Journal of Monetary Economics 135 (2023) 116–133

Contents lists available at ScienceDirect

Journal of Monetary Economics

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

How does caste affect entrepreneurship? birth versus worth[☆]

Sampreet Singh Goraya[*]

Department of Economics, Stockholm School of Economics, Sveavägen 65, Stockholm 113 83, Sweden

ARTICLE INFO

Article history:
Received 27 October 2021
Revised 13 January 2023
Accepted 17 January 2023
Available online 20 January 2023

JEL classification:
O11
E44
D24
D61

Keywords:
Misallocation
Caste system
Credit constraints
Entrepreneurship

ABSTRACT

Informal institutions play an important role in resource allocation across entrepreneurs in developing countries. I focus on the caste system in India and document three stylized facts. First, entrepreneurs of historically disadvantaged castes have a higher average revenue product of capital, *arpk*, relative to high-caste enterprises. Second, cross-caste differences in *arpk* are driven by small enterprises. Third, the majority of these differences are concentrated in financially underdeveloped regions. In a model of entrepreneurship, I find that the cross-caste differences in *arpk* are explained by differences in access to credit and that such asymmetries reduce output per capita by 5.6%.

1. Introduction

A large body of literature has argued that the misallocation of resources explains a substantial fraction of cross-country differences in aggregate productivity.[1] A number of market-oriented distortions, such as financial frictions, labor market regulations, and size-dependent policies, among others, have been proposed as being responsible for resource misallocation. However, we lack systematic evidence about the quantitative importance of informal institutions in generating aggregate misallocation.

This paper quantifies the effects of one such institution – the caste system in India – on aggregate productivity. In particular, I explore the hypothesis that "birth and not worth" – that is, the caste and not the productivity of individuals – determines how resources are allocated in the economy. Historically, the caste system sorted people into different occupations

☆ This paper is a substantially revised version of Chapter 1 of my PhD dissertation. I thank the editor Urban Jermann, the associate editor Pierre-Daniel Sarte, an anonymous referee for their helpful comments. I also thank my advisors Julian di Giovanni, Manuel Garcia Santana and Alessandro Tarozzi for their support, and Andrea Caggese, Christian Fons-Rosen, and Christopher Woodruff, as well as participants of CREI international lunch, 2nd CESC conference, V winter workshop MOVE/UAB, NEUDC Dartmouth 2020, PEDL conference 2020, Royal Economic Society conference 2021, MWIEDC Northwestern 2021, EUDN 2021, ASWEDE 2021, ACEGD 2021, World Bank ETIFE, and NBER Entrepreneurship 2022 for helpful discussions. I would like to thank Lakshmi Iyer for generously sharing state-level aggregates of the Economic Census.

* Corresponding author.:
E-mail address: sampreet.goraya@hhs.se

[1] See Banerjee and Duflo (2005), Restuccia and Rogerson (2008), Guner et al. (2008), De Mel et al. (2008), and Hsieh and Klenow (2009). See Hall and Jones (1999) and Caselli and Feyrer (2007) for cross-country productivity differences.

https://doi.org/10.1016/j.jmoneco.2023.01.007
0304-3932/

and constrained mobility, suppressing the entrepreneurial potential of a vast portion of society. While mobility restrictions for certain castes have weakened over time, the caste system remains a salient feature of India.[2]

I use firm-level and find outcomes that are consistent with the presence of high levels of *caste-driven* resource misallocation. My empirical analysis exploits data from the Micro, Small and Medium Enterprises (MSME) survey of 2006–2007. This dataset contains a representative sample of MSMEs together with an exhaustive list of balance sheet variables, and information on the caste of the enterprise owner and employees, a feature missing in other commonly used firm-level datasets in India.[3] Using this dataset, three stylized facts are established.

First, on average within a sector, low-caste (LC) and middle-caste (MC) entrepreneurs have 25%-30% and 13%-22% higher average revenue product of capital (*arpk*), respectively, relative to high-caste (HC) entrepreneurs with similar characteristics. Furthermore, non-HC entrepreneurs are characterized by lower credit-to-capital and credit-to-output ratios relative to HC entrepreneurs. Second, most of the cross-caste *dispersion* in *arpk* is driven by small entrepreneurs. In particular, moving from the smallest to the largest entrepreneur in the economy, *arpk* for LC entrepreneurs declines from being 52% higher than that of HC enterprises to approximately 12%.

Third, cross-caste *arpk* differences negatively correlate with regional financial development. In particular, state-level credit-to-output ratio is used as a measure of financial development. The observed differences in *arpk* across castes shrink as the credit-output ratio increases.[4] It is observed that LC enterprises have an *arpk* that is double the value of HC enterprises in the states with the lowest financial development, such as Bihar, Jharkhand, and Uttar Pradesh, whereas no such differences are observed in states with well-functioning financial markets.

To rationalize these facts, I build a quantitative model of entrepreneurship in the spirit of Lucas (1978), Quadrini (2000), and Buera et al. (2011). The model has two sectors of production: the corporate sector and the noncorporate sector (also called the entrepreneurial sector), and both produce a single good. The corporate sector comprises of a representative firm, whereas the noncorporate sector is made up of entrepreneurs.

The households from different castes can choose to become either entrepreneurs or workers in the context of *caste-dependent access to credit*. Moreover, the model allows for intertemporal savings to capture the self-financing channel, as in Banerjee and Moll (2010), Midrigan and Xu (2014), and Moll (2014).[5] The quantitative predictions crucially depend on the identification of three parameters that govern the degree of financial frictions in the economy. I use caste-specific credit-to-output ratios to calibrate these parameters.

The model finds substantial differences in access to credit across castes. In the benchmark calibration, the model identifies the degree of financial frictions as 56% and 51% more stringent for LC and MC entrepreneurs, respectively, relative to those of HC entrepreneurs. Furthermore, the model can explain the majority of the value computed in the data for the cross-caste differences in *arpk*. In particular, the *arpk* of LC and MC entrepreneurs is 31.6% and 17.7% higher, respectively, than that of HC entrepreneurs.

In line with the data, the model predicts lower household income and wealth for non-HC households relative to HC households. Moreover, the model predicts more concentration of wealth and income among HC households relative to non-HC households. The difference is a result of the lower wealth creation among the entrepreneurs of non-HC households, as they face stricter borrowing constraints. Therefore, to a certain extent, the model rationalizes the persistent cross-caste disparities in wealth and income.

Finally, regional financial development is modeled as a shock to credit supply that affects all castes proportionately. I consider each region as a separate general equilibrium economy and compare the steady-state levels of *arpk* differences across castes. The model predicts a declining cross-caste *arpk* differences with a rise in regional financial development, as it particularly benefits relatively more constrained entrepreneurs, who are concentrated among non-HC castes.

In what follows, I conduct various counterfactual exercises using the model. First, non-HC entrepreneurs are granted a borrowing capacity that is similar to their HC counterparts. The model identifies gains of 5.6% in output per capita. Further, the model is used to decompose output gains at the extensive and intensive margins. The reallocation of capital from unproductive HC entrepreneurs to more productive non-HC entrepreneurs increases the allocative efficiency of the economy; as a result, cross-caste *arpk* differences fall significantly, and the overall dispersion in *arpk* declines by 24.1%. These changes result in a 3.1% rise in output per capita. The reduction in borrowing constraints induces the entry of more non-HC entrepreneurs. The share of LC enterprises increases from 16.0% in the benchmark economy to 22.0%. Moreover, the excess entry of entrepreneurs increases demand for capital and labor. This higher demand would induce an increase in the interest rate by approximately 0.5 percentage points, which further leads to the exit of unproductive but wealthy HC entrepreneurs. This improvement in selection increases output per capita by a further 2.5%. In light of these results, I conclude that these caste-specific asymmetries in the financial markets are an important source of misallocation in India, as highlighted in Hsieh and Klenow (2009).

[2] See Munshi (2016). Traditionally, entrepreneurship and financial intermediation belonged to one group called "Vaishyas"; however, these occupations have spilled over to other high castes such as "Brahmins" and "Kshatriyas"; see Damodaran (2008). The high castes represent 31.2% of the population, whereas low and middle castes represent 29.5% and 39.3% of the population, respectively.

[3] The most commonly used datasets are the Annual Survey of Industries and Prowess. They do not include the caste of the enterprise owner.

[4] The results remain qualitatively similar with other proxies of regional financial development such as the number of rural and commercial banks per capita, share of households with bank accounts, and share of households with loans.

[5] See also Gopinath et al. (2017) and Buera and Shin (2013).

Literature review: This paper contributes to the literature on race and ethnicity-based distortions, and misallocation. Hsieh et al. (2019), and Cassan et al. (2021) focus on the allocation of talent across occupations.[6] Hjort (2014) shows that inter-ethnic rivalries generate misallocation. Banerjee and Munshi (2004) document the misallocation of capital across communities in Tirupur (India). This paper quantifies the *aggregate* effects of caste-specific misallocation.

This paper also builds on the work of Thorat and Sadana (2009), Iyer et al. (2013), Deshpande and Sharma (2013), and Jodhka (2010), who document substantial caste differences in entrepreneurship rates, employment, and growth rates in India. Fisman et al. (2017) provide evidence on the importance of caste match between lender and borrower for the access to credit.[7] This paper formalizes the idea of caste-specific borrowing limits.

This paper is related to the papers that link finance, wealth, and entrepreneurship, such as Quadrini (2000) and Cagetti and De Nardi (2006), among others. It quantifies the effects of caste-dependent financial frictions on entrepreneurship, wealth and income inequality.

My methodology can be used to quantify the effects of discrimination arising from various other informal institutions and norms. In this sense, this paper also links to a broader literature regarding the economics of discrimination: see, for instance, Becker (1957), Phelps (1972), Arrow (1973), and Akerlof (1976).

The remainder of the paper is organized as follows. Section 2 describes the institutional setup. Section 3 describes the data. Section 4 documents empirical facts. Section 5 presents the theoretical framework, and Section 6 presents the main results. Section 7 concludes.

2. Institutional setup: The caste system

The caste system is a form of social stratification that divides people into rigid hierarchical groups based on their occupation. For centuries, caste dictated customary social interaction, exclusion, and endogamy. In a high to low hierarchical order these groups are the Brahmins (priests and teachers), Kshatriyas (rulers and soldiers), Vaishyas (merchants and traders), and the Sudras (laborers and artisans). Further, there are two additional groups that fall outside the caste system. The first one embodies the group of people traditionally known as Dalits.[8] The second group of people is known as Scheduled Tribes. Scheduled Castes and Scheduled Tribes have been subject to various forms of discrimination including barriers to access capital and firm creation. For the remainder of the paper, I ignore the micro structure of the caste system and primarily focus on a very broad definition. The low-caste individuals are denoted by “LC,” which includes the Scheduled Castes and Scheduled Tribes. The middle-caste individuals are denoted by “MC,” which includes the Sudras (also known as Other Backward Castes, OBC, which fall between the traditional upper castes and the lowest castes). The high caste, denoted by “HC,” covers the rest.

