Column 3 they are now statistically significant. Column 4 which includes only community and year fixed effects shows absence of a positive interaction between initial network size and the linear time trend. Hence there is no evidence in favor of the alternative hypothesis based on positive selection.[11]

5.2. Evidence from the longer time span 1860–1910

Fig. 9 extends the time plot of presence of different communities until 1910 in the two regions surrounding Bombay and Ahmedabad respectively. It shows the emergence of the Ahmedabad region as a new location for the industry. There is a steep surge of entry between 1895–1900 that is accounted for mainly by the Vania community, who became the largest group in the industry. The process of entry of the Vanias resembles a cascade, consistent with the hypothesis of strong community networks.

Vanias in the Ahmedabad region were unlikely to be significantly connected with the Vania community located in Bombay. Other assumptions of the network model would also cease to be valid for the Bombay-based industry; after the passage of almost three decades it may have entered the mature stage, with no scope for entry rates among the Bombay-based communities to rise any further. Fig. 9 shows that the presence of different communities in the Bombay region tended to grow more slowly after 1900 and exhibited parallel trends, while the newly emerging Ahmedabad based industry experienced the kind of nonlinear divergence between Vanias and other communities that the Bombay industry had experienced earlier during the 1865–90 period. It is therefore not surprising that the significance of the network effects that we found during the phase 1865–90 for the Bombay based industry does not extend when we run the corresponding regressions

[11] When we replace the dependent variable in Tables 9, 10 by log of entering firm size (rather than capital per entrepreneur), the coefficients of $n_0^c * t^2$ continues to be negative, though statistically indistinguishable from zero. Hence there is no evidence that entering capital size of firms exhibited a larger upward trend among communities with higher initial presence. This addresses the possible concern that the preceding results were driven by increasing entry flows of entrepreneurs per firm, which dominated higher time trends in entering capital size per firm among communities with higher initial presence.

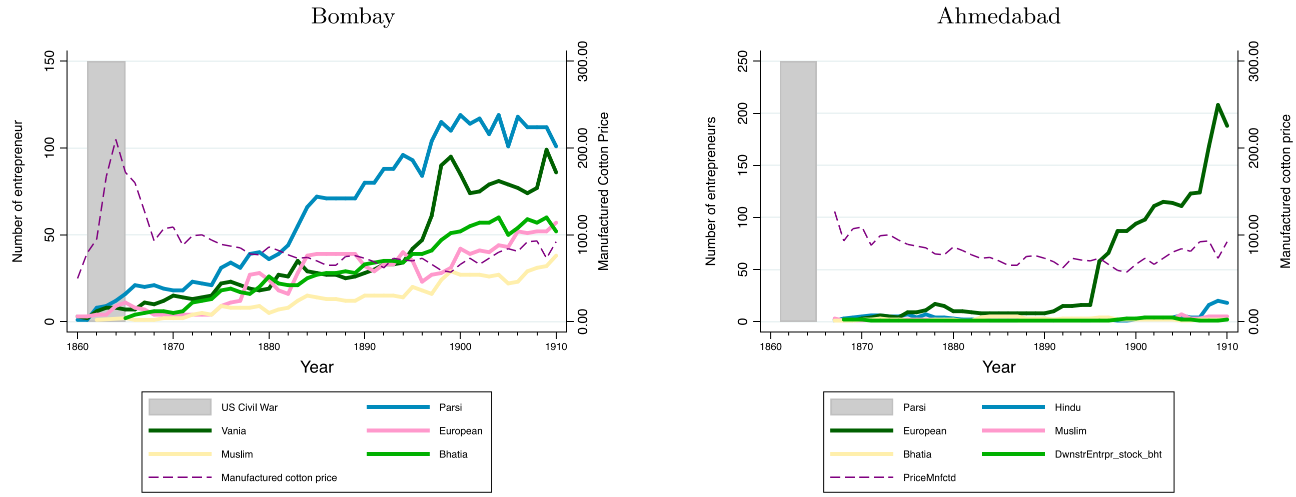


Fig. 9. Entrepreneur stocks in the downstream cotton textile industry by location (1860–1910).
Source: Entrepreneur stocks: own data; prices: see Appendix A.

