American Economic Journal: Applied Economics 3 (January 2011): 239–263
http://www.aeaweb.org/articles.php?doi=10.1257/app.3.1.239

Caste as an Impediment to Trade[†]

By Siwan Anderson*

We compare outcomes across two types of villages in rural India. Villages vary by which caste is dominant (owns the majority of land): either a low or high caste. The key finding is that income is substantially higher for low-caste households residing in villages dominated by a low caste. This seems to be due to a trade breakdown in irrigation water across caste groups. All else equal, lower caste water buyers have agricultural yields which are 45 percent higher if they reside in a village where water sellers are of the same caste compared to one where they are not. (*JEL* O12, O13, O17, O18, Q15, R23, Z13)

This paper compares outcomes across two types of villages in a poor region of rural India. We exploit dramatic village level variation in caste composition and land ownership, that is historically and exogenously determined, in order to identify effects of village level caste differentiation on household level outcomes. Villages vary markedly by the identity of their *dominant* caste group.[1] The notion of a dominant caste used here borrows from previous sociological and anthropological work. M. N. Srinivas (1955) first defined the term "dominant caste" to refer to the caste in the village that is numerically strong and also wields preponderant economic and political power. Dumont (1970) later insisted that dominance arises solely from economic power rather than factors like numerical preponderance, and that this power flows exclusively from control of land. This latter definition of caste dominance is the one now commonly used in the literature and is the one used here. Dominant caste refers to the caste group owning the majority of land.

Approximately 48 percent of the Hindu villages in the sample are dominated by an upper caste and 42 percent are dominated by a lower backward agricultural caste (BAC). The differences across village type are dramatic. In the BAC dominated villages there are almost never upper caste households present. Upper caste dominated

* Department of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, BC, V6T 1Z1, Canada, and Canadian Institute for Advanced Research (CIFAR) (e-mail: siwander@interchange.ubc.ca). I thank Patrick Francois, Ashok Kotwal, Thomas Fujiwara, Debraj Ray, Abhijit Banerjee, Roger Myerson, Rohini Somanathan, and two anonymous referees for very useful comments. This paper has also benefited from seminar participants of the LSE-UCL development seminar, University of Michigan, Queen's University, Yale University, New York University, Rice University, Centre interuniversitaire sur le risque, les politiques économiques et l'emploi (CIRPEE) (Quebec), Groupement de Recherche en Economie Quantitative d'Aix Marseille (GREQAM Marseilles), MacArthur Development and Inequality Meeting (Namur), and the CIFAR Institutions, Organizations, and Growth group. Financial help from Social Sciences and Humanities Research Council of Canada (SSHRC) and CIFAR is gratefully acknowledged.

[†] To comment on this article in the online discussion forum, or to view additional materials, visit the article page at http://www.aeaweb.org/articles.php?doi=10.1257/app.3.1.239.

[1] All Hindus, the major religious group in India, are divided into a number of hereditary caste groups. Long-standing rules govern interaction within and across caste groups. These include strict endogamy and restrictions on the sharing of food and drinking water and other social interactions (Louis Dumont 1970).

villages, in contrast, include all main caste groups in the sample: upper castes, BAC, other backward castes (OBC), and scheduled castes (SC).[2] The analysis compares outcomes of lower caste (BAC, OBC, SC) households residing in both types of villages and finds substantially higher income for low-caste households residing in villages dominated by BACs. The aim of this paper is to understand why. The answer's implications are potentially more far-reaching than the Indian context.

It is not that surprising to find some measure of social fragmentation significantly impacting individual well-being. The variation in caste dominance here could be picking up a type of ethnic heterogeneity; the villages where the upper castes are also present (the upper caste dominated villages) are more ethnically heterogeneous than those where only lower castes reside (the BAC dominated villages). Much previous work has demonstrated a negative correlation between ethnic diversity and economic outcomes, consistent with the findings here (refer to Alberto Alesina and Eliana La Ferrara 2005). It is thought that more ethnically diverse communities have greater difficulty sharing public goods and resources, and are less able to impose social sanctions that prevent collective action failures. Previous empirical work, particular to India, has demonstrated that ethno-linguistic fragmentation, applied to caste and religious divisions, negatively correlates with access to public goods (Abhijit Banerjee and Rohini Somanathan 2007; Banerjee, Lakshmi Iyer, and Somanathan 2005).

The hierarchical Indian social structure could also explain the finding that lower caste incomes are higher in low-caste dominated villages. The traditional village economy revolved around a hereditary caste hierarchy that prescribed individuals' occupations. Upper castes were the land owners, middle ranked (backward) castes the farmers and artisans, and the lowest ranked (scheduled) castes were the labourers and performers of menial tasks (André Béteille 1996). Given these historical patterns, we may well expect lower castes to fair better in villages where no upper castes are present. Via tenancy or credit relations, upper caste landlords might be able to exploit the lower castes, so that they are better off in those villages where land is exclusively lower caste.

However, neither low public good access or exploitative tenancy or credit relations explain the large losses visited upon lower castes residing in high-caste dominated villages. Moreover, low-caste losses do not appear directly related to the political economy environment.

The main cause of poorer low-caste outcomes in high-caste dominated villages appears to be a pervasive breakdown in the functioning of private groundwater markets. These markets are ubiquitous and highly important in arid areas, but the empirical results suggest that upper caste water sellers are unable to easily trade with lower caste water buyers. As a result, in villages where the dominant caste, who own the majority of the private groundwater extraction mechanisms, is an upper caste, there appears to be a severe inefficiency in the distribution of groundwater. The implications of this trade breakdown in a poverty stricken part of India are dramatic.

[2] The BAC and OBC categories are broadly both from the middle-ranking caste in the overall hierarchy. The BAC group represent the traditional farming castes, and the OBC group represent the traditional artisan castes. The BAC group is ranked higher than OBC. The SC are the lowest in the caste ranking, formerly known as the untouchable castes.

All else equal, lower caste water buyers have agricultural yields that are 45 percent higher if they reside in a village where the majority of water sellers are of the same caste compared to one where they are not.

This paper provides some empirical support that significant trading opportunities, even for a relatively simple and homogeneous good, can remain unrealized due to social or cultural distance.[3] Relative to trade in complex, heterogeneous, quality varying goods or services, trade in private water markets should be relatively simple. Such trade usually consists of a simple bilateral agreement between two individuals residing within close proximity (a proximity usually shared by their families for generations). That trade breaks down, particularly as the documented gains from trade are enormous, is at least suggestive that underlying distrust, as proxied by our measure of social distance, may be crucial in the development of markets in other more complex contexts.

The paper's main identification strategy relies on village level variation in caste dominance. It is therefore crucial to establish the exogeneity of this variation with regard to economic outcomes today. Section I provides evidence supporting this. Section II demonstrates the main empirical finding that low-caste households have significantly higher agricultural income if they reside in a village where the large landowning castes are from a low-caste group compared to a high-caste group. Estimations at the household level in Section III point to the importance of private groundwater markets in explaining these differences across village types. In particular, the positive effect of village caste dominance on agricultural income seems strictly related to being a buyer in the private groundwater markets. That is, low-caste water buyers seem to gain better access to irrigation if they are in a village that is, in turn, dominated by a lower caste. This finding is robust to an instrumenting strategy that treats private water market activities as endogenous. Alternative explanations are considered in Section IV, and Section V concludes.