3. Description of the data

My main data source is the Micro, Small and Medium Enterprises (MSME) data from 2006 to 2007 (MSME, 2009), the details of which are provided below. The details for other supplementary datasets are in Appendix.[9]

MSME Dataset: The MSME dataset contains a representative sample of the *MSME sector* that comprises enterprises that invest less than INR 100 million (manufacturing enterprises) and INR 50 million (service enterprises).[10] An enterprise is a firm. The MSME dataset has two parts: a census of registered MSMEs and a sample survey of unregistered MSMEs (see Appendix for more details). A total number of 126,169 enterprises are surveyed to capture a representative sample of unregistered MSMEs. There are 1.65 million observations in total, of which 179,041 do not provide information on the wage bill. Finally, 1.45 million observations (1.12 million in the manufacturing sector) are left after the cleaning process, which is described in detail in Appendix. The descriptive statistics for manufacturing are provided in Table 1. As shown in Panel A.1, HC, MC, and LC represent 51%, 39.1%, and 9.8% of observations in the sample. In terms of total output, HC, MC, and LC enterprises produce 61.3%, 30.7% and 8%, respectively, whereas in terms of total credit, 75.1% is allocated to HC enterprises and only 4.7% to LC enterprises.

The enterprise-level statistics are provided in Panel A.2 of Table 1. LC and MC enterprises are smaller and have lower leverage (defined as the ratio of credit to capital stock) than HC enterprises. Next, the retrospective information on output is exploited from the dataset to compute growth rates. The average growth rate of output is around 9% and does not vary

[6] See Erosa et al. (2017), and Morazzoni and Sy (2022) for misallocation across gender in the US.

[7] They find that a lender of a given caste increases credit access and reduces collateral requirements for a borrower of the same caste. In general, it is more likely that an owner of a bank, a bank manager, or a loan officer is an HC individual.

[8] In the Indian constitution, Dalits have fallen under the category of Scheduled Castes since 1947.

[9] I also compute various statistics with the Economic Census of India (EC) 2005, Indian Human Development Survey 2005 (Desai, Vanneman and National Council of Applied Economic Research, 2018), the Annual Survey of Industries (ASI) 2006, and the National Sample Survey (NSS) 2005–2006. The main advantage of using the MSME and EC is that they provide the caste of the enterprise owner together with other financial variables. However, MSME is not as commonly used as the ASI-NSS or the CMIE Prowess because it omits large enterprises (above a certain threshold of capital stock), nor does it have a panel dimension.

[10] The MSME sector is estimated to employ about 59 million individuals in over 26.1 million enterprises throughout the country. Further, 1.5 million (5.94%) are registered MSMEs and 24.5 million (94.06%) are unregistered MSMEs. The registration is under the Factories Act of 1948 (See Appendix for more details).

Table 1
Summary Statistics: Manufacturing Sector & Population Shares.

	HC			MC		LC		Overall
	Total	%		Total	%	Total	%	Total
	PANEL A.1: MSME 2006–2007 - Aggregate Statistics							
Observations (000s)	572	51%		438	39.1%	109	9.8%	1121
Enterprises (000s)	2915	39%		3504	46%	1144	15%	7563
	Mean	%		Mean	%	Mean	%	Mean
Employment	3.41	46.7%		2.53	41.5%	2.18	11.7%	2.82
Output	285,495	61.3%		119,040	30.7%	95,227	8.0%	179,594
Capital	679,811	70.3%		195,465	24.3%	133,282	5.4%	372,573
Credit	27,689	75.1%		6158	20.1%	4462	4.7%	14,200
	PANEL A.2: MSME 2006-2007- Enterprise-level Statistics							
	HC		MC		LC	All		
	Mean (S.D.)		Mean (S.D.)		Mean (S.D.)	Mean (S.D.)		
log(Output/Capital)	-0.21 (1.36)		0.02 (1.22)		0.25 (1.22)	-0.03 (1.29)		
log(Output/Labor)	0.63 (1.00)		0.60 (0.92)		0.88 (1.08)	0.66 (0.99)		
Credit/capital	3.6%		1.6%		2.1%	2.4%		
Credit/Output	8.4%		4.2%		2.6%	5.6%		
Δ Output (Mean)	9.4% (0.29)		9.0% (0.23)		9.2% (0.25)	9.2% (0.26)		
Default Rate	13.6%		13.1%		14.5%	13.8%		
	PANEL B: Economic Census 2005							
	HC			MC		LC		Overall
	Total	%		Total	%	Total	%	
Enterprises (000s)	2924	36.5%		3818	47.5%	1286	16%	8029
	Mean	%		Mean	%	Mean	%	
Employment	3.86	49%		2.39	40%	2.0	11%	2.9
PANEL C: ASI-NSS 2005–2006								
	Overall		Noncorporate sector					
	Output	Labor	Output	Labor				
Average (all)	437,258	3.3	191,864	2.9				
Average (MSME)	161,107	2.9	133,048	2.8				
Share of MSME in ASI-NSS	36.9%	89.8%	70%	97.3%				
PANEL D: IHDS								
	HC	MC	LC					
Population Share	31.2%	39.3%	29.5%					
Share of Entrepreneurs	5.5%	4.6%	2.3%					

Notes: The table reports descriptive statistics. LC, MC, and HC represent the low-caste, middle-caste, and high-caste entrepreneurs, respectively. **Panel A.1** reports statistics for the MSME dataset. For each caste, the first column reports the total number in the case of Observations, and Enterprises, and the mean value for Employment, Output, Capital, and Credit. The second column reports respective caste shares in the overall dataset. The "Observations" report the number of enterprises available in the dataset. For total "Enterprises", I compute a total number of MSME enterprises by multiplying the observations with their respective sampling multiplier provided in the data. Employment is measured as the number of employees (e.g., the average number of employees for HC enterprises is 3.41, and HC enterprises employ 46.7% of the overall workforce), and output, capital, and credit are in Indian rupees. Sampling multipliers are applied.

Panel A.2 reports enterprise-level statistics for MSMEs. S.D. is the standard deviation. The credit-output and credit-capital ratios are reported for all enterprises (denoted as zero if no credit). Output is value-added and credit is the number of outstanding loans. Sampling multipliers are applied.

Panel B provides summary statistics for manufacturing enterprises in the Economic Census 2005. Enterprises represent the total number of productive units for each caste in the data (and their share). Employment reports the mean value of the number of employees for each caste (and the share of the workforce employed by entrepreneurs of each caste).

Panel C provides statistics for the ASI-NSS 2005. The noncorporate sector includes privately owned and private limited enterprises, Khadi and village industries, and handlooms. Sampling multipliers are applied for MSME and ASI-NSS datasets. MSMEs in the ASI-NSS dataset are enterprises with Investment< INR 100 million.

Panel D provides statistics for population shares as computed with IHDS data. I compute population shares as the average between two datasets (IHDS 2004 and 2012). The share of entrepreneurs is computed as the number of entrepreneurs (individuals who report business as their main activity) as the share of the population (variable is available in IHDS 2012).

a lot across castes, whereas the dispersion in output growth is higher for HC enterprises relative to LC and MC enterprises. Further, the default rates for each caste are computed. Around 14.5% of LC enterprises and 13.6% of HC enterprises defaulted in the last 12 months.[11]

Here, the quality of MSME data is assessed. I compare the coverage of MSME data with other representative firm-level datasets, such as the Economic Census 2005 and the ASI-NSS 2005–2006. In the Economic Census 2005, enterprise ownership across castes is similar to the MSME dataset (see Panel B of Table 1).[12] Enterprises under the MSME threshold in the ASI-NSS dataset represent 37% of the total output and 70% of output in the noncorporate sector (see Panel C in Table 1). The noncorporate sector includes proprietary or Hindu joint family firms, partnership firms, private limited companies, khadi,

[11] These numbers suggest that, at least among MSMEs, riskiness may not be the primary driver of low credit allocation to LC entrepreneurs.

[12] In this dataset, I cannot separate out the MSMEs as there is no information on the capital stock; therefore, results for the whole manufacturing sector are reported.

Table 2
Sectoral Shares of Output across Castes.

Sector	NIC	LC Share	MC Share	HC Share	Sector Share
Manufacture of food products and beverages	15	0.016	0.054	0.106	0.090
Manufacture of wearing apparel	18	0.013	0.049	0.052	0.035
Manufacture of tobacco products	17	0.009	0.040	0.079	0.081
Manufacture of furniture; manufacturing n.e.c.	36	0.008	0.035	0.036	0.036
Manufacture of fabricated metal products	28	0.006	0.027	0.057	0.038
Manufacture of wood and of products of wood	20	0.006	0.014	0.017	0.015
Manufacture of other non-metallic mineral products	26	0.003	0.017	0.043	0.051
Manufacture of machinery and equipment n.e.c.	29	0.003	0.014	0.044	0.057
Tanning and dressing of leather	19	0.002	0.003	0.008	0.008
Manufacture of rubber and plastics products	25	0.002	0.006	0.024	0.024
Manufacture of chemicals and chemical products	24	0.002	0.009	0.037	0.133
Manufacture of basic metals	27	0.002	0.006	0.028	0.113
Publishing, printing and reproduction of recorded media	22	0.002	0.008	0.014	0.017
Manufacture of tobacco products	16	0.001	0.004	0.007	0.020
Manufacture of electrical machinery and apparatus n.e.c.	31	0.001	0.007	0.017	0.035
Manufacture of radio, television and communication equipment	32	0.001	0.002	0.004	0.015
Manufacture of other transport equipment	35	0.001	0.003	0.011	0.025
Manufacture of medical, precision and optical instruments	33	0.001	0.003	0.007	0.008
Manufacture of paper and paper products	21	0.000	0.003	0.012	0.014
Manufacture of motor vehicles, trailers and semi-trailers	34	0.000	0.002	0.005	0.064
Manufacture of coke, refined petroleum products and nuclear fuel	23	0.000	0.001	0.002	0.115
Manufacture of office, accounting and computing machinery	30	0.000	0.000	0.002	0.007
Recycling	37	0.000	0.000	0.000	0.000
Total		0.080	0.307	0.613	1.000

Notes: HC, MC, and LC stands for high-caste, middle-caste and low-caste, respectively. The table is sorted in terms of LC's output share (largest to smallest). The variables "LC share","MC share" and "HC share" contain output share of LC, MC, and HC enterprises in that sector, respectively (computed in MSME dataset). The "sector share" contains sectors' output share in the national manufacturing output (computed with ASI-NSS dataset). The Sampling multipliers are applied. The sector names are truncated for presentation purposes and to see the full name, refer to NIC 2004 classification.

and village industries and handlooms, and the rest is placed under the corporate sector. In terms of enterprise-level characteristics, the average number of employees for MSMEs in the ASI-NSS dataset is 3.0.[13]

4. Empirical analysis

In this section, the allocation of entrepreneurs across castes and sectors in the Indian manufacturing is documented. Further, it illustrates the observed differences in the average product of capital, *arpk*, and the average product of labor, *arpl*, across castes.

The population shares of HC, MC and LC are 31.2%, 39.3%, and 29.5%, respectively (see Panel D of Table 1). The share of entrepreneurs (computed as the ratio of entrepreneurs over the total population) for each caste is 5.5% (HC), 4.6% (MC), and 2.3% (LC). In Table 2, the allocation of entrepreneurs across sectors is documented. There are twenty-three 2-digit sectors within Indian manufacturing. The sectoral output share is similar across castes. For instance, "Manufacture of food products and beverages" represents the biggest chunk of output produced in the manufacturing sector and also represents the largest share of output produced by enterprises of each caste. Next I delve into the caste-specific difference in the average product of inputs.

Fact 1: arpk is high for LC and MC enterprises.