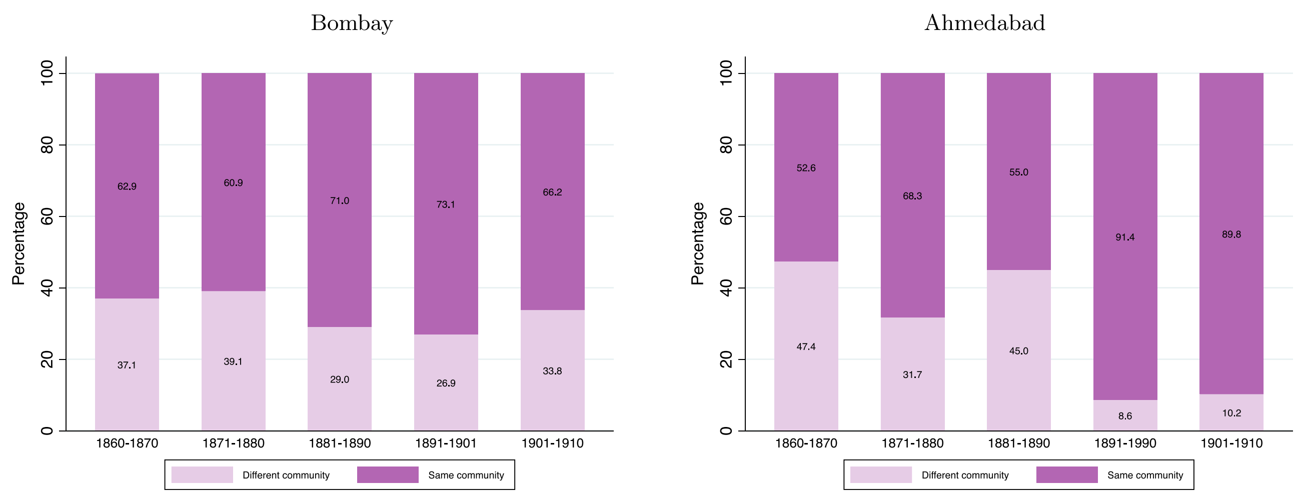


Fig. 10. Community concentration of entrants in downstream cotton textile firms at entry by location (1860–1910).
Source: Own data. For definitions of Same and Different community shares, see notes to Fig. 7.

for network size and entering capital per entrepreneur for the longer period 1865–1910.[12]

However, the growth of the industry beyond 1890 still featured aspects of the importance of community networks. We find evidence of clustering by community within firms until 1910. In Fig. 10, we show this by location. Entry into the Ahmedabad region after 1890 exhibits even larger clustering than the Bombay region, with nearly 90% of entrepreneurs from the same community. Moreover, prior upstream experience was even less important for the late entrants than for the previous entrants in both locations (Fig. 11).

6. Entry patterns in the jute industry

Unlike Bombay, the commercial life of Calcutta was dominated by British firms. While Bombay was the hub of Indian capital, Calcutta was the centre of British firms in tea, coal and jute. Many investors were British, resident in India. The Calcutta industry we focus on is jute, for a number of reasons. Until 1930 this was the most important instance of entry by Indian entrepreneurs after cotton. Jute is also closest to cotton textiles in terms of technology and capital requirements. It developed with British entrepreneurship around the same time as cotton textiles in Bombay. But unlike cotton, jute was mainly exported and British firms had an advantage in the export trade. In 1866, there were four British firms in jute. From the mid 1870s, the industry expanded rapidly and by 1900, there were 32 British firms, but no Indian firms. The situation changed with the first World War, when Indian entrepreneurship appeared for the first time.

Indian presence in Calcutta's commercial sector had been low compared to Bombay. The Marwari traders, the dominant group among Indian communities in Eastern India, were involved in money lending, trade, brokerage with European companies and speculation in different markets (Tripathi, 2004, p.166). The futures market in opium, gold specie and later raw jute and hessian was started by the Marwaris and became the focus of speculation. Several Marwaris first created their wealth in the opium trade. Among them, Birla and Hukumchand were the key players, who later became the pioneers in establishing new jute firms (Timberg, 1978, p. 160–61). While the export trade in raw jute and jute textiles had been controlled by the Europeans before the war, after 1914 Marwari traders Birla and Hukumchand became involved in this trade (Bagchi, 1994, p.179).

British Managing Agency Houses (which managed and controlled firms across various industries, including jute) had Marwaris as brokers.

[12] We do not show the regressions for the longer time span, in order to conserve space.

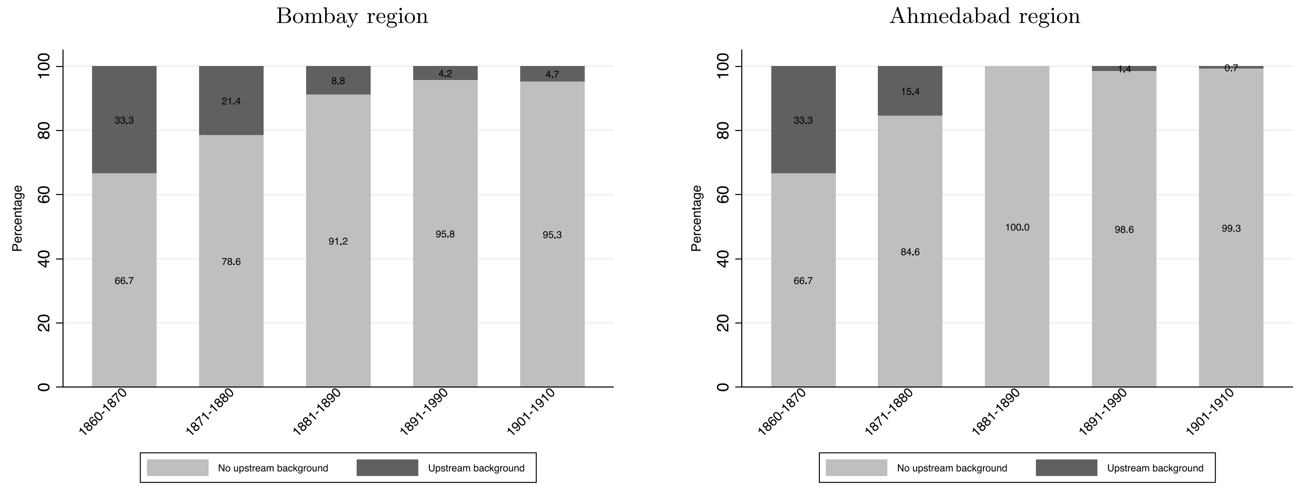


Fig. 11. Relevance of upstream background in downstream entry into cotton textiles (1860–1910).
Source: Own data. See notes to Fig. 8 for definition of upstream experience.