I. Village Caste Dominance

This section provides evidence to establish the exogeneity of village level variation in caste dominance with regard to economic outcomes today. The identification strategy relies on two main claims. The first claim is that village level caste composition and land settlement patterns have remained essentially unchanged for centuries. This claim is related to the identifying strategy used by Banerjee and Somanathan (2007) at the district level to understand country-wide differences. Here, I present a discussion pertaining to the level of this data. The second claim is that the only important change that has occurred is a statewide land reform in 1950. This claim is also related to the identification strategy in previous work by Timothy Besley and Robin Burgess (2000), who exploit statewide variation in Indian land reform policies to understand country-wide differences today. The land reform in question exogenously altered land ownership rights of the different caste groups and, in turn, the village level caste dominance that we observe today. I present a

[3] George A. Akerlof (1976) demonstrates how sanctions of a caste system can prevent efficient outcomes in a theoretical framework.

detailed discussion of this *zamindari* reform in the present context, and document precisely how it impacted village caste dominance.

The data come from the same region of northeast India, the two bordering states of Uttar Pradesh and Bihar. There are essentially two dominant caste groups: the upper castes (primarily made up of Brahmins and Rajputs) and the BACs (mainly Yadavs). These large caste groups (Brahmins, Rajputs, and Yadavs) tend to be evenly spread throughout both Uttar Pradesh and Bihar.[4] Variation in caste dominance is not unusual in the region under study. Hetukar Jha (1991), for instance, analyses colonial documentation from the early 1900s on the caste composition of 5,475 villages located in Bihar and finds that approximately 39 percent of the Hindu villages have no upper caste members in residence. This is exactly the variation exploited in the current data where for 42 percent of the villages there are no upper castes present. That is, what determines if a lower caste group is dominant is simply the lack of presence of an upper caste group. In other words, in villages where upper castes are present, they own the majority of the land and are therefore dominant. On the other hand, in villages where upper castes are not present, a lower caste group owns the majority of the land and forms the dominant caste. Consistent with the colonial documentation, the most likely hypothesis is that there has never been a large upper caste presence in these low-caste dominated villages.

Village anthropological studies reveal that the origins of the distribution of caste groups at the village level go back hundreds of years. Settlement of the area under study can be traced back many centuries to Aryan occupation (which dates to 1500 BC). According to Thomas R. Metcalf (1979), basic elements of the village system and various cultivating castes, such as the Yadavs (the main BAC group in our study), were established early in the sixth century. During subsequent centuries, cultivation slowly extended across the fertile plains. These resident cultivators, together with their artisan (now classified as OBCs) and untouchable dependents (now SCs), generated the wealth that sustained society. Members of the noncultivating Brahmin caste (priests) were also present in the villages. The Muslim invasion of western India, beginning in the twelfth century, led to a mass arrival of dispossessed Rajput colonies into the region. During the next three centuries these Rajput exiles spread successfully, so that by the time the British arrived in the late eighteenth century, the Rajput caste owned and controlled the majority of the land.

Under colonial rule, the zamindari system of land tenure was in place in Uttar Pradesh and Bihar. *Zamindars* (landlords) were declared proprietors of land who paid revenue to the government.[5] The status of the zamindars was initially determined by their pre-colonial position of domination, and the zamindars mainly comprised members of the Rajput caste. New regulations during the colonial period initiated a slow decline of Rajput territorial power, opened up zamindari rights to members of the Brahmin caste, and opened occupancy rights to the cultivating castes (BACs). The land ownership of zamindars varied significantly and extremely

[4] These proportions by caste are according to the 1931 census, which is the last national census with detailed information on caste groupings.

[5] Other states in colonial India instead gave land rights to cultivators. Banerjee and Iyer (2005) exploit this historic statewide variation in land rights to explain agricultural outcomes today.

large landholdings would extend over several villages. When this happened, these high-caste landlords were absentee, i.e., resided elsewhere (Metcalf 1979).

At the time of independence, the states of India legislated large scale land reforms. In Bihar and Uttar Pradesh, the "Zamindari Abolition" took place just after independence in 1950. This land reform stripped the large zamindars of the majority of their landholdings. The land rights of the former zamindars were transferred to the former permanent tenants (primarily the BACs), thus giving rise to a new class of landowners comprised mainly of the BACs. After the reform, these former tenants owned more than two-thirds of the total state land (Walter C. Neale 1962).

Consequently, the 1950s saw a significant redistribution of land ownership across caste groups, though the actual land being cultivated by the tenant caste (BAC) did not change. Since the land reform, other changes in land ownership and distribution have been almost entirely due only to the process of inheritance and partition (land is typically divided amongst sons), with the combined ownership of each dynasty remaining fairly constant. Formal sales of land are rare. Village level studies from the region estimate that less than 1 percent of land is sold each year (Jean Drèze, Peter Lanjouw, and Naresh Sharma 1999).

Given this history, the caste composition and land settlement patterns across villages in our sample have likely been maintained for centuries. Whereas the variation in caste dominance was determined by the Zamindari Abolition. In the villages where upper caste groups are present today (high-caste dominated villages), these upper castes were possibly resident landlords in the colonial period or inherited their land. In contrast, the low-caste dominated villages in our sample were most likely controlled by absentee high-caste landlords. After independence, tenants (the BACs) were given ownership rights to the land they had always cultivated, so that these lower castes (BACs) now make up a significant proportion of the land owning households in all villages. The key distinction is that in villages where there are no upper castes present (low-caste dominated villages), the dominant caste in these villages is now a low caste. By contrast, in villages where upper castes are present, although the lower castes now also own their land, the upper caste group owns more and hence form the dominant caste.

In the main analysis of the paper we will treat the historically determined village level caste dominance described above as exogenous to economic outcomes today. A concern with this assumption is the possibility of caste-based migration in response to economic outcomes, or possibly to policy changes since independence, which would in turn directly alter village level caste composition today. This concern is not warranted here. Given the strict rules governing hereditary caste rankings, there is virtually no mobility of individuals across the different caste groups. Moreover, as recently analysed by Kaivan Munshi and Mark Rosenzweig (2005), there is also very little caste-based migration in India (refer to Anderson 2009 for more details). This is conjectured to be primarily because of reliance on sub-caste networks of mutual insurance, which do not transgress village boundaries. In general, there is no evidence of a large scale caste-based migration, which would explain why in 42 percent of villages of the current sample, there are no upper castes present. Especially as for almost the same proportion of villages (39 percent) in the colonial period of 1900, there were also no upper castes present.

It seems then that the caste composition of the villages under question has not changed systematically for some time. Though there is no historical village level data, we can confirm that the caste composition by district of the current data matches that of the 1921 census (refer to Anderson 2009).

Empirical Evidence.—In this section, we verify that there are no important differences across the two types of villages: those where total land ownership is dominated by the upper castes, and those where total land ownership is dominated by a lower caste (BACs).

The primary data used in this paper were collected by a team of researchers based at the World Bank and in India in 1997–1998. The villages of study are located in south and southeastern Uttar Pradesh and north and central Bihar. Uttar Pradesh and Bihar, together with Madhya Pradesh, have been referred to as India's "poverty belt." All three states are characterized by unusually large populations with per capita expenditure levels far below the poverty line. Eastern and Southern Uttar Pradesh, from where the study villages were drawn, are generally poorer than the western part of the state, and poverty levels have been rising in recent years. Bihar, which lies just east of Uttar Pradesh, has the lowest per capita rural income in India, and is the most rural state in the country. Overall poverty levels are even higher in Bihar than in Uttar Pradesh, and highest in the Northern region.