For enterprise i with owner of caste s, $arpk_{is} := \ln(ARPK_{is}) = \ln(Y_{is}) - \ln(K_{is})$ and $arpl_{is} := \ln(ARPL_{is}) = \ln(Y_{is}) - \ln(L_{is})$. The variable Y_{is} is gross value added, K_{is} is capital, and L_{is} is labor input, measured as the wage bill.

To evaluate *arpk* differences and to control for sector and regional heterogeneity, the following regression is used

$$ln\, Y_i = \beta_0 + \beta_1 \mathbf{1}_{L-CASTE} + \beta_2 \mathbf{1}_{M-CASTE} + \sigma' \Gamma + \epsilon_i. \quad (1)$$

The dependent variables are $\{arpk, arpl\}$. The explanatory variables are the dummies for low castes, $\mathbf{1}_{L-CASTE}$, and the middle castes, $\mathbf{1}_{M-CASTE}$, whose corresponding coefficients are β_1 and β_2. The estimators $\widehat{\beta_1}$ and $\widehat{\beta_2}$ are interpreted as the percentage difference in the dependent variable between low- and high-caste enterprises and the middle- and high-caste enterprises, respectively. The regressions include four-digit sector and district fixed effects.[14]

[13] According to the reports generated by the Ministry of Micro, Small and Medium Enterprises in India, the threshold for a firm to be defined as MSME is in "cumulative investment in plant and machinery (original cost)" (see Garcia-Santana and Pijoan-Mas (2014) for more details). This variable is available in the ASI-NSS dataset as the *value of plant and machinery owned by the firm*.

[14] Additionally, there is a vector of controls, Γ, that includes gender and religion of the owner fixed effects.

Table 3
ARPK and ARPL across Castes.

	(1) log(ARPK)	(2) log(ARPL)
M-caste	0.13***	-0.05***
	(0.02)	(0.00)
L-caste	0.25***	-0.02
	(0.02)	(0.35)
Constant	-0.67***	0.61***
	(0.01)	(0.01)
Obs.	1,121,610	1,084,632
R-squared	0.51	0.32
District FE	✓	✓
NIC4 FE	✓	✓

Notes: The table shows results from the enterprise level regression using Eq. 1 for manufacturing sector. The variables log(ARPK) and log(ARPL) are the average products of capital and labor, respectively. M-caste is the dummy variable for the middle-caste enterprises. L-caste is the dummy variable for the low-caste enterprises. All specifications include sector, district (586 in total), gender of the owner and religion fixed effects. Standard errors are in parentheses, clustered at the caste, state and sector level. To see the regression estimates with sampling weights, refer to Table in Appendix. In that case, MC and LC enterprises have 22% and 30% higher *arpk* relative to HC enterprises. Further, differences in *arpl* are also larger but remain smaller than that of *arpk*. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

My estimates show that MC and LC enterprises have 13% and 25% higher *arpk*, whereas *arpl* is 5% higher for MC enterprises. No significant differences in *arpl* for LC enterprises relative to HC enterprises are observed (see Table 3). Further, the results of regressions with sampling weights are that MC and LC enterprises have 22% and 30% higher *arpk* relative to HC enterprises (refer to Table in Appendix). The robustness checks are provided by including control variables that may be correlated with caste. Here, enterprise-level wages, mean schooling rates, land holdings, and dispersion in output growth, computed at the sector-state-caste level are included. The *arpk* differences across castes remain high (see Table in Appendix).

To compare the cross-caste *arpk* differences for different sectors, I compute *arpk* differences in the 2-digit sectors. The *arpk* is also substantially higher relative to HC in almost all the sectors. Further, in the sectors where LC output share is largest, the *arpk* is also substantially higher relative to HC enterprises (see Table in Appendix).

Finally, given the fact that cross-caste differences in *arpk* are much larger and more robust across various specifications compared with *arpl*, from here onward, the analysis will primarily focus on the observed differences in *arpk* and explores its various facets.[15]

Fact 2: ARPK differences across castes decline with size.

Here, I document how *arpk* differs according to enterprises size. To do so, enterprises are divided into five different size bins, defined by output, and β_1 and β_2 are computed for each size bin using the regression model described in Eq. 1 (See Table in Appendix for distribution of enterprises across different size bins). As shown in Fig. 1 a, large *arpk* differences persist among smaller enterprises, but these differences in *arpk* shrink as enterprise size increases. In fact, the *arpk* is highest for LC enterprises in the lowest decile of the size distribution, +52% relative to that of HC enterprises, and it declines by 40 percentage points (p.p.) and stands at +12% for enterprises in the top decile of the size distribution. A similar decline is observed for MC enterprises as well (see Fig. 1 b). In this case, the *arpk* differences between MC and HC enterprises decline by 14 p.p as enterprise size increases.[16] To a certain extent, these results are consistent with the existing evidence on financial frictions and enterprise size (Hadlock and Pierce, 2010), which suggests that small entrepreneurs are relatively more constrained. These results will also motivate the inclusion of the size-dependent borrowing capacity of entrepreneurs in the quantitative analysis.

Fact 3: ARPK differences decline with financial development.

This section provides evidence for how regional financial development changes cross-caste dispersion in *arpk*, the capital-labor ratio, and profitability. Ayyagari et al. (2014) document large and persistent differences in financial development across states in India. The credit-to-output ratio for each state is used as a measure of financial development.[17] The regression

[15] Oh (2019) shows how caste affects labor supply in India. However, the inquiry of caste-specific labor misallocation is left for future projects.

[16] A similar but less steep convergence is also documented over enterprise age; however, in this case, substantial *arpk* differences remain even for older enterprises in the sample. For instance, the *arpk* is highest for young LC entrepreneurs, +26% relative to that of HC enterprises, and it declines by 4 p.p. over age (see Appendix for details).

[17] State-wise indicators on GDP and domestic credit are taken from the RBI's Handbook of Statistics on Indian states.

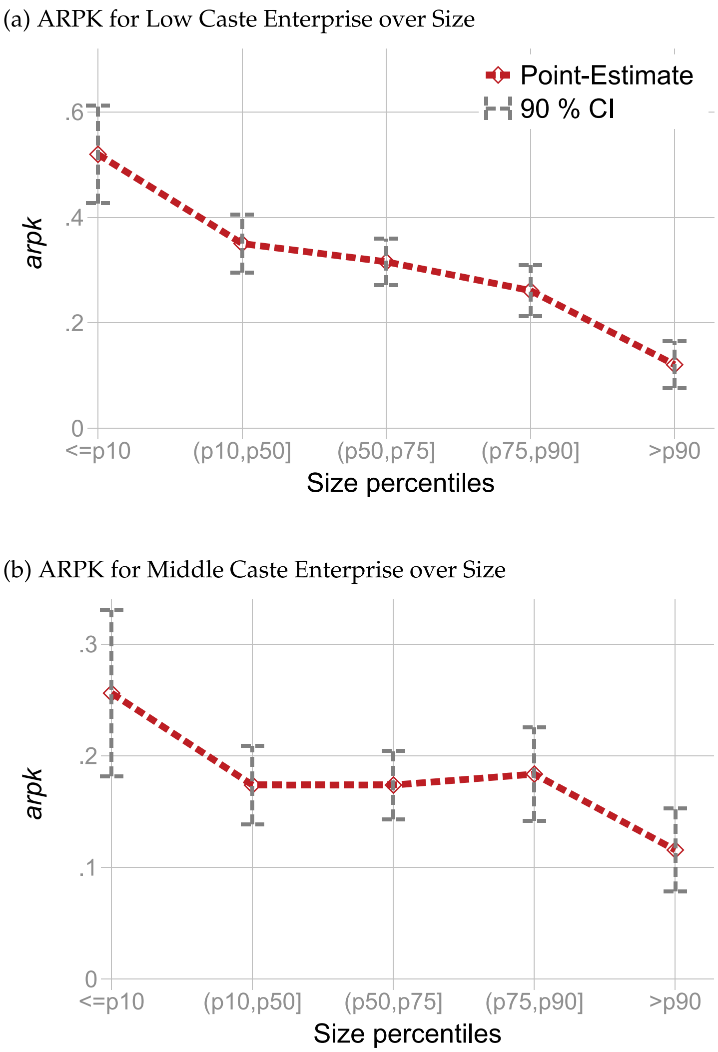

Fig. 1. Caste Differences in ARPK over Enterprise Size Note: This figures plots the evolution of ARPK over enterprise size (defined as output). The size bins are defined as follows: (i) p10: enterprises below the 10th percentile, (ii) p10-p50: enterprises between the 10th and 50th percentile (iii) p50-p75: enterprises between the 50th and 75th percentiles, (iv) p75-p90: enterprises between the 75th and 90th percentile, and (v) p90: enterprises above the 90th percentile. I run regression specification 1 for each size bin and plot β_1 in Panel a, and β_2 in Panel b. The standard errors are clustered at the caste, state and sector level. To see the regression estimates with sampling weights, refer to Fig. in Appendix. The results are qualitatively similar.

model is used to pin down the elasticity of $\{arpk, k/l, profits\}$ to financial development for LC and MC entrepreneurs, where the caste dummies are interacted with the financial development of states Fd_s. The regression specification is

$$\begin{aligned} ln\,Y_i = \hat{\gamma}_0 + \hat{\gamma}_1 \mathbf{1}_{L-CASTE} + \hat{\gamma}_2 \mathbf{1}_{L-CASTE} \times Fd_s + \hat{\gamma}_3 \mathbf{1}_{M-CASTE} \\ + \hat{\gamma}_4 \mathbf{1}_{M-CASTE} \times Fd_s + \hat{\gamma}_5 Fd_s + \Gamma + \epsilon_i, \end{aligned} \quad (2)$$

where $\hat{\gamma}_2$ and $\hat{\gamma}_4$ represent the elasticity of the dependent variable to Fd_S for LC and MC entrepreneurs, respectively, with respect to HC entrepreneurs. The value of $\hat{\gamma}_2$ and $\hat{\gamma}_4$ is negative and significant for *arpk*, suggesting an improved allocation of capital across castes in regions with high financial development (see Table 4). Further, LC and MC entrepreneurs, who are less capital intensive relative to HC entrepreneurs on average, increase their capital-labor ratio as regional financial development increases.

Finally, I analyze how better financial conditions affect enterprise profitability.[18] In principle, fewer financial constraints lead to more entry of enterprises at the margin (Cagetti and De Nardi, 2006), which analyzes the role of borrowing constraints as determinants of entrepreneurial decisions). This implies that as LC enterprises are able to access easy credit, one should expect more entry of enterprises and lower profitability on average, as the productivity threshold to become an en-

[18] The profitability is $\frac{Y-wL-RK}{Y}$, where Y is value added, wL is the wage bill, and RK is the cost of capital, respectively.