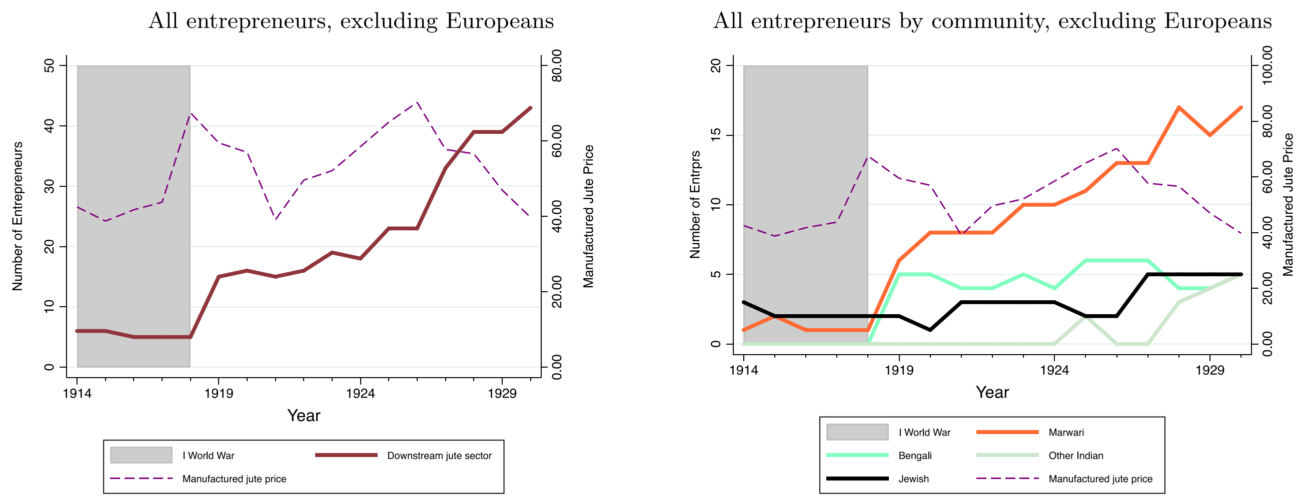


Fig. 12. Entrepreneur stocks in the Jute downstream sector,1914–1930.
Source: Entrepreneur stocks: own data; prices: see Appendix A.

The first World War was a demand shock for the jute industry, just as the American civil war had been in the case of cotton half a century earlier. As demand for jute goods increased, raw jute prices rose; share prices of jute firms rose 8–10 times (Goswami, 1985). The close contact of the Marwari families with British firms opened up a channel of entry into the industry. British shareholders sold shares to their Marwari contacts. The Marwaris also lent short term capital to British firms against a collateral of shares. Both interactions allowed the Marwaris to acquire shares in British firms and get elected on the boards of British Companies. In 1918, of the 114 directors in British owned jute firms, only 3 were Marwaris. By 1924, of the 46 British firms, 19 had Marwari directors on their boards (Goswami, 1985). The first Indian-owned firms were set up in 1918 by Birla and Hukumchand. While Hukumchand's firm was primarily self-financed, capital for Birla's firm was raised more widely (Timberg, 1978, p.171). As in cotton, the initial stock of entrepreneurs opened up the way for further entry, which continued until the beginning of the Great Depression. Unlike cotton textiles, the entry of Indian traders into the jute industry happened both through acquisition of shares in British firms and starting new firms.

Fig. 12 shows entry into the downstream jute industry from 1914 to 1930. Until 1918, only a handful of Indian entrepreneurs from three communities were present: Marwaris, Baghdadi Jews and Bengalis. After 1918, there was a steep rise in Marwari entry.

Fig. 13 shows that Indian entrants did not have upstream experience. Analogous to the cotton industry, we define jute baling as the upstream sector. Fig. 14 shows the extent of clustering by communities within firms. More than half the partners within the average firm belonged to the same community. The clustering was the highest (58%) among the Marwaris, although the extent of clustering was lower than that observed in the early phases of the cotton textile industry. This is perhaps explained by the different path of entry compared to cotton textiles. Many Marwari entrepreneurs entered by acquiring shares in existing firms from British investors initially.

Finally, we test the model of network-based dynamics by using a regression specification analogous to that used for the cotton industry in Table 11. The results are very similar to those for the early stage of Indian entrepreneurship in downstream cotton in Bombay. The short time span for the jute data, however, do not allow us to estimate the corresponding regressions for capital size of entrants.