The field survey was administered in villages drawn at random from 12 districts in Uttar Pradesh and 13 districts in Bihar. A total of 120 villages, with an overall sample size of 2,250 households, were sampled: 57 villages in Bihar and 63 in Uttar Pradesh. All of the study villages are rural and the economies in these areas are primarily dependent on agriculture.

Information on the village caste composition comes from village level questionnaires. Land ownership is either in the hands of the upper castes or the lower castes (the BACs) (48 percent and 42 percent respectively). This is the key variation in the data, which we will exploit. We will compare villages where upper castes own most of the land (termed high-caste villages in the tables that follow), to villages where the BACs own most of the land (termed low-caste villages in the tables that follow).[6] Table 1 compares village-level characteristics across the two types of villages.

Aside from the total number of households and area of the village, the top part of Table 1 reflects all of the information available regarding the quality of infrastructure and public goods at the village level. We see that, on average, high-caste dominated villages fair slightly better than low caste ones, but that the differences are never significant. The lower part of the table lists agricultural characteristics, such as cropping patterns and prices and land/soil quality and prices. Again the two types of villages look extremely similar along these observable dimensions.

[6] We do not analyze Muslim dominated villages (which form 10 percent of our total village sample) and also drop Muslim households from the analysis (which comprises only 2 percent of the sample in Hindu dominated villages).

Table 1—Village Characteristics by Caste Dominance

	High-caste village	Low-caste village	Equivalence of means
Number of households	246.7 (173.2)	278.3 (163.2)	31.6 (35.6)
Area (hectares)	275.7 (501.5)	274.7 (260.3)	−1.1 (88.5)
Landless households	0.22 (0.18)	0.22 (0.20)	0.0 (0.04)
Electrified households	0.38 (0.29)	0.35 (0.28)	−0.03 (0.1)
Main drinking source—hand pump	0.44 (0.50)	0.36 (0.48)	−0.1 (0.1)
Main drinking source—well	0.54 (0.50)	0.64 (0.48)	0.1 (0.1)
Paved road	0.60 (0.49)	0.45 (0.50)	−0.15 (0.1)
Access to natural water sources	0.35 (0.48)	0.45 (0.50)	0.09 (0.10)
Bus stop (km)	3.6 (3.9)	3.9 (2.9)	0.2 (0.7)
Telephone service (km)	5.4 (6.2)	8.1 (8.6)	2.7 (1.6)
Police station (km)	7.7 (4.6)	8.3 (5.3)	0.6 (1.0)
Bank (km)	5.1 (4.5)	5.4 (5.0)	0.3 (1.0)
PDS shop (km)	1.3 (1.4)	1.8 (2.7)	0.5 (0.5)
Primary school (km)	0.5 (0.7)	0.7 (1.0)	0.2 (0.2)
Middle school (km)	2.8 (2.3)	2.9 (2.5)	0.1 (0.5)
Secondary school (km)	5.1 (4.1)	5.6 (4.6)	0.5 (0.9)
Hospital (km)	20.7 (15.7)	21.0 (16.5)	0.3 (3.4)
PHC (km)	5.0 (5.4)	5.9 (4.9)	0.9 (1.1)
Price of irrigated land (per acre)	119,213 (82,923)	123,367 (142,281)	4,153 (24,187)
Price of non-irrigated land (per acre)	55,950 (34,526)	48,964 (36,615)	−6,986 (7,504)
Main crop is paddy	0.62 (0.49)	0.62 (0.49)	−0.01 (0.1)
Main crop is wheat	0.21 (0.41)	0.21 (0.41)	0.01 (0.09)
Price of paddy (Rs/100 kg)	323.4 (56.8)	344.2 (109.1)	20.8 (19.8)
Price of wheat (Rs/100 kg)	455.7 (70.6)	455.0 (71.5)	−0.7 (15.1)
Almost no land suffers from floods	0.46 (0.50)	0.40 (0.50)	−0.05 (0.11)
Almost no land suffers from alkalinity	0.71 (0.46)	0.74 (0.44)	0.03 (0.1)
Almost no land suffers from waterlogging	0.48 (0.50)	0.57 (0.50)	0.09 (0.11)
Almost no land suffers from soil erosion	0.75 (0.44)	0.71 (0.46)	−0.04 (0.09)
Observations	48	42	

Notes: Area of the village and access to natural water sources come from the village amenities data of the Census of India 2001. Standard deviations are in parentheses in columns 1 and 2, standard errors are in column 3. PDS refers to the Public Distribution System of food grains to the poor at subsidized prices and PHC refers to Primary Health Care Facility.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Source: UP-Bihar LSMS World Bank (Village Questionnaire).

As a robustness check, we looked also to the village level amenities data from the 2001 Census of India for the villages in our sample (refer to Anderson 2009). Again there were no significant differences across the two types of villages along many dimensions at the village level.

These findings are somewhat in contrast to Banerjee and Somanathan (2007). They find: areas with a larger proportion of high castes have better public good provision; and those with more caste heterogeneity have worse. One likely reason for the differences here is that we are exploiting village level variation in one small

Table 2—Household Characteristics by Caste Dominance

Variable	High-caste village	Low-caste village	Equivalence of means
Literate	0.33 (0.47)	0.47 (0.50)	0.14 (0.03)***
Total income	3,999.0 (6,590.6)	7,524.5 (16,173.3)	3,525.5 (709.6)***
Crop income	1,891.6 (5,921.0)	5,201.2 (14,755.4)	3,309.6 (646.2)***
Land owned	3.0 (6.1)	3.1 (4.5)	0.1 (0.3)
Land value	83,630 (59,019)	88,122 (103,418)	4,492 (5,788)
Crop income/acre	1,602.1 (2,093.2)	2,813.1(3,012.1)	1,211.0 (249.6)***
Total yields	446.9 (780.9)	1,087.3 (1,703.7)	640.4 (91.3)***
Crop variety	3.2 (1.3)	4.0 (1.7)	0.7 (0.1)***
Percent land irrigated	0.78 (0.33)	0.85 (0.27)	0.06 (0.02)**
Tubewell irrigation	0.40 (0.49)	0.46 (0.50)	0.06 (0.03)**
Buy water	0.67 (0.47)	0.75 (0.43)	0.08 (0.03)***
Tubewell owner	0.13 (0.34)	0.24 (0.42)	0.11 (0.03)***
Tenant	0.43 (0.50)	0.34 (0.47)	−0.09 (0.04)**
Landlord	0.08 (0.02)	0.07 (0.01)	0.02 (0.02)
Borrow from same caste/relative	0.24 (0.43)	0.37 (0.48)	0.13 (0.03)***
Borrow from higher caste	0.44 (0.50)	0.21 (0.40)	−0.23 (0.03)***
Observations	592	705	

Notes: Standard deviations are in parentheses in columns 1 and 2, standard errors are in column 3. Sample is low caste (BAC, OBC, SC). Total annual income is the sum of wage income (for all household members), household enterprise income, total crop income, transfers into the household (typically from relatives), and the total value of home production of in-kind receipts of crops and food. Income is measured in rupees, there are approximately 40 rupees to the US dollar. Crop income is total value of sales of all crops over the past year and excludes crops for the purpose of household consumption. The amount of land owned (in acres) reported is conditional on owning land at all. Crop income per acre is equal to the total value of sales of all crops over the past year divided by the total land. Crop yields are measured as the value of produce sales from each crop per acre of land cultivated under that crop. Crop variety is the total number of different crop types grown. Tenant is a dummy variable equal to one if any land was sharecropped, rented, or received as wages in the past year. Landlord is equal to one if any land was sharecropped or rented out or given as a wage payment. These two variables are reported conditional on being a cultivator. Borrowing from same caste/relative or higher caste are reported conditional on borrowing. The alternative is predominately a money-lender.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Source: UP-Bihar LSMS World Bank (Household Questionnaire).

region of India, whereas Banerjee and Somanathan (2007) are exploiting caste variation at the district level for the entire country. Secondly, the high-caste dominated villages in our sample have the highest proportion of higher castes, but also are the more heterogeneous villages in terms of caste populations. Therefore, it is also possible that the two effects found in Banerjee and Somanathan (2007) cancel each other out in our context.