Table 4
Regional Financial Development and Capital Allocation .

	log(ARPK)	log(ARPK)	log(K/L)	log(K/L)	Profit Rate	Profit Rate
M-caste	0.13***	1.28***	-0.26***	-1.28***	0.02**	0.70***
	(0.02)	(0.21)	(0.02)	(0.18)	(0.01)	(0.14)
L-caste	0.25***	1.32***	-0.49***	-1.43***	0.05***	0.68***
	(0.03)	(0.21)	(0.03)	(0.19)	(0.02)	(0.15)
FD		0.82***		-0.28		0.69***
		(0.20)		(0.17)		(0.14)
M-caste × FD		-1.37***		1.08***		-0.81***
		(0.23)		(0.22)		(0.15)
L-caste × FD		-1.16***		0.90***		-0.69***
		(0.23)		(0.22)		(0.15)
Observations	1,119,508	1,119,508	1,119,508	1,119,508	1,119,508	1,119,508
R-squared	0.51	0.13	0.55	0.23	0.46	0.09
District FE	✓	-	✓	-	✓	-
NIC4 FE	✓	✓	✓	✓	✓	✓

Note: Results from the enterprise-level regression using Eq. 2. The Profit Rate is IHS transformed and shown in column headings. *Fd* is an index of financial development across states. The variables ARPK and K/L are the average products of capital and the capital-labor ratio. M-caste is the dummy variable for middle-caste enterprises and L-caste is the dummy variable for low-caste enterprises. All specifications include gender of owner and religion fixed effects. The standard errors are in parentheses, clustered at the caste, state and sector level. Robust standard errors are in parentheses, clustered at the caste, region and sector level. To see the regression estimates with sampling weights, refer to Table in Appendix. The results are qualitatively similar. *** p<0.01, ** p<0.05, * p<0.1.

trepreneur decreases. In the data, LC and MC enterprises are more profitable than HC enterprises on average; however, these differences shrink as the regional financial condition improves (see Table 4).

Several factors other than the credit-to-output ratio may differ across regions. For instance, dispersion in the capital-to-output ratio varies substantially across regions (see Table in the appendix). In Table in the appendix, the estimates of regression model 2 with a set of regions with a comparable dispersion in capital-to-output ratio are reported.[19] Further, the distribution of LC households across regions is explored. The caste composition of the population for each region is reported in Table in the appendix. The LC's population share is used as a control variable in regression 2 and estimates are reported in the Table in the appendix. The results remain qualitatively similar. Finally, several other proxies of regional financial development are used in the regression model 2 such as the number of rural and commercial banks per capita, share of households with bank accounts, and share of households with loans. Similar results are documented across various specifications (see Table B.14, B.15, and B.16 in Appendix B).

In this section, three stylized facts are documented that suggest that capital is misallocated across castes. However, the welfare implications of these facts are not straightforward and depend on various general equilibrium forces. In the next section, I describe a quantitative model that will be used to illustrate the macroeconomic implications of caste-specific distortions in the capital markets.

5. Theoretical framework

The model is an extension of the frameworks used in Buera and Shin (2013) and Quadrini (2000). Time is discrete, and there is a measure M of infinitely lived agents that are heterogeneous across productivity z, assets a, and caste s. Every period, agents choose to become either workers or entrepreneurs based on their wealth a and entrepreneurial productivity z, and this occupational choice is represented by o_t. Financial wealth is determined endogenously, whereas productivity follows a stochastic process such that agents retain their last-period productivity z_t with probability ψ, and with probability $1-\psi$, they draw their new productivity from a Pareto distribution whose probability density is $\eta z^{-(\eta+1)}$ for $z \geq 1$. The parameter ψ represents the persistence, whereas η captures the dispersion in the productivity process. The households' caste is constant over time. There are three sectors: the household sector, the production sector, and the intermediation sector. I start with a description of the household sector.

5.1. Households

Preferences: Agents' utility functions are strictly increasing and concave and satisfy standard Inada conditions. Agents discount their future utility at a discount rate ρ, and at any point in time t, their preferences are represented by the function

$$\mathbb{E}\sum_{t=o}^{\infty}\rho^t\frac{c_t^{1-\gamma}-1}{1-\gamma}.$$

[19] The restricted sample about 89% of the enterprises in the data.

5.2. Production sector

There are two sectors of production: the noncorporate sector and the corporate sector. These two sectors differ in terms of the technology they employ to produce a single homogeneous good that is used for both consumption and investment. The output price is normalized to one. The differentiation between these sectors tries to capture the nature of enterprises. The noncorporate sector is made up of entrepreneurs who will operate small and risky enterprises, whereas the corporate sector represents large conglomerates (e.g., enterprises listed on the stock market). I assume perfect competition in the output market.

Corporate Sector: Following Quadrini (2000), the corporate sector has a representative enterprise that uses constant returns to scale production technology, given by

$$Y_c = F(K_c, N_c) = K^{\theta_c} N^{1-\theta_c},$$

where Y_c is output produced, K_c is the capital hired, and N_c is the amount of labor hired. The capital depreciates at the rate δ.

Noncorporate Sector: The noncorporate sector is populated by households that choose to become entrepreneurs. These entrepreneurs have a decreasing returns to scale production function $f(z, k, l) = z\,(k^{\alpha} l^{\beta})^{1-\nu}$, where $\alpha + \beta = 1$ and $1 - \nu$ is the span-of-control parameter that varies between 0 and 1. An entrepreneur rents capital k in the financial market and hires labor l to produce y units of a single good. Entrepreneurs incur a caste-dependent per-period fixed cost of operation κ_s.

5.3. Financial markets

There is an intermediary that receives deposits from savers and lends these funds to entrepreneurs and the corporate sector. The lending activity is based on a constant returns to scale technology with a proportional cost per unit of funds intermediated. While this cost is zero for the corporate sector, lending to entrepreneurs has a proportional cost d. The cost of capital for entrepreneurs is $r_t + \delta + d$ in period t, where δ is a time-invariant depreciation cost and r_t is the deposit rate, and the cost of capital for the corporate sector is $r_t + \delta$. The corporate sector is unconstrained in its ability to borrow. However, entrepreneurs face borrowing constraints of the form:[20]

$$(1 + r_t + \delta + d)k_t \leqslant \phi\, \lambda_s\, a_t; \quad a_t \geqslant 0, \tag{3}$$

where $\phi \in [\underline{\phi}, \overline{\phi}]$ is a common component and $\lambda_s \in [1, \infty)$ is a caste-specific component of the degree of credit constraints for entrepreneurs. The ϕ can be interpreted as a common shock to the credit environment that affects all castes proportionally. For a given ϕ, a large value of λ_c means tat entrepreneurs face low credit constraints and have high borrowing capacity.[21] For the benchmark economy, $\phi = 1$.

5.4. Recursive formulation of Individuals' problem

Agents maximize their expected utility for a given set of factor prices $\{w, r\}$, their asset base a, productivity z, and a vector of probabilities corresponding to future productivity z' given by $d\Upsilon(z'|s)$, such that the resource constraint always binds. The value function that agents maximize is

$$V(a, z, s) = \max\{V^w(a, z, s), V^e(a, z, s)\}. \tag{4}$$

The workers' value function is given by

$$\begin{aligned} V^w(a, z, s) = &\max_{c, a' \geqslant 0} u(c) + \rho\{\psi V(a', z, s) + (1 - \psi) \int_{z'} V(a', z', s) d\Upsilon(z'|s)\} \\ &\text{s.t.} \quad c + a' \leqslant w + (1 + r)a. \end{aligned} \tag{5}$$

The entrepreneurs' value function is given by

$$\begin{aligned} V^e(a, z, s) = &\max_{c, k, l, a' \geqslant 0} u(c) + \rho\{\psi V(a', z, s) + (1 - \psi) \int_{z'} V(a', z', s) d\Upsilon(z'|s)\} \\ &\text{s.t.} \quad c + a' \leqslant z\,(k^{\alpha} l^{\beta})^{1-\nu} - wl - (r + \delta + d)k - \kappa_s + (1 + r)a \\ &(1 + r + \delta + d)k \leqslant \phi\, \lambda_s\, a. \end{aligned} \tag{6}$$

[20] Collateral-based financial constraints are more common in India, as the majority of loans are based on collateral and not on cash flows. For example, in 2014, the majority of the loans required collateral in India according to the World Enterprise Survey 2014.

[21] Similar to Buera and Shin (2013), any borrowing for intertemporal consumption smoothing is ruled out by assuming $a_t \geqslant 0$. This constraint is binding for workers, whereas it does not matter for entrepreneurs as they need to have a sufficiently large asset base to fund their capital requirements.

5.5. Recursive competitive equilibrium

Equilibrium: At time 0, given the distribution $\Lambda_0(a,z,s)$, the equilibrium of the economy is characterized by a sequence of allocations $\{o_t, c_t, a_{t+1}, k_t, l_t, K_c, N_c\}_{t=0}^{\infty}$, factor prices $\{w_t, r_t\}_{t=0}^{\infty}$, and $\Lambda_t(a,z,s)_{t=1}^{\infty}$ such that

1. $\{o_t, c_t, a_{t+1}, k_t, l_t\}_{t=0}^{\infty}$ solves the individuals' policy functions for given factor prices $\{w_t, r_t\}_{t=0}^{\infty}$.
2. The capital, labor, and goods markets clear in each period

$$\int_{o_t(a,z,s)=e} k_t\, d\Lambda_t(a,z,s) + K_c - \int a_t\, d\Lambda_t(a,z,s) = 0,$$

$$\int_{o_t(a,z,s)=e} l_t\, d\Lambda_t(a,z,s) + N_c - \int_{o_t(a,z,s)=w} d\Lambda_t(a,z,s) = 0,$$

$$\int_{o_t(a,z,s)=e} [z_t(k_t^{\alpha} l_t^{\beta})^{1-\nu} - \kappa_s] d\Lambda_t(a,z,s) + K_c^{\theta_c} N_c^{1-\theta_c} = \int c_t d\Lambda_t(a,z,s) + (\delta + d)K + \delta K_c;$$

3. The joint distribution of productivity and assets for each caste $\Lambda_t(a,z,s)_{t=1}^{\infty}$ evolve according to the equilibrium mapping

$$\Lambda_{t+1}(a,z,s) = \psi \int_{\{z, a_{t+1}(a,z,s)<a\}} \Lambda_t(da,dz,s) + (1-\psi) \int_{\{z' \leqslant z, a'(a,z,s) \leqslant a\}} \Lambda_t(da,dz,s) d\Upsilon_t(z'|s).$$

6. Quantitative analysis

In this section, the role of access to credit is evaluated in explaining the cross-caste dispersion in the *arpk*. Further, the output losses due to resource misallocation generated by such asymmetries at the extensive and intensive margins are quantified. The section starts by calibrating the model to India's manufacturing sector.[22] Then, the model is assessed on several nontargeted moments to validate the calibrated parameters.

6.1. Calibration

The calibrated parameters are provided in Table 5. The calibration strategy is based on Buera and Shin (2013) and Quadrini (2000). Overall, one needs to specify values for 19 parameters: span-of-control of production technologies, dispersion and persistence in ability distributions, degree of financial frictions for each caste, fixed cost of operation for each caste, discount factor, coefficient of risk aversion, capital depreciation rate, physical capital share in the corporate and non-corporate sectors, size of the corporate sector, intermediation cost, and population shares. The parameterization proceeds in two steps. First, a set of parameters that are fixed outside the model are chosen.[23] Second, given the values of these fixed parameters, the remaining parameters are calibrated to match the salient features of the economy.

Fixed Parameters: A model period is one year. The annual depreciation rate for capital is $\delta = 0.06$, the capital income share in both sectors is $\alpha = \theta_c = 0.33$, and the coefficient of risk aversion is $\gamma = 1.5$, following Hsieh and Klenow (2009) and Cagetti and De Nardi (2006). The share of corporate sector capital in total capital is $\kappa_c = 0.68$. The population share for high castes is 31.2%, whereas for low and middle castes it is 29.5% and 39.3%, respectively, as computed in Indian Human Development Survey (IHDS) data.[24] Finally, the intermediation cost is fixed at $d = 0.033$ to match the net interest rate margin, as in Gu (2021).