7. Concluding comments

This paper has presented evidence from colonial India on the role played by community networks in early stages of development of downstream cotton and jute industries, when market institutions were weak and state support was missing. We have put together data on individual

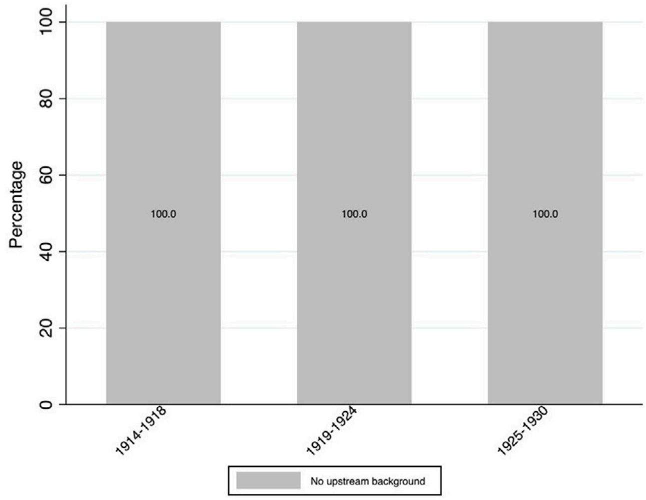


Fig. 13. Upstream experience of downstream entrants into Jute sector,1914–1930.
Source: Own data. The figure shows the percentage of entrepreneurs in the downstream jute industry between 1914–30 who respectively were and were not listed in any previous year in an upstream jute firm.

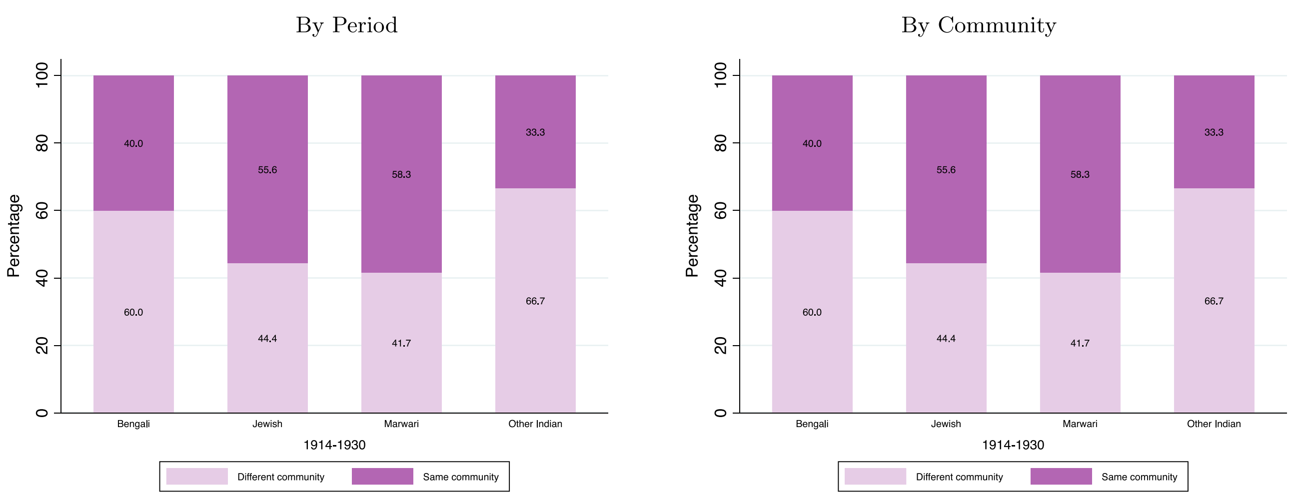


Fig. 14. Community concentration of entrants into downstream jute firms,1914–1930 .
Source: Own data. 'Same community' share equals the largest share of entrepreneurs from any community within a firm, averaged across downstream jute firms listed during each of the time periods (in the left panel) and averaged across communities for the entire period (right panel). 'Different community' share equals 100% minus Same community share, i.e., refers to average intra-firm share of entrepreneurs from a community different from the community with the largest share.

entrepreneurs in the cotton and jute industry, thus filling a gap in the literature (e.g., Bagchi (1994), p.185) states: "No detailed figures on the distributions of directorships among various Indian communities are available"). We use this data to show that community connections were important in early entry. Entrants tended to join firms dominated by entrepreneurs from the same community. Presence of communities in the early stages predict their presence several decades later, consistent with the hypothesis that help provided by early incumbents facilitated entry of entrepreneurs from the same community, controlling for fixed differences in unobserved community characteristics, varying prices and other time varying shocks in the industry.

We also found evidence of misallocation of talent resulting from the reliance on community networks, thereby ruling out alternative explanations of observed entry dynamics based on faster growth in education or wealth of communities with higher initial presence. Entrants in communities with higher initial and subsequent presence tended to enter with lower capital sizes, indicating they were of lower ability compared to potential entrants from other communities. Owing to this it is difficult to make any inferences concerning the aggregate welfare impact of the network effects. On the one hand, the help provided by incumbents from the same community encouraged entry, thereby alleviating the underinvestment associated with lack of market institutions or infant industry protection from the state. This is likely to have constituted a first order positive welfare effect, which would have to be set against the welfare losses resulting from the misallocation due to lack of uniformity in networks across communities.