II. Household Outcomes by Caste Dominance

In this section, we compare outcomes of lower caste (BAC, OBC, SC) households residing in both types of villages. All castes generate the majority of their income from agricultural cultivation activities. Household characteristics for the lower caste group (BAC, OBC, SC) by village type are described in Table 2.

The striking finding from Table 2 is that lower castes fair significantly better, in terms of household income, if they reside in villages where a lower caste is dominant.[7] This difference in income is mainly driven by agricultural income, where the median value is almost doubled when residing in low-caste dominated villages. If we look only to cultivator households, which form 75 percent of our total sample, we see that average crop yields are significantly higher in low-caste dominated villages, and that households tend to select a larger variety of crops.[8] Though not reported here, individual crop yields are higher for lower castes residing in low-caste dominated villages for almost all crops. These differences in agricultural income by caste dominance also persist across different land holding groupings and for the separate caste groupings (BAC, OBC, and SC).

We will now turn to estimations of agricultural income to demonstrate the robustness of this finding. That is, controlling for many factors, agricultural income is significantly higher in low-caste dominated villages. Even if we accept that village caste dominance is exogenous, there is still an econometric concern that the characteristics of lower caste individuals vary systematically across the two village types. In particular, they may have different endowments if they reside in low compared to high-caste dominated villages. As we see from Table 2, however, for the most important agricultural endowments, land ownership and land quality (as measured by the value of their land), there are no significant differences.[9] Another important endowment, education, however, does vary systematically. From Table 2, literacy is higher for low castes residing in low-caste dominated villages. We control for literacy in the estimations and we will see, however, that it is not the reason for why we observe higher agricultural yields in low-caste dominated villages.

Estimations.—The main estimating equation is as follows:

$$Y_{ivds} = \beta_0 + \beta_1 X_{ivds} + \beta_2 D_{vds} + \alpha_{ds} + \gamma_s + \varepsilon_{ivds}\,. \tag{1}$$

Subscript i refers to a household from the lower caste groups who reside in both types of villages (BACs, OBCs, and SCs). Y_{ivds} is crop income per acre of total land of household i, residing in village v, district d, and state s. X_{ivds} includes exogenous household controls such as education, land ownership, and caste identity. D_{vds} is our key variable of interest, which is equal to 1 if a village v (in district d and state s) is dominated by a lower caste, and equal to 0 if the village is instead dominated by an upper caste. α_{ds} and γ_s are district and state fixed effects, respectively, and ε_{ivds} is a regression disturbance term clustered at the village level.

[7] All results pertaining to household income also hold for per capita income.

[8] The group of noncultivators are essentially the landless households.

[9] The value of land is the response to the question of how much would it cost per acre to buy this type of land. Although, actual land sales are few in the area, taxes from land revenue are collected by the local governments. Therefore, estimates of the value of land are likely known and this information can be gathered from government officials appointed to the villages.

TABLE 3—OLS ESTIMATIONS OF HOUSEHOLD CROP INCOME

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Literate	255.0	294.9	163.7	297.4	288.6	224.4
	(141.7)*	(115.4)**	(127.7)	(158.3)*	(145.6)*	(104.4)**
Total land	58.5	52.7	47.8	55.9	65.6	50.9
	(18.4)***	(18.1)***	(17.9)***	(19.6)***	(20.5)***	(15.6)***
Low-caste village	566.5	393.3	387.2	505.6	668.5	371.3
	(209.0)***	(191.6)**	(161.8)**	(198.6)**	(254.9)***	(167.6)**
Caste controls	Yes	Yes	Yes	Yes	Yes	Yes
State controls	Yes	Yes	Yes	Yes	Yes	Yes
District controls	No	Yes	No	No	No	No
Crop controls	No	No	Yes	No	No	No
Distance controls	No	No	No	Yes	No	No
Groundwater controls	No	No	No	No	Yes	No
Public goods controls	No	No	No	No	No	Yes
Observations	1,295	1,295	1,295	1,295	1,295	1,295
$\overline{R}^2$	0.13	0.28	0.20	0.14	0.17	0.12

Notes: The sample in the estimations are the lower castes (BAC, OBC, SC). Robust standard errors are in parentheses. Regression disturbance terms are clustered at the village level. Data for the groundwater controls come from the 2001 Census of India and the Central Water Board of India.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 3 reports the main estimations results from an ordinary least squares estimation of (1). The same results also persist if the dependent variable instead yields per acre of cultivated land, or alternative tobit estimations for the entire sample and also for only those households who received positive revenue for their crops.

The estimation results from Table 3 confirm the robustness of the positive relationship between agricultural income and residing in a low-caste dominated village. Columns 1 and 2 demonstrate that this result is robust to including exogenous household controls, such as education, land ownership, caste identity, and also district and state fixed effects. Column 3 includes crop choice controls that might be deemed an endogenous determinant of crop yields.[10] The estimation in column 6 includes measures of village-level public goods. The main results are also robust to including alternative measures of political power in terms of population and political leadership (refer to Anderson 2009). Therefore, standard political economy explanations for why economic outcomes might be related to village caste dominance do not seem to play a direct role here.

The evidence of this section raises questions as to why agricultural incomes are higher for low castes when they reside in low-caste dominated villages. We now turn to evidence suggesting that these households seem to gain better access to irrigation in low-caste dominated villages, and that this is the central reason for their significantly higher yields.

[10] The results are also robust to including all other relevant household controls such as value of land, inputs into production, and sharecropping measures (see Anderson 2009).

III. Access to Irrigation

This section provides evidence that points to the importance of irrigation access in explaining the observed household differences by caste dominance of the previous section.

In Section I, it was demonstrated that there are no significant differences in terms of most observable village features by caste dominance. There is, however, one key difference that does seem to exist. It appears that a greater proportion of cultivated land is irrigated in the low-caste dominated villages. The most important source of irrigation in both types of villages is private tubewells. Though not reported here, both the village level data and the more detailed village level amenities data from the 2001 census of India confirm that the total irrigated area and total irrigated area by private tubewells are both significantly higher in low-caste dominated villages, despite the fact that access to other natural water sources are, on average, the same relative to the high-caste dominated villages. The household level data, reported in Table 2, also demonstrates that the proportion of land irrigated is higher and that households are more likely to use private tubewell irrigation as their primary source in low-caste dominated villages. The proportion of low-caste households owning a private tubewell and pump is also systematically higher and so is the probability of being a private groundwater buyer in low-caste dominated villages. These relationships suggest that the distribution of private irrigation water seems to systematically vary across the two types of villages.

A. *OLS Estimation*

We now compare agricultural income by water market activities and village type to determine if these systematic differences in access to private irrigation are determining the higher yields in low-caste dominated villages. Table 4 reports analogous estimation results to Table 3 where the key dummy variable denoting low-caste village is interacted with water market activities.