Fitted Parameters: The remaining parameters are chosen such that a number of statistics computed using a panel of simulated data are close to their empirical counterparts. The simulated data are drawn from the stationary distribution. The matched moments and their empirical counterparts are shown in Table 5.

Because of the non linearities involved, it is not possible to match particular parameters and moments. However, the mechanics of the model clearly indicate which are the key parameters for each set of moments. What follows is a summary description of the algorithm that is used to select parameter values.

Column 6 of Table 5 shows the relevant moments in the Indian data. Further, Fig. 3 a shows the cumulative employment share distribution, and Fig. 3 b plots the cumulative distribution of income in the model and the data. In particular, the model targets the income share of the top 10%, 5%, and 1% of the population to capture the concentration of income at the top. Similarly, the model targets the top end of the employment share distribution to mimic the concentration of employment in the right tail.

[22] The manufacturing sector is chosen for two reasons. First, because of restrictions on the data side, I can evaluate the proportion of output that is linked to MSMEs in the manufacturing sector but not in the service sector. This is important as this moment is matched in the model to compute comparable statistics. Second; Hsieh and Klenow (2009) also evaluate the role of misallocation in the manufacturing sector in India, thus, focusing on the manufacturing sector helps me to review my results with respect to their findings.

[23] The fixed parameters are difficult to identify with the available data, so I use the values that are commonly used in the literature.

[24] I use the average between 2004 and 2012.

Table 5
Calibrated Parameters and Matched Moments.

Fixed	Value	Description			
δ	0.06	Annual depreciation rate			
α	0.33	Physical capital share: noncorporate sector			
α_c	0.33	Physical capital share: corporate sector			
γ	1.50	Coefficient of risk aversion			
κ_c	0.68	Share of capital in corporate sector			
d	0.03	Net interest rate margin			
P_{lc}	0.29	Population share: LC			
P_{mc}	0.39	Population share: MC			
P_{hc}	0.36	Population share: HC			
Fitted	Value	Description	Moments	Model	Data
λ_{hc}	4.20	Financial frictions HC	Overall Credit/output	0.45	0.44
λ_{mc}	2.05	Financial frictions MC	Rel. Credit/output: MC	0.54	0.53
λ_{lc}	1.86	Financial frictions LC	Rel. Credit/output: LC	0.46	0.48
ν	0.25	Span of control	Income distribution	See Fig. 3 a	
η	4.85	Scale Parameter productivity	Employment distribution	See Fig. 3 b	
			Annual Real Interest Rate Capital-output ratio	5.84%	5.7%
ρ	0.92	Discount rate		2.10	2.04
ψ	0.89	Persistence in productivity	Annual Enterprise Exit rate	10.1%	8.8%
κ_{hc}	0.40	Fixed cost of operating-HC	Share of enterprise owned-HC	37.9%	36.5%
κ_{mc}	0.30	Fixed cost of operating-MC	Share of enterprise owned-MC	46.1%	47.5%
κ_{lc}	0.67	Fixed cost of operating-LC	Share of enterprise owned-LC	16.0%	16.0%
$\bar{K}$	87.1	Capital threshold MSME sector	Share of MSME sector	0.72	0.70

Note. HC, MC, and LC stands for high-caste, middle-caste and low-caste, respectively. The net interest rate margin is taken from Gu (2021). The rel. Credit/output for MC is $\left(\frac{Credit}{output}\right)_{MC} \Big/ \left(\frac{Credit}{output}\right)_{HC}$, and the rel. Credit/output for LC is $\left(\frac{Credit}{output}\right)_{LC} \Big/ \left(\frac{Credit}{output}\right)_{HC}$. The households' income distribution in Fig. 3 a is taken from IHDS (see Table in Appendix). The data counterpart of employment share distribution is from the Economic Census. The capital-output ratio is taken from Feenstra et al. (2015). The share of the MSME sector in noncorporate sector is taken from Section 3.

The model targets an annual enterprise exit rate of 8.8%.[25] The overall credit-to-output ratio is 0.44 (domestic credit to private sector as percentage of GDP, see World Bank, 2021). The credit-to-output ratio for LC and MC entrepreneurs relative to HC entrepreneurs is 0.53 and 0.48, respectively.

For some empirical moments such as the relative credit-to-output ratio for different castes as listed in Table 5, MSMEs are used. Therefore, one needs to define MSME enterprises in the model as well. The MSME threshold $\bar{K}$ is chosen such that the total output produced by enterprises with a capital stock below $\bar{K}$ is the same as in the data (see Section 3 for a discussion on the MSME capital threshold).

Column 5 of Table 5 shows the moments simulated from the calibrated model. All moments in the model are jointly determined by the fitted parameters. To a certain degree, however, they tend to be differentially linked to the parameters. For instance, the discount factor is set at $\rho = 0.92$ to match the annual interest rate of 5.7% and the capital-output ratio of 2.04.[26] The lower discount rate makes households less patient, which in turn decreases savings in the economy and raises the interest rate.

I am thus left with nine non standard parameters: $1-\nu, \eta, \psi, \lambda_{lc}, \lambda_{mc}, \lambda_{hc}, \kappa_{hc}, \kappa_{mc}, \kappa_{lc}$. Closely following Buera and Shin (2013), given the returns to scale, $1-\nu$, the tail parameter of the entrepreneurial talent distribution, η, is chosen to match the right tail of the employment share distribution (or concentration at the top). This is done because, under a perfect credit benchmark ($\lambda \to \infty$), the employment distribution follows a Pareto distribution with a tail parameter proportional to the tail parameter of the entrepreneurial talent distribution.[27] One can then infer $1-\nu$ from the upper tail of the income distribution. The households at the top end of the distribution tend to be entrepreneurs in both the data and the model, and $1-\nu$ controls the share of output that goes to entrepreneurial input (see Fig. 3 a and Fig. 3 b).[28] The persistence of the productivity process ψ is calibrated to match the annual exit rate of 8.8% in the model. The entrepreneurs in the model exit only if their newly drawn productivity is below the equilibrium cutoff level. A lower persistence implies higher churning of entrepreneurs in the model. The parameters related to financial frictions $\{\lambda_{lc}, \lambda_{mc}, \lambda_{hc}\}$ are fixed such that the overall credit-to-output ratio and the credit-to-output ratios for LC and MC entrepreneurs relative to HC entrepreneurs are also matched. A higher λ_c implies a larger supply of credit in the economy and hence higher leverage in the economy. The fixed cost of operation κ_s is set to match the share of enterprises for each caste.

[25] Because of data access difficulties, I use the average exit rate for US manufacturing firms. The exit rates of firms in India and the US are not very different, as mentioned in Hsieh and Klenow (2014) on page 1043, paragraph 3. They compute the exit rate between 1992 to 1997 from the US Manufacturing Census and the exit rate between 1994–1995 to 2010–2011 from the Indian ASI-NSS.

[26] The annual real interest rate in India varied from 2% to 8.34% between 1999 and 2010. The capital-output ratio is taken from Feenstra et al. (2015).

[27] See Appendix for a derivation of the employment distribution.

[28] The within-caste distribution of income in the model and the data is presented in Fig. in Appendix.

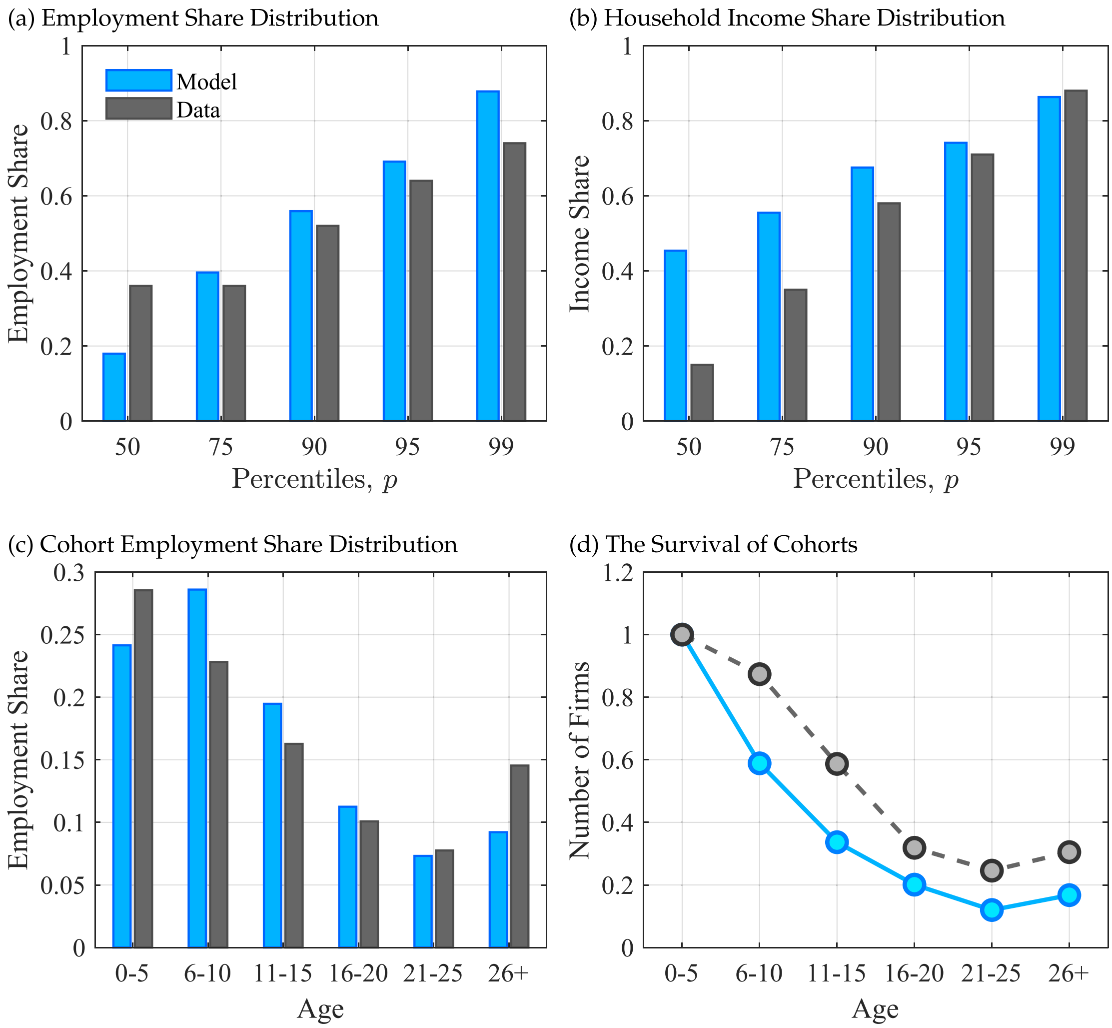

Fig. 3. Distribution of Enterprises, Employment and Income Note. Grey (dark) color represents data points and Blue (light) color represents model counterparts. Panel (a) documents the share of employment by enterprises below a certain percentile of the enterprise size (employees) distribution. Panel (b) shows the share of income owned by households below a certain percentile of the household income distribution. Panel (c) documents the employment share of enterprises in a certain age bin and Panel (d) depicts the share of enterprises in a certain age bin relative to the share of enterprises in the youngest age category (definition of age bins and relevant data points are taken from Akcigit, Alp and Peters 2021). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The calibration finds returns to scale equal to $1-\nu = 0.75$ and a scale parameter for productivity distribution $\eta = 4.85$. Further, the model identifies $\lambda_{lc} = 1.86$ and $\lambda_{mc} = 2.05$, which are significantly smaller than $\lambda_{hc} = 4.20$. This result is driven by the lower credit-to-output ratio of LC and MC entrepreneurs relative to that of HC entrepreneurs. The fixed cost of operation is set at $\kappa_{hc} = 0.40$, $\kappa_{mc} = 0.30$, and $\kappa_{lc} = 0.67$.