The results in this paper raise further interesting questions, pertaining to the role of community networks in later stages of evolution of these industries. After 1890, the cotton industry witnessed emergence of new locations and new communities. Our empirical methodology could not be applied to draw any clear inferences regarding the role of networks at these later stages. A related question concerns the causes of delayed entry by specific communities in specific regions. In the case of cotton, the later development of the Ahmedabad region, may be explained by the disappearance of the primary pre-industrial occupation (moneylending) of the relevant community (Hindu and Jain Vanias) following the Deccan riots in the 1870s. In similar vein, delayed entry into industrialization in South India by the community of

Table 11
Regression: Stock of active entrepreneurs at community level in downstream jute, 1919–1930.

VARIABLES	(1)	(2)	(3)	(4)
n_0^c downstream $* t^2$		0.005***	0.005***	0.005***
		(0.000)	(0.000)	(0.000)
n_0^c downstream $* t$		−0.077***	−0.077***	
		(0.000)	(0.000)	
n_0^c downstream		1.129***		
		(0.000)		
n_0^c upstream		0.614		
		(0.470)		
Year dummies	Y	Y	Y	Y
Community dummies	Y	N	Y	Y
Community $* t$	Y	N	N	Y
Community $* t^2$	Y	N	N	N
F test statistic: Community $* t^2$	2.80			
Prob > F	0.04			
Observations	60	60	60	60
R-squared	0.995	0.988	0.993	0.995

Unit of observation: community-year level. The dependent variable is the total number of active entrepreneurs for any given community in a given year in the downstream jute industry. Coefficients not reported in any column correspond to excluded regressors in that regression, owing to collinearity with controls.
Column 1, F-statistic refers to F-test for equality of coefficients of all community dummies interacted with t^2.
Owing to the small number of clusters, we correct standard errors using the wild bootstrap procedure of Cameron et al. (2008) and report bootstrapped p- values in parentheses. $^{***}: p < 0.01$, $^{**}: p < 0.05$, $^{*}: p < 0.1$.

Chettiars (after 1930s) is possibly explained by the decline of their moneylending activities in Burma in the aftermath of the Great Depression. Conversely, the early lead of the Bombay region may have owed to the earlier disappearance of profitable trading opportunities in opium and cotton enjoyed by Western Indian business communities. The lack of pre-industrial accumulation of wealth or skill in related trading sectors seems unlikely to explain delayed industrialization in other regions. On the contrary, these delays probably arose precisely owing to the continued profitability of these pre-industrial trading opportunities, which reduced the incentive of the concerned communities to enter industry. In other words, successful community networks can hinder progression into the next higher stage of development, an effect that did not manifest itself in the phenomena studied in this paper. Exploring this hypothesis remains an interesting task for future research.

Data availability

Data will be made available on request.

Appendix A

A.1. Data sources

We have several data sources which we collected and digitized. In this paper, we construct panels for cotton and jute industry in India, between 1850 and 1930 at the entrepreneur level.

Our data provides information on preindustrial activity from before the first waves of industralization, documenting the presence of cotton merchants and jute balers, and up to 1930 with the impact of the Great Depression. In cotton sector, industralization started right after the American civil war, while in jute sector Indian entry took place after 1914.

A.1.1. Cotton industry registers

"*The Bombay Almanac, Directory, and Register*", (1806–1868)
We construct the List of Merchants, from the records of firms and partners in firms organized by occupation for years: 1840, 1850, 1855, 1860, 1865. In this dataset is also available the list of Joint Stock Companies from 1860 onward.
" *Bombay Calendar and Almanac*", (1853–1861)
From this dataset we only use years 1860 and 1861, where we obtain: (i) the list of firms, location, and capital; (ii) the list of entrepreneurs and individuals by occupation.
"*Times of India Calendar and Directory (Bombay)*", (1862–1887)
We use this dataset for the period (1862–1884), where we collect: (i) the list of firms, location, and capital; (ii) the list of entrepreneurs.
"*Thacker's Indian Directory (Bombay)*", (1885–1960)
This is our main data source, which we use from year 1885 onward. From here we construct: (i) the list of firms, location, and capital; (ii) the list of entrepreneurs.

A.1.2. Jute industry registers

"*New Calcutta Directory*" (1856–1863)
We construct yearly cross sections between 1860 and 1869: (i) the list of firms, location, and capital; (ii) the list of entrepreneurs.
"*Thackers Bengal Directory*" (1863–1884)
From this dataset we construct yearly cohorts from 1870 to 1875, in addition to 1880. The information found in the directory is: (i) the list of firms, location, and capital; (ii) the list of entrepreneurs.
"*Thacker's Indian Directory (Calcutta)*", (1885–1960)
This dataset is used to obtain: (i) the list of jute balers; (ii) the list of firms, location, and capital; (iii) the list of entrepreneurs and individuals by occupation.
From 1885 to 1905 we collected data every five years, and form 1905 we construct our yearly cohorts up to 1930.