Interestingly, once we include interaction terms, the effect of residing in a low-caste village no longer significantly affects crop income on its own. The positive effect seems strictly related to being a water buyer in the private groundwater markets. By contrast, there is no significant difference for tubewell owners, and it is not this water market activity that is explaining the systematic differences in agricultural incomes across the two village types. It appears rather that low-caste water buyers gain better access to irrigation if they are in a village that is in turn dominated by a lower caste. Estimations on yields from specific crops provide some further evidence to support this hypothesis. In these estimations we find that the positive results for water buyers in low-caste villages are strongest for irrigation intensive crops such as sugarcane and paddy. The results reported in Table 4 are robust to including numerous controls and also, though not reported here, to including additional interaction terms with village caste dominance interacted with all household characteristics, and crop, caste, district, and state controls.

Table 4—OLS Estimations of Crop Income with Irrigation, Credit, and Tenancy Variables

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Low-caste village	346.2	−88.0	19.0	−162.1	48.1	654.4	740.7
	(196.2)*	(165.9)	(137.1)	(186.3)	(175.0)	(245.9)***	(259.7)***
Tubewell owner	1,119.6	865.7	755.8	841.0	703.3		
	(260.0)***	(474.4)*	(497.0)	(532.7)	(524.1)		
Water buyer	318.1	−143.0	43.4	−233.9	−225.7		
	(127.9)**	(174.9)	(193.0)	(171.7)	(172.6)		
Tenant						79.3	
						(106.1)	
Landlord						−324.2	
						(173.0)*	
Borrow same caste							−234.8
							(125.0)*
Borrow higher caste							−28.6
							(161.0)
LCV × tubewell owner		388.5	365.3	773.2	654.8		
		(500.4)	(593.9)	(682.0)	(600.1)		
LCV × water buyer		850.9	602.1	980.7	901.8		
		(275.0)***	(254.4)**	(333.9)***	(273.7)***		
LCV × tenant						−420.7	
						(289.0)	
LCV × landlord						189.6	
						(323.2)	
LCV × bor. same caste							−482.7
							(292.2)
LCV × bor. high caste							−542.5
							(318.7)*
Caste controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District controls	Yes	Yes	No	No	No	No	No
Crop controls	No	No	Yes	No	No	No	No
Distance controls	No	No	No	Yes	No	No	No
Groundwater controls	No	No	No	No	Yes	No	No
Observations	1,295	1,295	1,295	1,295	1,295	1,295	1,295
$\overline{R^2}$	0.31	0.31	0.22	0.19	0.21	0.13	0.14

Notes: The sample is the lower castes (BAC, OBC, SC). Regression disturbance terms are clustered at the village level. All estimations include the exogenous household controls (literacy and land ownership). LCV refers to low-caste village.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

B. *Instrumental Variables Estimation*

The above OLS estimates demonstrate that the positive effects of village caste dominance on agricultural yields is directly linked to being a water buyer. In these estimations, there is still a concern that unobservable endowments of water buyers may systematically vary across low- and high-caste dominated villages. If this is so, a difference in endowments could instead explain the positive coefficient of the interaction between water buyer status and residing in a low-caste village, rather than the conjecture that low-caste water buyers gain better access to irrigation in these villages. To address this issue we instrument for water buyer status.

For this purpose, we use village level information on access to natural water sources and total area of the village. We would expect that both of these variables reduce private groundwater purchases. Access to alternative water sources should

reduce the need to purchase private groundwater. Relatedly, the larger the area of the village, the higher the transaction costs in water trade (in terms of distance covered to transport the water), and water purchases should decline. In the estimations, we need to instrument for both water buyer status and its interaction with village caste dominance. Therefore, it is very important to establish that these two village-level instruments do not significantly vary by village caste dominance. This is demonstrated in Table 1, where we see there are no significant differences in access to natural water sources and total area of the village across low and high-caste dominated villages. The spirit of the identification strategy is that village level access to natural water sources and higher transaction costs in groundwater trade both cause private groundwater markets to be thin. The difficulties in groundwater trade that low castes face in high-caste dominated villages are worse when these markets are thin.

More specifically, in the second stage we estimate an analogous regression to (1):

$$(2) \qquad Y_{ivds} = \gamma_0 + \gamma_1 X_{ivds} + \gamma_2 D_{vds} + \gamma_3 WB_{ivds} + \gamma_4 D_{vds} \times WB_{ivds} + \theta_{ds} + \sigma_s + \epsilon_{ivds},$$

where Y_{ivds}, X_{ivds}, and D_{vds} are defined as in (1). WB_{ivds} is equal to one if household i (residing in village v, district d, and state s) is a water buyer, and zero otherwise. θ_{ds} and σ_s are district and state fixed effects, respectively, and ϵ_{ivds} is a regression disturbance term clustered at the village level.

In the estimation of (2), we treat WB_{ivds} and its interaction with village caste dominance, $D_{vds} \times WB_{ivds}$, as endogenous regressors. To this end, as recommended by Joshua D. Angrist and Jorn-Steffen Pischke (2009, 191), we first estimate the following:

$$(3) \qquad WB_{ivds} = \alpha_0 + \alpha_1 X_{ivds} + \alpha_2 D_{vds} + \alpha_3 Z_{vds} + \delta_{ds} + \phi_s + \xi_{ivds},$$

where Z_{vds} is a vector of two instruments: whether a village, v (in district d, and state s) has access to natural water sources (canal, river, lake, or pond), and the total area of the village (in hectares). δ_{ds} and ϕ_s are district and state fixed effects, respectively, and ξ_{ivds} is a regression disturbance term clustered at the village level.

We then use the predicted value, $\widehat{WB}_{ivds}$, from (3), and its interaction with village caste dominance, $D_{vds} \times \widehat{WB}_{ivds}$, as instruments in the two first-stage estimations of WB_{ivds} and $D_{vds} \times WB_{ivds}$, respectively (4) and (5) below, in a conventional 2SLS procedure.

$$(4) \qquad WB_{ivds} = \pi_0 + \pi_1 X_{ivds} + \pi_2 D_{vds} + \pi_3 \widehat{WB}_{ivds} + \pi_4 D_{vds} \times \widehat{WB}_{ivds} + \tau_{ds} + \rho_s + \eta_{ivds}.$$

$$(5) \qquad D \times WB_{ivds} = \lambda_0 + \lambda_1 X_{ivds} + \lambda_2 D_{vds} + \lambda_3 \widehat{WB}_{ivds} + \lambda_4 D_{vds} \times \widehat{WB}_{ivds} + \psi_{ds} + \omega_s + \upsilon_{ivds},$$

TABLE 5—IV-2SLS ESTIMATIONS OF CROP INCOME WITH WATER BUYER VARIABLE

Variable	First-stage water buyer	First-stage water buyer	First-stage LCV × water buyer	Second-stage crop income
Low-caste village	0.070	0.17	0.12	−1,522.8
	(0.048)	(0.11)	(0.09)	(814.1)*
Water buyer				202.1
				(931.6)
LCV × water buyer				3,519.5
				(1,413.7)**
Area	−0.0002			
	(0.00004)***			
Natural water sources	−0.19			
	(0.04)***			
$\widehat{\text{Water buyer}}$		1.03	−0.06	
		(0.16)***	(0.12)	
LCV × $\widehat{\text{water buyer}}$		−0.26	0.83	
		(0.19)	(0.14)***	
Caste controls	Yes	Yes	Yes	Yes
State controls	Yes	Yes	Yes	Yes
District controls	Yes	Yes	Yes	Yes
F-stat on instruments	27.9	21.3	19.7	
Observations	1,127	1,127	1,127	1,127

Notes: The sample is the lower castes (BAC, OBC, SC). Regression disturbance terms are clustered at the village level. All estimations include household land ownership. LCV refers to low-caste village.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

where τ_{ds} and ψ_{ds} are district fixed effects; ρ_s and ω_s are state fixed effects; and η_{ivds} and υ_{ivds} are regression disturbance terms clustered at the village level.