6.2. Wealth, consumption, and entrepreneurship

This section discusses the implications of financial frictions on individual decision making and how they affect castes differently.

The model imposes $a \geq 0$. This constraint is binding only for individuals who choose to be workers; it has no direct effect on the behavior of entrepreneurs of any castes, who need to hold assets for production because of the collateral constraint. A relatively higher degree of financial frictions for low castes distorts their occupational choice and makes them more likely to be workers than high castes. Therefore, average asset holdings are lower for LC households relative to HC households.

Further, Fig. 2 highlights the differences in various policy functions of individuals and discuss how they originate in the model. First, a comparison of consumption profiles across individuals over the state space of assets is presented. As expected, wealthier agents consume relatively more. However, when I compare high productivity individuals across castes, low castes consume less than high castes, whereas no such difference appears among low productivity individuals. This

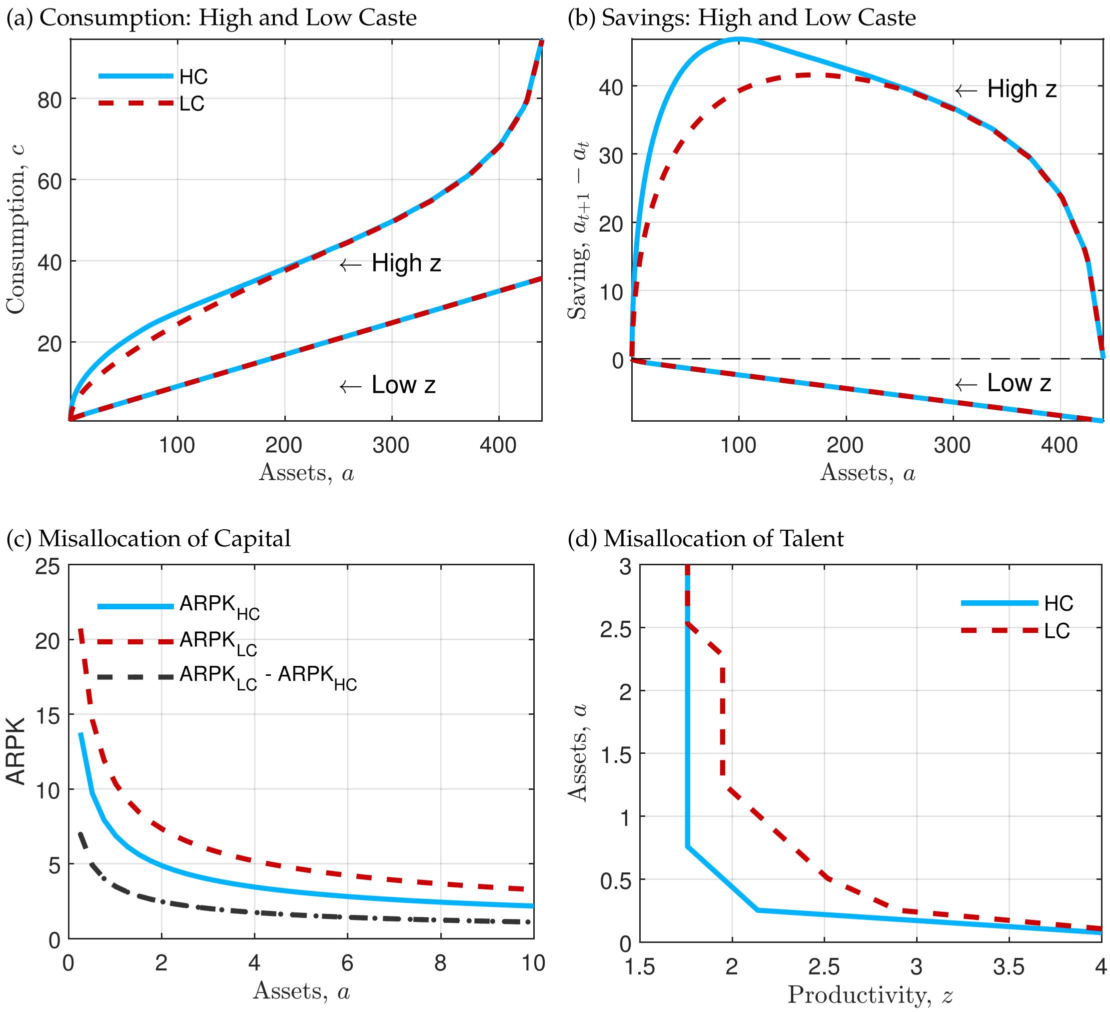

Fig. 2. Consumption, Asset Growth and Misallocation across Castes: Model Note. LC and HC represents low-caste and high-caste households, respectively. In the Panels (a) and (b), z represents productivity. These panels have policy functions of consumption and savings for high productivity z and low productivity z households in the model. In Panels (c) and (d) I truncate the state space of productivity z and assets a to clearly highlight the cross-caste differences. Panel (c) presents the average revenue product of capital *arpk* over assets for entrepreneurs of different castes. Panel (d) presents the occupation choice policy function for every productivity z and assets a. The lines represent pairs (z, a) of a marginal entrant upon whose right everyone becomes an entrepreneur.

result comes from the fact that high productivity individuals choose entrepreneurship, whereas low productivity individuals become workers. The workers are paid the same wage regardless of their caste, but entrepreneurial income depends on the ability of different castes to borrow from the financial markets. Low castes can borrow less relative to high castes, which in turn reduces their entrepreneurial output and income. A similar pattern is observed for asset growth rates across individuals of different castes (see Fig. 2 b). The distributional impact of caste-specific borrowing constraints on income and wealth will be discussed in Section 6.5.2.

6.2.1. Misallocation across castes

The literature has stressed the role of financial frictions on two different margins of misallocation: the extensive margin and the intensive margin.

The *intensive margin* refers to the overall dispersion in *arpk*; however, this paper is primarily concerned with the *arpk* dispersion across castes. The literature refers to the dispersion in the marginal revenue product of capital (*mrpk*) as a misallocation of capital. The model implies that the *mrpk* is directly proportional to the *arpk*, $mrpk = \alpha(1-\nu)arpk$, where $\alpha(1-\nu)$ is constant across castes. Therefore, dispersion in *mrpk* directly translates into dispersion in *arpk*. For constrained entrepreneurs, *mrpk* is not equalized to the marginal cost and thus generates a wedge. This wedge decreases as the asset base of the entrepreneurs increases because it increases their borrowing capacity.

Fig. 2 c plots the *arpk* of high- and low-caste entrepreneurs with the same productivity and their difference over assets in the model. The *arpk* is higher for low castes relative to high castes for every value of assets, and this difference stems

Table 6
Model Assessment on Nontargeted Moments.

	Model	Data		Model	Data			
Panel A.1: ARPK, Capital Intensity and Profitability								
log(ARPK)-MC	+17.7%	+13%(22%)	$sd(log(ARPK))$-HC	0.44	1.36			
log(ARPK)-LC	+31.6%	+25%(30%)	$sd(log(ARPK))$-MC	0.54	1.22			
log(K/L)-MC	−17.7%	−26%(27%)	$sd(log(ARPK))$-LC	0.64	1.22			
log(K/L)-LC	−31.6%	−49%(45%)						
Profitability-MC	+10.4%	+2%(10%)						
Profitability-LC	+6.0%	+5%(14%)						
Panel A.2: ARPK over Firm Size for LC Enterprises								
Small Enterprises: Lowest Decile	+55%	+52%						
Large Enterprises: Top Decile	+15%	+12%						
PANEL B: Regional Financial Development and ARPK Differences								
	Model					Data		
Financial Development	Lowest	Benchmark			Highest	Low	Medium	High
Regions	(i)	(ii)	(iii)	(iv)	(v)	(a)	(b)	(c)
Intermdiation cost	$d = 0$	$d = 0.03$	$d = 0$	$d = 0$	$d = 0$			
Regional credit shifter	$\phi = 0.6$	$\phi = 1$	$\phi = 1$	$\phi = 2$	$\phi = 3$			
Credit-output ratio	0.33	0.45	0.60	0.82	0.90	0.28	0.62	1.21
log(ARPK)-MC	+34.2%	+17.1%	+19.4%	+14.4%	+9.1%	+19.1%(30.8%)	+9.7%(10.1%)	+8.5%(9.7%)
log(ARPK)-LC	+45.0%	+30.1%	+33.1%	+21.4%	+13.6%	+30.3%(45.8%)	+22.8%(14.5%)	+11.5%(9.3%)

Note: HC, MC, and LC stands for high-caste, middle-caste and low-caste, respectively. **Panel A.1**: The measures of ARPK, capital-labor ratio (K/L), and profitability (is defined as $\frac{Y-wL-RK}{Y}$, where Y is value added, wL is the wage bill, and RK is the cost of capital, respectively) are computed for MSMEs in both the data and the model and represent their respective values with respect to HC enterprises. For ARPK, K/L, and profitability, the table also contains, in brackets, the caste differences estimated using the regression model with sampling multiplier. $sd(log(ARPK))$ report the standard deviation for each caste. **Panel A.2** presents the *arpk* for LC enterprise relative to HC enterprise for small enterprises (enterprises below 10th percentile of the size (in terms of output) distribution) and for large enterprises (enterprises above 90th percentile of the size distribution). **Panel B**: The measures of ARPK represent their respective values with respect to HC enterprises for different regions. In the model, financial constraint is given by $(1 + r_t + \delta + d)k_t \leqslant \phi\, \lambda_s\, a_t$, where regional credit supply shifter is ϕ and intermediation cost is d. The *arpk* differences are computed for all enterprises in the model. For data, I divide regions into three levels of financial development: low, medium and high and report the *arpk* differences across castes (in brackets, the estimates from the regression model with sampling multiplier are reported). The aggregate credit-output ratio is computed for these three groups and is reported in row five.

from their lower ability to borrow, $\lambda_{lc} < \lambda_{hc}$. Further, the difference between the *arpk* of low- and high-caste entrepreneurs declines with assets. On average, enterprises with a large asset base are large in size as well; therefore, a decline in the relative *arpk* over assets in the model is in line with Fact 2 in Section 4.

The *extensive margin* refers to the distorted occupation choice. In particular, the occupation is determined by Eq. 4, and this decision is dynamic. The occupational thresholds above which individuals choose to be entrepreneurs are shown in Fig. 2 d. The occupational threshold depends on two state variables; exogenous productivity z and endogenous assets a. In a world without financial frictions, the occupational choice is determined by productivity and fixed cost of operation. Financial frictions affect the marginal entrants, as they deter the entry of poor but marginally productive individuals and promote the entry of wealthy but marginally unproductive individuals. The stricter borrowing constraints for low castes exacerbate these forces and distort their occupational decisions relatively more on the margin. High fixed costs of operation for LC households further deter entry into entrepreneurship.