A.1.3. Cotton and jute prices

"*Index Number of Indian Prices*" India Office Records (1861–1931)
The most relevant pieces of information extracted from this source are the following:

Exported articles: (i) Cotton, Raw Broach (Bombay) per candy of 784 lb; (ii) Cotton, Manufactured Yarn 20 s (Bombay) per lb and T. cloth (Bombay) per lb (only between 1874–1931); (iii) Jute, Raw Picked & Ordinary (Calcutta). Per bale of 400 lb.; (iv) Jute, Manufactured: Gunny bags (Calcutta). Per bale of 100 lb
Imported articles: (i) Cotton, Manufactured Grey shirtings (Calcutta) and Grey yarn Banner Mill (Calcutta)

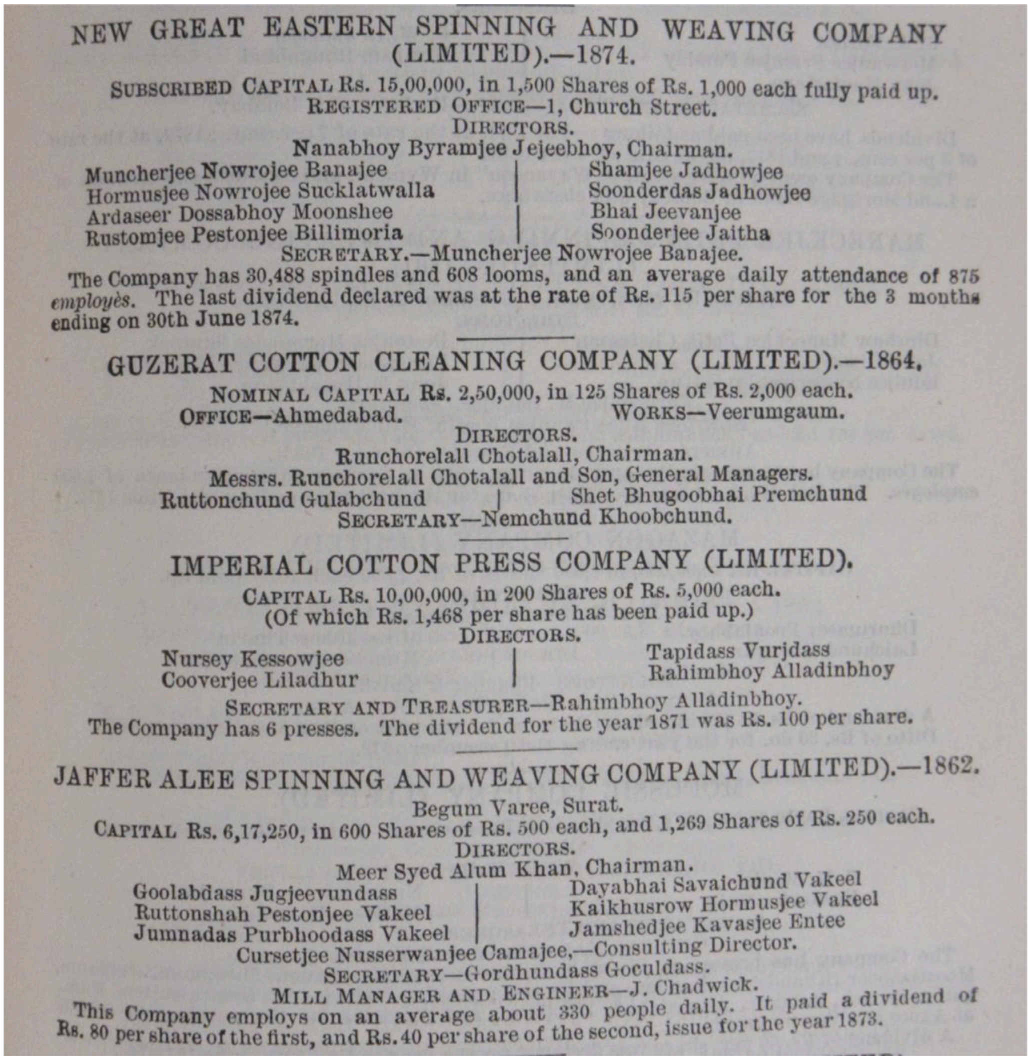
A.1.4. Data sample

Fig. 15 provides a sample of our data pertaining to a cotton spinning and weaving company in 1874.

A.1.5. Entrepreneur and firm panels construction

We first pool yearly cross-section cohorts with all firms in upstream and downstream sector. From each firm we obtain the names of partners, directors and managing agents. Each of them were associated to a community origin, which defined the firm's community by simple majority. Raw and manufactured cotton (and jute) prices from the *Index Number of Indian Prices* were yearly assigned to entrepreneurs and firms panels.

The Entrepreneur panel collapses yearly information at individual level, as a single entrepreneur could be part of more than one firm in the same or different sector during the same year. Therefore, in our panel we are able to identify whether an entrepreneur is incumbent in one sector and later entrant in the other. Information on capital requirement is summarized at individual level as well.



NEW GREAT EASTERN SPINNING AND WEAVING COMPANY (LIMITED).—1874.

SUBSCRIBED CAPITAL Rs. 15,00,000, in 1,500 Shares of Rs. 1,000 each fully paid up.
REGISTERED OFFICE—1, Church Street.
DIRECTORS.
Nanabhoy Byramjee Jejeebhoy, Chairman.