The estimation results for the first-stage estimations of (3), (4), and (5) are reported in the first three columns of Table 5 respectively.

We see that our instruments (village area and access to natural water sources) are indeed strong negative predictors of water buyer status. The fourth column of Table 5 reports the results from the second-stage estimation, equation (2), on agricultural income. There it is demonstrated that the key finding of the OLS estimations (Table 4) is robust to this instrumenting strategy. That is, the coefficient on the instrumented variable of the interaction term between water buyer and low-caste dominance, γ_4 in equation (2), is positive and significant at the 5 percent level.

We now explore further water markets in the area to better understand these findings.

C. *Groundwater Markets*

The study area is located in what is known as the Ganga basin. This region has enormous groundwater potential, and informal groundwater markets have emerged as an extremely important institution over the last three decades (Aditi Mukherji 2004). Although a number of public tubewells were installed in the region in the

1970–1980s, most of them have become nonoperational due to erratic and inadequate supply of electricity and a lack of repair and maintenance (T. Prasad 1993). As a result, groundwater utilization for agricultural purposes in this region is predominantly through private tubewells.[11] The exploitation of groundwater through the use of tubewells converted fallow land of the dry season into fertile paddy fields well suited to the seed-fertilizer technology available through the wave of the Green Revolution. As a result, both cropping intensities and patterns vastly improved, and yield rates witnessed a tremendous upward swing (Pant 2004; Mukherji 2004).

Most of the groundwater development, which took place through private water extraction mechanisms, was skewed in favor of larger farmers with a higher ability to invest (D. R. Singh and R. P. Singh 2003; Prasad 1993). Nevertheless, the emergence of groundwater markets has been seen as an opportunity for more equitable access to groundwater irrigation for resource poor small and marginal farmers (K. Shankar 1992). In spite of the inequities in terms of pump and tubewell ownership, poorer farmers do tend to fair better with the development of private tubewells, where having access to groundwater at all has been a key to success (Tushaar Shah 1993). Several studies show that cropping intensities and yields of tubewell owners and water buyers are comparable, suggesting that buyers are in fact receiving reliable and adequate irrigation water (Shah and Vishwa Ballabh 1997; Kishore 2004). Shah (1993) cites studies that show more than half of the total area irrigated by private irrigation systems, in many parts of India, belong to water purchasers.

This previous research highlights the very large returns to groundwater irrigation for water buyers. The main empirical finding here, that lower caste water buyers have agricultural yields 45 percent higher if they reside in a village dominated (in terms of land ownership) by a lower caste compared to a higher caste, suggests that buyers are getting better access to water in these former villages. Given that it is the larger landowners who own the irrigation pumps, a key difference between the two types of villages is that in the high-caste dominated villages, the majority of water sellers are from the high caste, whereas in the low-caste dominated villages, the water sellers are correspondingly of a lower caste.[12] The empirical results are therefore consistent with a scenario where lower caste water buyers obtain better access to groundwater when the majority of sellers are also of a lower caste, compared to when the majority of sellers are from a higher caste. We now turn to exploring why private groundwater trade may break down across caste groups.

D. *Groundwater Contracts and Trade Breakdown*

Groundwater markets are characterized by barriers to entry that arise from the lumpiness of the tubewell investment coupled with credit constraints. [13] Moreover,

[11] Niranjan Pant (2004) documents stupendous growth in private tubewells in Uttar Pradesh. The estimated total was 3,000 in 1951, 600,000 in 1977, and 1.05 million by 1980. Between 1986–1987 and 1992–1993, the density of tubewell and pump sets increased fourfold in Bihar (Avinash Kishore 2004).

[12] In the high-caste dominated villages, 59.5 percent of the pump owners are from the upper castes. In the BAC dominated villages, 69.9 percent of the pump owners are BAC.

[13] Tubewell installation costs amount to roughly a year's income for the average rural household (Hanan G. Jacoby, Rinku Murgai, and Saeed ur Rehman 2004).

due to conveyance costs involved in irrigating fields that are far, there is only a limited area surrounding the tubewell that can economically be irrigated by a given tubewell. As a result, market competition is likely to be weak and typically a given tubewell owner accommodates a small number of buyers within close physical proximity, and often water buyers are restricted to the choice of a single seller (Shankar 1992). Water is generally transported to the buyer's field either through unlined or lined field channels. Matters of water conveyance, clearing the channel, and closing in between outlets are the responsibility of the water buyer. The main responsibility of the tubewell owner is to switch on and off the water pump. Bilateral oral contracts generally dictate the terms of exchange of groundwater transactions between buyers and sellers.

In addition to driving up the price, monopoly power can affect quality of service with regards to adequacy and reliability of supply provided by sellers, where buyers have little recourse. Tubewell owners often follow a schedule of rotation for irrigating the fields of all of the buyers (Vikas Rawal 2002 and Navroz K. Dubash 2000). Despite this institutionalized system, sellers are often reported to discriminate between buyers, particularly in times of irrigation shortages. There are reports of harassment of water buyers as sellers angle to extract more rents by threatening to reduce the water supply (Rawal 2002). Due to the moral hazard problems involved, exchange relationships are retained through village level institutions and norms (Dubash 2000). Studies report that the ability to pay the price does not guarantee access to groundwater, farmers must be networked (G. Wood 1999). In general, transactions are not impersonal, but are part of inter-linkages where sellers tend to give preference to relatives or members of their own caste, either through lower water rates or priority for service (Mukherji 2004; Dubash 2000; Wood 1999). Lower water prices often imply a more balanced relationship of mutuality between buyers and sellers (Pant 2004; Shah 1993). Kei Kajisa and Takeshi Sakurai (2000), using a sample of villages in Madhya Pradesh, find that 62 percent of water transactions are conducted between buyers and sellers of the same caste, and of the 38 percent of transactions conducted between different caste groups, trade occurs between groups with the least social distance, as measured by caste rank. In a study from northeast Bihar, Wood (1999) similarly reports that traders operate strictly within their own caste, and in the case of a numerically dominant landholding caste, trade is further restricted within particular extended lineages. Valerie Kozel and Barbara Parker (1999) report similar concerns prevailing between water sellers and buyers throughout the present study area.

This evidence suggests that caste ties or dominance may play a significant role in the enforcement of informal groundwater contracts within villages and that, in particular, trade is more common between members of the same caste. A very simple framework could explain why caste could matter to these bilateral oral contracts. In the context of private groundwater markets, the buyer and seller share some surplus from interaction where the monopolistic seller sets the price. The environment is akin to the classical hold-up problem where agents undertake match specific investments prior to the exchange.[14] In this case, buyers must incur the fixed

[14] Refer to recent paper by Jacoby and Ghazala Mansuri (2008) who find evidence of the hold-up problem among land tenants in rural Pakistan. They demonstrate that land specific investments are lower on leased plots.

costs of the water channels between their field and that of the seller. If contracts are not enforceable between the buyer and seller, then the seller will obtain all of the surplus to trade and request an ex post price at which it is never worthwhile for the buyer to incur the match specific investment and trade will not occur. This hold-up problem may well exist throughout the cropping season. Water intensive crops typically require regular irrigation, and water sellers can turn off the switch to the pump at any time during the season. Water sellers likely cannot commit to a given volume of water over the cropping season. On the buyer's side, match specific investments may well be quite large. Given that buyers often face only one seller, match specific investments on their own land could include costs associated with making water intensive crop choices, such as sugarcane, not simply the costs of building water channels.