6.3. Nontargeted moments

This section discusses the life-cycle implications of the model. Fig. 3 compares the distributions produced by the model with a representative empirical distribution constructed using Indian data. The enterprises are divided into different cohorts following Akcigit et al. (2021). Fig. 3 c reports the aggregate employment share by age for Indian manufacturing enterprises. In line with the literature, it is clear that the aggregate importance of old enterprises is small in India. While enterprises that are older than 25 years account for 55% of employment in the US, the corresponding figure is less than 20% in India. Further, I compare the model with respect to the degree of selection. Fig. 3 d depicts the survival rate, i.e. the share of enterprises by age relative to the share of enterprises in the youngest age category. The rate of enterprise survival in the model is similar to that in the data.[29]

6.3.1. ARPK Dispersion across castes

In a final validation exercise, I compare *arpk* averages and their dynamics across castes. The model predicts that low and middle castes have 31.6% and 17.7% higher *arpk* than high castes (see Panel A.1 in Table 6). These values are in line with the empirical estimates, as reported in Column 3 of Panel A.1 in Table 6.

[29] Because of data limitations, the distributions are constructed for all the enterprises in the data instead of the noncorporate sector, which generate bias in the distributions.

Moreover, in the spirit of Fact 2 in Section 4, I compare the *arpk* estimates across small and large enterprises in the model with those in the data. The *arpk* of low-caste enterprises drops from being 55% higher in the bottom decile to 15% higher (see Panel A.2 in Table 6) in the top decile of the firm size (measured as output) distribution relative to high castes. This finding is qualitatively similar to the findings in Fact 2 in Section 4.

In terms of the standard deviation $sd(arpk)$, the model underpredicts the value (less than half of the observed value) for each caste relative to the data (see Panel A.1 in Table 6). This could be because other forces – such as imperfect competition, heterogeneous input shares, measurement error, among others – affect $sd(arpk)$ but are not present in the model.[30] In this way, one can interpret the $sd(arpk)$ predicted by the model as a fraction that can be explained by the financial frictions (together with perfect competition) in the context of the Indian manufacturing sector.[31]

Further, the presence of heterogeneous borrowing constraints across castes also distorts their capital-labor ratio. Low castes experience a high shadow cost of capital that lowers their capital-labor ratio, as documented in Fact 3 in Section 4. Here, I compare those estimates with that of the model. The model predicts a 31.6% and 17.7% lower capital-labor ratio for low- and middle-caste enterprises relative to high castes (see Panel A.1 in Table 6).

Finally, the model predicts that LC entrepreneurs are 6.0% and MC entrepreneurs are 10.4% more profitable than HC entrepreneurs, respectively, a finding that is in line with the data (see Panel A.1 in Table 6). Profitability is determined by a combination of forces. Stricter financial constraints increase the selection effect; however, they also limit the enterprise size and decrease profits. In the quantitative exercise, the selection effect dominates.

6.4. Quantitative regional analysis

This section revisits Fact 3 in the spirit of the model. In the data, states that are more financially developed (i.e., those with a high credit-to-output ratio) have lower cross-caste dispersion in *arpk*. To assess the model's ability to capture this correlation, I vary the intermediation cost d and the common component of the financial constraints ϕ and report steady-state statistics.

We can view different values of credit supply shifter ϕ or intermediation cost d as representing different regions in India.[32] For instance, high ϕ or low d implies a region with developed financial markets. Panel B in Table 6 reports the results. Region (ii) represents the benchmark economy. In region (iii), where intermediation cost is reduced to $d = 0$, an increase in the credit-output ratio is observed. However, a slight increase in the *arpk* difference across castes relative to the region (ii) is found as well. The *arpk* in absolute terms declines for all castes as the cost of capital is lower, but *arpk* declines relatively more for HC entrepreneurs.

Further, in regions (iv) and (v), value of credit supply shifter ϕ is increased. The values of ϕ used in the simulation generate variation in the credit-output ratio that is consistent with the data.[33] I find that it substantially increases the credit-output ratio and decreases cross-caste *arpk* differences. This finding suggests a decline in the misallocation of resources, a result that is in line with Buera et al. (2011). These results are driven by the fact that large values of credit supply shifter ϕ push the majority of entrepreneurs out of credit constraints such that the credit supply shifter lowers the negative impact of caste differences in financial frictions λ_c. Therefore, it disproportionately benefits non-HC enterprises as they were relatively more credit constrained in the benchmark economy.

Finally, I compare these model predictions with data and find consistent results. In particular, Panel B in Table 6 reports the cross-caste *arpk* differences for regions with low (a), medium (b), and high (c) levels of financial development. Relative to HC entrepreneurs, the *arpk* of LC entrepreneurs is 30%-45% higher in less financially developed regions, whereas it is only 9.3%-11.5% higher in regions with high levels of financial development.

In the model, the differential impact of changes in ϕ and d on cross-caste *arpk* differences may be driven by the fact that a decline in d directly increases entrepreneurs' borrowing capacity and decreases the cost of capital – consequently increasing the desired level of capital, whereas an increase in ϕ increases entrepreneurs' borrowing capacity only in partial equilibrium. Therefore, the former is more potent in decreasing the gap between the actual and desired levels of capital.

6.5. Counterfactual analysis

This section highlights the potential productivity and output gains if capital were homogeneously allocated across castes. In counterfactual economy CF, the degree of financial frictions for non-HC individuals is similar to those of HC individuals (i.e., $\lambda_{lc} = \lambda_{mc} = \lambda_{hc}$). The model predicts output per capita gains of 5.6% (see Panel A in Table 7).

These gains come from two main sources in the counterfactual economy. First, the reallocation of capital from unproductive HC entrepreneurs to more productive non-HC entrepreneurs increases the allocative efficiency. As a result, the dispersion in *arpk* declines by 24.1% (see Panel B in Table 7). This decline in the dispersion in *arpk* is driven by the decline in

[30] See, e.g., Atkin et al. (2019), Rotemberg and White (2021), and Bils et al. (2021).

[31] Further, the model predicts higher $sd(arpk)$ for LC entrepreneurs relative to HC entrepreneurs, whereas it is opposite in the data. This may be because of the differences in the underlying productivity process for each caste (e.g., Asker et al. (2014) show that the underlying productivity process is an important driver of $sd(arpk)$). In this paper, the productivity process is assumed to be the same across castes.

[32] The approach of viewing different regions as different levels of credit supply shifter ϕ is more closely related to the exercise performed in Buera et al. (2011). Each region is considered as a separate general equilibrium economy.

[33] The regional credit-output ratio varies between 0.10 to 1.91, whereas the ratio for the majority of regions is below 1.

Table 7
A Comparison of the Benchmark and Counterfactual Economy.

	BM	CF	Change			
Panel A. Overall Economy						
Output per capita	1.43	1.51	+5.6%			
Capital Intensity	3.14	3.29	+4.7%			
Interest Rate	5.86%	6.35%	0.5 p.p.			
Panel B. Noncorporate Sector						
Output per worker	1.76	1.86	+5.7%			
$\sigma(arpk)$	0.54	0.41	-24.1%			
Capital Intensity	2.41	2.63	+9.1%			
Credit/Output	0.46	0.60	+30.4%			
Panel C. Caste-Level						
	Middle-Caste			Low-Caste		
	BM	CF	Change	BM	CF	Change
Capital Intensity	2.34	2.64	12.8%	2.28	2.64	15.8%
Output per worker	1.76	1.86	+5.7%	1.76	1.86	+5.7%
Enterprise Ownership	46%	46%	+0 p.p.	16%	22%	+6.0 p.p.
Credit/Output	0.35	0.60	71%	0.29	0.58	100%
$\sigma(arpk)$	0.54	0.41	-24.0%	0.64	0.45	-29.7%

Note. This table compares the results between the benchmark economy (BM) and the counterfactual economy (CF) for the overall economy and the noncorporate sector. Relative to the BM economy, the degree of financial constraints is equalized across castes in the CF economy. **Panel A** presents the output per capita, capital intensity and interest rate in the overall economy. **Panel B** reports the output per worker, $\sigma(arpk)$ (standard deviation), capital intensity and credit-output ratio for the noncorporate sector. **Panel C** presents capital intensity (aggregate), output per worker (aggregate), enterprise ownership, credit-output ratio (aggregate) and $\sigma(arpk)$ for low-caste and middle-caste enterprises. The changes in enterprise ownership rates and interest rates are in percentage points (p.p.).

both within- and across-caste groups. For instance, the relative difference in average $arpk$ for low- and high-caste enterprises falls from +31.6% in the benchmark economy to +4% in the counterfactual economy (see Panel C in Table 7). As a result, the economy becomes more capital intensive, with gains of +9.1% in the capital-labor ratio in the noncorporate sector. These gains come mostly from low- and middle-caste entrepreneurs, which increases their aggregate capital-labor ratio by 12.7% and 15.8%, respectively.

Second, the reduction in borrowing constraints induces the entry of more non-HC entrepreneurs at the margin (occupational choice, as shown in Fig. 2 d, becomes symmetrical across castes). The share of LC enterprises increases from 16% in the benchmark economy to 22%. Moreover, because of the excess entry of entrepreneurs, demand for capital and labor increases. The interest rate increases by 0.5 p.p. and wages increase by 6.1%. This implies a higher cost of inputs of production in CF, which leads to the exit of unproductive HC enterprises.

Finally, two more counterfactual exercises are performed to highlight the importance of misallocation at the extensive and intensive margins and to disentangle the gains from these two sources. I start with the stationary distribution $\Lambda(z, a, c)$ and redistribute capital across entrepreneurs (equalizing λ_c), while the distribution of enterprises, total capital, and labor supply are the same. The reallocation of capital toward non-HC entrepreneurs increases output per capita by 3.1%. Next, I allow the distribution $\Lambda(z, a, c)$ to adjust to changes in factor prices. This allows productive but poor non-HC entrepreneurs to enter the market and forces unproductive but wealthy entrepreneurs to exit the market. The reduction in talent misallocation increases output per capita by 2.5%.

6.5.1. *Enterprise ownership rates and population shares across castes*

In the data, low castes are 29.5%, middle castes are 39.3%, and high castes are 31.2% of the total population, whereas LC owns 16%, MC owns 47.5%, and HC owns 36.5% of total enterprises (see Table 1). In the counterfactual economy, where the degree of financial constraints is equalized, LC owns 22%, MC owns 46%, and HC owns 32% of total enterprises (see Panel C in Table 7). The asymmetries in access to credit reduce the share of enterprises owned by LC households by 6 p.p.

6.5.2. *Wealth and income inequality across castes*

Hnatkovska et al. (2012) document that wage earnings for median (mean) LC households are 21% (29%) lower than in non-LC households. Using the All India Debt and Investment Survey, I report vast disparities in wealth holding across castes (see Table).[34] Further, I document that wealth inequality is higher among HC households relative to non-HC households (see Table in Appendix). Using IHDS data, I report that, on average, LC household income is half that of HC household income and income inequality is higher among HC households relative to non-HC households (see Table and Table in Appendices and).