Muncherjee Nowrojee Banajee	Shamjee Jadhowjee
Hormusjee Nowrojee Sucklatwalla	Soonderdas Jadhowjee
Ardaseer Dossabhoy Moonshee	Bhai Jeevanjee
Rustomjee Pestonjee Billimoria	Soonderjee Jaitha

SECRETARY.—Muncherjee Nowrojee Banajee.
The Company has 30,488 spindles and 608 looms, and an average daily attendance of 875 *employès*. The last dividend declared was at the rate of Rs. 115 per share for the 3 months ending on 30th June 1874.

GUZERAT COTTON CLEANING COMPANY (LIMITED).—1864.

NOMINAL CAPITAL Rs. 2,50,000, in 125 Shares of Rs. 2,000 each.
OFFICE—Ahmedabad. WORKS—Veerumgaum.
DIRECTORS.
Runchorelall Chotalall, Chairman.
Messrs. Runchorelall Chotalall and Son, General Managers.
Ruttonchund Gulabchund | Shet Bhugoobhai Premchund
SECRETARY—Nemchund Khoobchund.

IMPERIAL COTTON PRESS COMPANY (LIMITED).

CAPITAL Rs. 10,00,000, in 200 Shares of Rs. 5,000 each.
(Of which Rs. 1,468 per share has been paid up.)
DIRECTORS.

Nursey Kessowjee	Tapidass Vurjdass
Cooverjee Liladhur	Rahimbhoy Alladinbhoy

SECRETARY AND TREASURER—Rahimbhoy Alladinbhoy.
The Company has 6 presses. The dividend for the year 1871 was Rs. 100 per share.

JAFFER ALEE SPINNING AND WEAVING COMPANY (LIMITED).—1862.

Begum Varee, Surat.
CAPITAL Rs. 6,17,250, in 600 Shares of Rs. 500 each, and 1,269 Shares of Rs. 250 each.
DIRECTORS.
Meer Syed Alum Khan, Chairman.

Goolabdass Jugjeevundass	Dayabhai Savaichund Vakeel
Ruttonshah Pestonjee Vakeel	Kaikhusrow Hormusjee Vakeel
Jumnadas Purbhoodass Vakeel	Jamshedjee Kavasjee Entee

Cursetjee Nusserwanjee Camajee,—Consulting Director.
SECRETARY—Gordhundass Goculdass.
MILL MANAGER AND ENGINEER—J. Chadwick.
This Company employs on an average about 330 people daily. It paid a dividend of Rs. 80 per share of the first, and Rs. 40 per share of the second, issue for the year 1873.

Fig. 15. Data Sample.
Source: Own data.

The Firm panel gathers all relevant information at entrepreneurs level. We assign a community to each entrepreneur using his name. We calculate the proportion of entrepreneurs belonging to the same community in a firm and the firm's community identity is based on the community of the majority of entrepreneurs. For each firm we assign a code or ID, which enables to track the same firm despite of firm's names changes. The location of firms is obtained from registered office address, and mills or presses location.

Cross section cohorts for cotton industry panels: (1860–61) *Bombay Calendar and Almanac*; (1862–1884) *Times of India Calendar and Directory (Bombay)*; (1885–1910) *Thacker's Indian Directory (Bombay)*;
Cross section cohorts for cotton jute panels: (1860–1869) *New Calcutta Directory*; (1870–1875) *Thackers Bengal Directory*; (1885–1905) for every five years *Thacker's Indian Directory (Calcutta)*; (1905–1930) *Thacker's Indian Directory (Calcutta)*

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jdeveco.2022.102973.

References

Anstey, V., 1949. The Economic Development of India. Longmans Green.
Bagchi, A.K., 1987. The Evolution of the State Bank of India: The Era of the Presidency Banks, 1876-1920 (Vol. 2). Oxford University Press.
Bagchi, A.K., 1994. European and Indian entrepreneurship in India 1900-30. In: Ray, R.K. (Ed.), Entrepreneurship and Industry in India 1800-1947. Oxford University Press, Delhi.
Banerjee, A., Munshi, K., 2004. How efficiently is capital allocated? Evidence from the knitted garment industry in Tirupur. Rev. Econom. Stud. 71 (1), 19–42.
Banerjee, A., Newman, A., 1993. Occupational choice and the process of development. J. Polit. Econ. 101 (2), 274–298, (Apr., 1993).
Beaman, L., Magruder, J., 2012. Who gets the job referral? Evidence from a social networks experiment. Amer. Econ. Rev. 102 (7), 3574–3593.
Beteille, A., 1996. Varna and jati. Sociol. Bull. 45 (1), 15–27.
Buchanan, D.H., 1934. The Development of Capitalist Enterprise in India. New York:Macmillan.
Burhop, C., 2006. Did banks cause the german industrialization? Explor. Econ. Hist. 43 (1), 39–63.
Cameron, A., Gelbach, J., Miller, D., 2008. Bootstrap-based improvements for inference with clustered errors. Rev. Econ. Stat. 90 (3), 414–427.
Dai, R., Mookherjee, D., Munshi, K., Zhang, X., 2018. Community Networks and the Growth of Private Enterprise in China. Working Paper, Institute for Economic Development, Boston University.
Das Gupta, U. (Ed.), 2001. The World of the Indian Ocean Merchant, 1500-1800: Collected Essays of Ashin Das Gupta. Oxford University Press, USA.
Desai, A.V., 1968. The origins of parsi enterprise. Indian Econ. Soc. Hist. Rev. 5 (4), 307–317.