The claim in this paper is that trade breaks down between high-caste water sellers and low-caste water buyers in high-caste dominated villages, whereas in low-caste dominated villages, trade occurs between low-caste water sellers and buyers. There are several ways in which this hypothesis is consistent with the implications of the simple framework outlined above in the context of the data. Generally the dominant caste in the village is responsible for resolving disputes between trading partners through their control over the village panchayats (governments) or councils (Rawal 2002). Moreover, the dominant caste typically favors members of its own caste (Srinivas 1987). In this sense, a seller of the dominant caste who deviates ex post from its committed price is more likely to be punished if the buyer is also from the dominant caste compared to when the buyer is from a lower ranked caste. In other words, there are no instruments for a lower caste buyer to punish a higher caste seller when the upper castes maintain the political power in the village. Given the inherent social hierarchy of the caste system, it is not even necessary for the higher castes to dominate the village in order for this to be true. There are numerous historical and social reasons for why it would be very difficult for a lower caste member to punish someone from a higher caste. As a result, trade between high-caste water sellers and low-caste water buyers will break down because contracts are not enforceable, whereas trade between buyers and sellers of the same caste are more easily enforced.

The simple framework outlined here points to the inability of sellers to credibly commit to a price to explain why trade can break down across caste groups. This contrast with the previous literature on groundwater contracts has emphasized the moral hazard on the buyers side and highlighted the importance of interlinkages between tenancy and credit relations to solve this enforcement problem. In this scenario, high-caste landlords (or lenders) could potentially use their long-standing power to better enforce agreements with tenants or borrowers who are of lower caste. As a result, we would expect more groundwater trade across castes. This reasoning would suggest that because trade seems to break down across castes in the context of our data, it is the moral hazard on the sellers' side that may be the more important limit to water trade. Further support for this hypothesis will be seen in Section IVC, where it is demonstrated that inter-linkages between tenancy and credit relations do not play a central role here.

Table 6—Water Markets at the Village Level

	All castes			Low castes		
	High-caste villages	Low-caste villages	Equiv. of means	High-caste villages	Low-caste villages	Equiv. of means
Acres cultivated	883.8	915.2	−31.3	433.4	886.0	−452.6
	(825.7)	(803.2)	(172.3)	(571.8)	(784.7)	(143.5)***
Acres owned	468.6	436.4	32.2	172.3	415.8	−243.5
	(436.5)	(414.4)	(90.1)	(277.5)	(403.8)	(72.3)***
Population	208.1	236.9	−28.7	140.0	230.8	−90.7
	(127.2)	(140.1)	(28.2)	(92.9)	(136.6)	(24.4)***
Buyers per capita	0.70	0.75	−0.05	0.73	0.75	−0.02
	(0.27)	(0.22)	(0.05)	(0.31)	(0.23)	(0.06)
Buyers/land cultivated	0.21	0.22	−0.003	0.42	0.22	0.20
	(0.24)	(0.15)	(0.04)	(0.54)	(0.16)	(0.09)**
Buyers/land owned	0.62	0.46	0.16	1.45	0.47	0.99
	(1.90)	(0.33)	(0.30)	(2.32)	(0.33)	(0.36)***
Sellers per capita	0.10	0.12	−0.02	0.06	0.11	−0.05
	(0.11)	(0.11)	(0.02)	(0.11)	(0.12)	(0.02)**
Sellers/land cultivated	0.02	0.03	−0.01	0.02	0.03	−0.015
	(0.03)	(0.04)	(0.007)	(0.03)	(0.04)	(0.007)**
Sellers/land owned	0.05	0.06	−0.01	0.04	0.06	−0.02
	(0.07)	(0.06)	(0.01)	(0.08)	(0.06)	(0.01)
Pumps per capita	0.15	0.18	−0.03	0.09	0.23	−0.14
	(0.11)	(0.14)	(0.03)	(0.13)	(0.17)	(0.03)***
Pumps/land cultivated	0.04	0.06	−0.015	0.02	0.06	−0.03
	(0.03)	(0.06)	(0.01)	(0.03)	(0.06)	(0.01)***
Pumps/land owned	0.09	0.11	−0.027	0.06	0.11	−0.05
	(0.08)	(0.09)	(0.02)	(0.09)	(0.09)	(0.02)**
Observations	48	42		48	42	

Notes: Standard deviations are in parentheses except in the third column where standard errors are in parentheses. The sample is all castes (upper, BAC, OBC, SC) in the first panel of the table and the sample is low castes (BAC, OBC, SC) in the second panel.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

E. *Aggregate Water Market Activity*

The idea that groundwater trade across caste groups could breakdown in villages where upper castes dominate is consistent with the empirical results of Section III. There it was demonstrated that lower caste water buyers have better access to groundwater irrigation if they reside in villages where their own caste dominates and where the majority of sellers are from their same caste. Further empirical evidence consistent with the claim that a trade breakdown across caste groups is an explanation for the high yields of water buyers in low-caste dominated villages comes from looking at aggregated water market activity.

First suppose that caste status is irrelevant to groundwater trade arrangements. With this premise in mind, Table 6 attempts to capture the demand and supply in water markets at the village level. The total cultivated and owned land across two types of villages is computed using average land values, from the household level

data, multiplied by the total number of households within each village, using the village level data. Similarly, the number of water buyers, sellers, and pump owners relies on averages computed using the household level data of cultivators.[15]

The first panel of Table 6 considers the entire sample of all castes. We see that there is no evidence that the total amount of cultivated land is significantly different across the two types of villages. This also holds true if we look at the total acres cultivated by crop. Nor are there significant differences between the total number of buyers, sellers, and pump and tubewell owners. In other words, just looking at aggregate water market activity across the two types of villages does not explain why lower caste buyers are obtaining better access to groundwater irrigation in the low-caste dominated villages. Given these village comparisons, there is no evidence that the price or supply of water is different across the two types of villages.

In the second panel of Table 6, we construct the same comparisons using only the lower castes (BAC, OBC, SC). These results, on the other hand, reveal significant differences. In the aggregate, it appears that the number of water sellers (and pump owners) is significantly higher in low-caste dominated villages, whereas the number of buyers is significantly lower relative to the high-caste dominated villages, if we only consider the lower castes. A plausible explanation arises if we suppose that upper castes do not trade water with lower castes. Then the aggregate information, provided in the second panel of Table 6, implies that water prices faced by the lower castes should be lower in the low-caste villages. This could explain why low-caste water buyers are able to gain better access to irrigation, and consequently produce significantly higher yields when residing in low-caste dominated villages. The implications are that the presence of upper castes, who own a substantial proportion of the private tubewells and pumps, causes an inefficiency in the distribution of groundwater as they do not readily trade with lower castes in villages.

The conjecture that same caste individuals more easily trade is consistent with the recent work of Munshi and Rosenzweig (2005, 2008) who emphasize the importance of same caste networks in determining economic and political outcomes. Their work focuses on networks at the sub-caste or *jati* level. Here the analysis has instead focused on the broad caste groupings of upper castes compared to BACs. It is important to note though that in this particular context, these two groups are essentially defined at the sub-caste level. That is, the upper caste group is made up primarily of Brahmins and Rajputs and the BAC group is composed mainly of Yadavs.