In the benchmark calibration, for LC households on average income is 9.0% and wealth is 26.0% lower, relative to HC households. The model produces results that are qualitatively similar in these two aspects; however, it substantially under-

[34] See Bharti (2018) for more discussion on wealth inequality.

estimates cross-caste inequality (see Table in Appendix). This means that caste-specific borrowing constraints do make low castes poorer on average, but are not enough to match the cross-caste disparities in income and wealth as documented in the data.[35]

Furthermore, within caste distribution of households' income between the model and the data is compared in Fig.. Given that wages are the same across households in the model, the differences in the income distribution are driven by differences in profits and interest income on assets. The income inequality is lower among LC households as 25.3% of their income is owned by the top 5% of households, whereas 27.6% of the total HC income is earned by the top 5% of HC households. Higher wealth inequality is found among HC households relative to LC households. Among the HC entrepreneurs, 47.7% of their wealth is owned by the richest 5%, whereas, among the LC entrepreneurs, 42.4% of their wealth is owned by the richest 5% (see Table in Appendix).

In the counterfactual experiment, cross-caste wealth inequality declines as LC households become richer on average relative to the benchmark economy (wealth increases by 16.4% for LC households on average). Moreover, when LC enterprises are allowed to borrow more, wealth inequality within LC households increases, as the upper tail of the distribution is able to create more wealth (see Table in Appendix).

Cross-caste income inequality decreases as the average LC household is richer in the counterfactual economy (income increases by 8.8% for LC on average). Furthermore, the income inequality increases within LC households (in Appendix).

6.6. Comparison of results with existing literature

Here, I focus on the quantitative results and discuss how they relate to the existing literature. Hsieh and Klenow (2009) argue that if capital and labor were efficiently allocated in India, then output would increase significantly. Through the lens of the model, I conclude that caste-specific asymmetries in the financial markets in India are important, and the removal of such obstacles could increase output per capita by 5.6%. Further, in the case of financial frictions, Moll (2014) shows that the losses due to misallocation of capital depend on the persistence in the productivity process, ψ, and even small changes in ψ can lead to large differences in capital misallocation in the steady state. The estimates of persistence in this paper are in line with those found in Buera et al. (2011). Moreover, the estimated fixed cost of operation is much lower in my calibration than what is estimated in Buera et al. (2011). This factor may explain the bigger role for misallocation of capital in explaining productivity losses in this context.[36]

7. Conclusion

It is well established that the misallocation of resources can explain a large portion of the productivity differences across countries. However, the sources of misallocation are still under investigation, and several firm-level distortions have been proposed. This paper studies whether the caste system has a distortionary effect on capital and talent allocation in the economy, and quantify its contribution to aggregate output losses. It documents that enterprises owned by historically disadvantaged castes exhibit a higher *arpk*, higher profitability, and a lower capital-labor ratio relative to high-caste owned enterprises in India. A quantitative model of entrepreneurship is used to understand the implications of cross-caste dispersion in *arpk* for overall output per capita. The model identifies a high degree of financial constraints for non-HC entrepreneurs and shows that such asymmetries reduce output per capita by 5.6%. Given the findings of this paper, a natural next step would be to understand the implications of the caste system on long-run growth.

Data Availability

Data will be made available on request.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jmoneco.2023.01.007.

[35] To a certain extent, low cross-caste income inequality in the model is driven by equal wages for workers of all castes. If heterogeneous wages are introduced in the model for each caste (particularly, relatively low wages for non-HC workers), then one would expect higher enterprise ownership for non-HC households as the opportunity cost is lower relative to HC households. This would imply that even higher fixed costs are required to match the enterprise ownership across castes relative to the benchmark economy. High fixed costs would increase both income and wealth inequality. Understanding the different determinants of cross-caste income and wealth inequality is left for future research.

[36] The sources of misallocation losses depend on the fixed cost of operation as mentioned in Buera et al. (2011), where sectors characterized by a low fixed cost tend to have major productivity losses due to misallocation of capital. Buera et al. (2011) find the per-period fixed cost of operation in manufacturing to be about three times the equilibrium wage in the perfect-credit benchmark, whereas this paper finds the per-period fixed costs of operation to be approximately half of the equilibrium wages.

References

Akcigit, U., Alp, H., Peters, M., 2021. Lack of selection and limits to delegation: firm dynamics in developing countries. Am. Econ. Rev. 111 (1), 231–275.
Akerlof, G., 1976. The economics of caste and of the rat race and other woeful tales. Q. J. Econ. 599–617.
Arrow, K.J., 1973. The theory of discrimination. Princeton University Press.
Asker, J., Collard-Wexler, A., De Loecker, J., 2014. Dynamic inputs and resource (mis) allocation. J. Polit. Econ. 122 (5), 1013–1063.
Atkin, D., Khandelwal, A.K., Osman, A., 2019. Measuring productivity: Lessons from tailored surveys and productivity benchmarking. In: AEA Papers and Proceedings, Vol. 109, pp. 444–449.
Ayyagari, M., Demirguc-Kunt, A., Maksimovic, V., 2014. Does local financial development matter for firm lifecycle in India? The World Bank.
Banerjee, A., Munshi, K., 2004. How efficiently is capital allocated? evidence from the knitted garment industry in Tirupur. Rev. Econ. Stud. 71 (1), 19–42.
Banerjee, A.V., Duflo, E., 2005. Growth theory through the lens of development economics. Handbook Econ. Growth 1, 473–552.
Banerjee, A.V., Moll, B., 2010. Why does misallocation persist? Am. Econ. J.: Macroecon. 2 (1), 189–206.
Becker, G.S., 1957. The economics of discrimination. University of Chicago press.
Bharti, N.K., 2018. Wealth inequality, class and caste in India, 1961–2012. WID. Working Paper 14.
Bils, M., Klenow, P.J., Ruane, C., 2021. Misallocation or mismeasurement? J. Monet. Econ. 124, S39–S56.
Buera, F.J., Kaboski, J.P., Shin, Y., 2011. Finance and development: a tale of two sectors. Am. Econ. Rev. 101 (5), 1964–2002.
Buera, F.J., Shin, Y., 2013. Financial frictions and the persistence of history: aquantitative exploration. J. Polit. Econ. 121 (2), 221–272. http://www.jstor.org/stable/10.1086/670271
Cagetti, M., De Nardi, M., 2006. Entrepreneurship, frictions, and wealth. J. Polit. Econ. 114 (5), 835–870.
Caselli, F., Feyrer, J., 2007. The marginal product of capital. Q. J. Econ. 122 (2), 535–568.
Cassan, G., Keniston, D., Kleineberg, T., 2021. A division of laborers: Identity and efficiency in India. Technical Report. National Bureau of Economic Research.
Damodaran, H., 2008. India'S new capitalists: Caste, business, and industry in a modern nation. Springer.
De Mel, S., McKenzie, D., Woodruff, C., 2008. Returns to capital in microenterprises: evidence from a field experiment. Q. J. Econ. 123 (4), 1329–1372.
Desai, S., Vanneman, R., National Council of Applied Economic Research, N. D., 2018. India Human Development Survey (IHDS), 2005. Inter-university Consortium for Political and Social Research [distributor], 2018-08-08. https://doi.org/10.3886/ICPSR22626.v12.
Deshpande, A., Sharma, S., 2013. Entrepreneurship or survival? Caste and gender of small business in India. Technical Report. Centre for Development Economics Department of Economics, Delhi School of Economics.
Erosa, A., Fuster, L., Kambourov, G., Rogerson, R., 2017. Hours, occupations, and gender differences in labor market outcomes. NBER Working Paper (w23636).
Feenstra, R.C., Inklaar, R., Timmer, M.P., 2015. The next generation of the penn world table. Am. Econ. Rev. 105 (10), 3150–3182.
Fisman, R., Paravisini, D., Vig, V., 2017. Cultural proximity and loan outcomes. Am. Econ. Rev. 107 (2), 457–492.
Garcia-Santana, M., Pijoan-Mas, J., 2014. The reservation laws in India and the misallocation of production factors. J. Monet. Econ. 66, 193–209.
Gopinath, G., Kalemli-Özcan, c., Karabarbounis, L., Villegas-Sanchez, C., 2017. Capital allocation and productivity in South Europe. Q. J. Econ. 132 (4), 1915–1967.
Gu, J., 2021. Financial intermediation and occupational choice. J. Econ. Dyn. Control 133, 104238.
Guner, N., Ventura, G., Xu, Y., 2008. Macroeconomic implications of size-dependent policies. Rev. Econ. Dyn. 11 (4), 721–744.
Hadlock, C.J., Pierce, J.R., 2010. New evidence on measuring financial constraints: moving beyond the KZ index. Rev. Financ. Stud. 23 (5), 1909–1940.
Hall, R.E., Jones, C.I., 1999. Why do some countries produce so much more output per worker than others? Q. J. Econ. 114 (1), 83–116.
Hjort, J., 2014. Ethnic divisions and production in firms. Q. J. Econ. 129 (4), 1899–1946.
Hnatkovska, V., Lahiri, A., Paul, S., 2012. Castes and labor mobility. Am. Econ. J.: Appl. Econ. 4 (2), 274–307.
Hsieh, C.-T., Hurst, E., Jones, C.I., Klenow, P.J., 2019. The allocation of talent and US economic growth. Econometrica 87 (5), 1439–1474.
Hsieh, C.-T., Klenow, P.J., 2009. Misallocation and manufacturing TFP in China and India. Q. J. Econ. CXXIV (November).
Hsieh, C.-T., Klenow, P.J., 2014. The life cycle of plants in India and Mexico. Q. J. Econ. 129 (3), 1035–1084.
Iyer, L., Khanna, T., Varshney, A., 2013. Caste and entrepreneurship in India. Econ. Polit. Wkly 48 (6), 52–60.
Jodhka, S.S., 2010. Dalits in business: self-employed scheduled castes in north-west India. Econ. Polit. Wkly 41–48.
Lucas Jr, R.E., 1978. On the size distribution of business firms. Bell J. Econ. 508–523.
Midrigan, V., Xu, D.Y., 2014. Finance and misallocation: evidence from plant-level data. Am. Econ. Rev. 104 (2), 422–458.
Moll, B., 2014. Productivity losses from financial frictions: can self-financing undo capital misallocation? Am. Econ. Rev. 104 (10), 3186–3221.
Morazzoni, M., Sy, A., 2022. Female entrepreneurship, financial frictions and capital misallocation in the US. J. Monet. Econ..
MSME, 2009. Fourth all India census of micro, small and medium enterprises, 2006. Ministry of Micro, Small and Medium Enterprises. https://www.dcmsme.gov.in/ito_msme/censuses.htm
Munshi, K., 2016. Caste Networks in the Modern Indian Economy. In: Development in India. Springer, pp. 13–37.
Oh, S., 2019. Does identity affect labor supply. Job Marker Paper, Columbia University.
Phelps, E.S., 1972. The statistical theory of discrimination. Am. Econ. Rev. 62 (4), 659–661.
Quadrini, V., 2000. Entrepreneurship, saving, and social mobility. Rev. Econ. Dyn. 3 (1), 1–40.
Restuccia, D., Rogerson, R., 2008. Policy distortions and aggregate productivity with heterogeneous establishments. Rev. Econ. Dyn. 11 (4), 707–720.
Rotemberg, M., White, T.K., 2021. Plant-to-table(s and figures): Processed manufacturing data and measured misallocation. In: Mimeo.
Thorat, S., Sadana, N., 2009. Caste and ownership of private enterprises. Econ. Polit. Wkly 13–16.
World Bank, T., 2021. Domestic credit to private sector by banks (percent of GDP). Indicator Code: FD.AST.PRVT.GD.ZS. Source: International Monetary Fund, International Financial Statistics and data files, and World Bank and OECD GDP estimates.