Fafchamps, M., 2003. Market Institutions in Sub-Saharan Africa: Theory and Evidence. MIT Press.
Farooqui, A., 2005. Smuggling As Subversion: Colonialism, Indian Merchants, and the Politics of Opium, 1790-1843. Lexington Books.
Ghurye, G.S., 1961. Caste, Class and Occupation. Popular Book Depot, Bombay.
Goswami, O., 1985. Then came the marwaris: Some aspects of the changes in the pattern of industrial control in eastern India. Indian Econ. Soc. Hist. Rev. 22 (3), 225–249.
Greif, A., 2006. Institutions and the Path to the Modern Economy: Lessons from Medieval Trade. Cambridge University Press.
Guha, A., 1984. More about the Parsi Seths: Their roots, entrepreneurship and comprador role, 1650-1918. Econ. Political Wkly. 117–132.
Harnetty, P., 1972. Imperialism and Free Trade: Lancashire and India in the Mid-Nineteenth Century. Manchester University Press.
Irwin, D.A., 2007. Tariff incidence in America's gilded age. J. Econ. Hist. 67 (3), 582–607.
Kahan, A., 1967. Government policies and the industrialization of Russia. J. Econ. Hist. 27 (4), 460–477.
Kranton, R.E., Swamy, A.V., 1999. The hazards of piecemeal reform: British civil courts and the credit market in colonial India. J. Dev. Econ. 58 (1), 1–24.
Markevich, A., Nafziger, S., 2017. The state and market in Russian industrialization, 1870-2010. In: O'Rourke, K. H, Williamson, J.G. (Eds.), Spread of Modern Industry in the Poor Periphery Since 1871. Oxford University Press, pp. 33–62.
Marx, K., 1887. Capital (Vol. 1). Progress Publishers, Moscow..
Mehta, M., 1991. Indian Merchants and Entrepreneurs in Historical Perspective:With Spcial Reference to Shroffs of Gujarat, 17th to 19th Centuries. Academic Foundation, Delhi..
Mokyr, J., 2005. The intellectual origins of modern economic growth. J. Econ. Hist. 65 (2), 285–351.
Mokyr, J., 2009. Intellectual property rights, the industrial revolution, and the beginnings of modern economic growth. Am. Econ. Rev. 99 (2), 349–355.
Morris, M.D., 1965. The Emergence of an Industrial Labor Force in India: A Study of the Bombay Cotton Mills, 1854-1947. Univ of California Press.
Munshi, K., 2011. Strength in numbers: Networks as a solution to occupational traps. Rev. Econom. Stud. 78 (3), 1069–1101.
Munshi, K., Rosenzweig, M., 2016. Networks and misallocation: Insurance, migration and the rural-urban wage gap. Amer. Econ. Rev. 106 (1), 46–98.
North, D., Weingast, B.R., 1989. Constitutions and commitment: the evolution of institutions governing public choice in seventeenth century england. Journal of Economic History 49 (4), 803–832.
Panini, M.N., 1996. The political economy of caste. In: Srinivas, M.N. (Ed.), Caste: Its Twentieth Century Avatar. Penguin Books India.
Perkins, D., Tang, J.P., 2017. The east Asian pioneers: Japan, Korea and Taiwan. In: O'rourke, K.H, Williamson, J.G. (Eds.), Spread of Modern Industry in the Poor Periphery Since 1871. Oxford University Press, pp. 169–196.
Rauch, J.E., Trindade, V., 2002. Ethnic Chinese networks in international trade. Rev. Econ. Stat. 84 (1), 116–130.
Roy, T., Swamy, A.V., 2016. Law and the Economy in Colonial India. University of Chicago Press.
Rudner, D.W., 1994. Caste and Capitalism in Colonial India: The Nattukottai Chettiars. University of California Press, Berkeley, p. 93.
Rungta, S., 1970. The Rise of Business Corporations in India 1851-1900 (No. 8). CUP Archive.
Sutton, John, Kellow, Nebil, 2010. An Enterprise Map of Ethiopia. International Growth Centre, London.
Sutton, John, Kpentey, Bennet, 2012. An Enterprise Map of Ghana. International Growth Centre, London.
Timberg, T.A., 1978. The Marwaris, from Traders to Industrialists. Stosius Incorporated/Advent Books Division.
Tripathi, D., 2004. The Oxford History of Indian Business. Oxford University Press.
Tripathi, D., Mehta, M., 1990. Business Houses in Western India: A Study in Entrepreneurial Response. Jaya books, pp. 1950–1956.
Upadhyay, S.B., 1990. Cotton Mill Workers in Bombay, 1875 To 1918: Conditions of Work and Life. Economic and Political Weekly, July 28 1990, PE-87–99.
Webb, S.B., 1980. Tariffs, cartels, technology, and growth in the German steel industry, 1879 to 1914. J. Econ. Hist. 40 (2), 309–330.