IV. Alternative Explanations

The main empirical finding of this paper is that farm yields are systematically and substantially higher for low-caste households residing in villages dominated by lower castes compared to higher castes. Our favored interpretation is that agricultural yields are crucially determined by access to groundwater irrigation, which is distributed through private markets. These markets seem to work more efficiently in villages where the caste composition is more homogeneous. We conjecture that

[15] This aggregated data must be interpreted with caution as the total number of households per village in the household level data is only between 15 and 30.

social identity is playing a key role since bilateral contracts are more easily enforced between members of the same caste. Our empirical support for this hypothesis is limited as we do not have detailed data on the terms of trade for groundwater. There are indeed alternative explanations for a trade breakdown in groundwater that are not dependent on a measure of social distance. However, in what follows, we do not find support for these alternative explanations in our data.

A. *Geographic Distance between Buyer and Seller*

One possibility is that caste identity is picking up the importance of geographic distance rather than social distance in groundwater trade relations. Indian villages are spatially divided into *hamlets* (neighborhoods) that are often segregated by caste, with upper caste hamlets geographically distant from lower caste ones. Correspondingly, there can be systematic spatial variation by caste in the location of agricultural plots. As already emphasized, groundwater transaction costs are higher the greater the distance between the seller's tubewell and the buyer's field. Therefore, it is possible that trade in groundwater is breaking down across the caste groups, only because upper caste sellers are located further away from lower caste water buyers, relative to lower caste sellers. In other words, groundwater transaction costs are higher in the upper caste dominated villages. There are, however, several findings in the data that go against this conjecture. The first is that if transaction costs are indeed higher, then we should expect that overall water market activity is lower in the high-caste dominated villages. This goes against the findings of Table 6, where there are no significant aggregate differences in terms of number of water buyers and sellers per acre of land cultivated or owned.

Secondly, we can, to some extent, control for geographic distance between upper castes and lower castes in the estimations. There are a few variables that should be correlated with the geographic distance between the upper and lower caste groups. For one, the total area of the village should be positively correlated with the physical distance between castes. Two variables that reflect population density and should be negatively correlated with the caste geographic distance are the number of hamlets per hectare of the village and the proportion of hamlets that are mixed caste; that is, those in which upper and lower castes co-reside. However, though not reported here, there are no significant differences with regard to these distance measures across the two types of villages. Moreover, the main results are robust to the inclusion of these distance variables (Table 3 column 4 and Table 4 column 4). Additionally, though not reported here, these distance measures have no explanatory power when they are interacted with the dummy variable reflecting village caste dominance or the dummy variable for water buyer.

B. *Groundwater Quality*

Groundwater accessibility and quality is highly dependent on the nature of aquifers and ambient climatic conditions. It is therefore possible that, for some reason, low-caste dominated villages are naturally better endowed with groundwater resources. As already mentioned, villages in the data come from the Ganga Plain,

which contains one of the best reservoirs of groundwater in the world. Moreover, the majority of districts contained in this data (Eastern Uttar Pradesh and Northern Bihar) come from the best endowed hydrogeological areas with high-yielding aquifers. Hydrogeological characteristics of course transgress administrative boundaries, and within a district there is topographic variation that determines groundwater potential. According to the Central Water Board, there is indeed district-wise variation in groundwater development and potential. The main estimations are robust to including such district-level information on average rainfall and evaporization rates, groundwater levels (pre and post monsoon), and the stage of groundwater development (Table 3 column 5 and Table 4 column 5). It is important to recall that almost all of the districts in the sample contain both types of villages (dominated either by a high caste or low caste), and also that determinants of village caste composition predate groundwater development in this area.

At the village level, we have information on natural water sources and also soil characteristics that are correlated with groundwater supplies, such as the degree of alkalinity, salinity, flood proneness, waterlogging, and erosion. There are, however, no significant differences with regards to these environmental measures across the two types of villages (Table 1). Moreover, the main results are robust to the inclusion of these variables (Table 3 column 5 and Table 4 column 5). Though not reported here, the results are also robust to including these environmental variables and the groundwater variables, mentioned above, interacted with the dummy variable reflecting village caste dominance.

C. *Tenancy and Credit Relations*

There are also other plausible reasons for why yields are higher in low-caste dominated villages that are not related to caste or groundwater markets. In particular, one might expect that tenancy relations are worse in high-caste dominated villages; perhaps upper caste landlords treat low-caste tenants poorly relative to a lower caste landlord. Likewise, we may expect that credit markets function better in low-caste dominated villages where there is more scope to borrow from members of one's own caste (refer to Munshi and Rosenzweig 2005). In the data, we have information on both tenancy and credit relations. We see from Table 2 that inter-linkages do exist. That is, lower castes are more likely to sharecrop land in high-caste dominated villages, and they are also more likely to borrow from same caste members in low-caste dominated villages. However, unlike the water market activity variables, these systematic differences do not explain the differences in crop yields across the two types of villages. From Table 4 (columns 6 and 7), we see that the key variable of caste dominance remains robust when we include measures that reflect tenancy and credit relations. In other words, these latter variables do not explain why crop income is significantly higher in low-caste dominated villages. These results remain robust when including numerous controls as well as alternative measures of tenancy and credit relations.

It is possible though that these alternative hypotheses play a central role and are not entering into the estimations significantly on their own, but, instead, are somehow complementary to the groundwater market results. For example, we might

expect complementarity between water markets and tenancy relations, where, as found by Jacoby, Murgai, and Rehman (2004) for Pakistan, tenants receive lower prices and better access to water relative to other water buyers. In this case, we would expect to see the complementarity between water buyers and being a tenant driving the results. However, results from estimations with additional interactions terms between the tenant and landlord dummy variables and water market activities, though not reported here, demonstrate that there is no significant complementary relationship between being a water buyer and a tenant. Likewise, there are no significant complementary effects between credit relations, water market activity, and village level caste composition.[16]

D. *Land Quality*

A main claim of this paper is that village caste dominance is directly affecting economic outcomes. However, there is another plausible explanation for why agricultural yields may be higher in villages dominated by the lower castes that have nothing to do with caste (or inter-caste relations) per se. Yields may be higher in low-caste dominated villages simply because lower caste households residing in these villages have higher quality land compared to their low-caste counterparts residing in the high-caste dominated villages. However, we do not find this in the raw data, where the quality of land (measured by the value of land) is not significantly different for low castes residing either in low or high-caste dominated villages (refer to Table 2). The main empirical results are robust to including controls that capture the quality of land (measured by the value of land). In these estimations, land quality is significantly positively related to agricultural income, however, it does not explain the differences in agricultural incomes by caste dominance. Moreover, it is pump owners who have the highest quality land, therefore we should expect to see their yields significantly higher if land quality is a key factor in explaining the differences. Instead, what we observe is that it is water buyers who are strictly better off.

It is possible though that this alternative hypothesis does play a central role and is not entering into the estimations significantly on its own, but instead is somehow complementary to the groundwater market results. For example, perhaps higher land quality is driving the results and estimations pick up its importance via groundwater markets only because the demand for irrigation is complementary to the quality of land. Although, if there is indeed complementarity between irrigation and land quality, we would again expect that the significant determinant of crop yields in low-caste dominated villages should not be via water buyers but pump owners. The wealthier households have the higher quality land and can afford to incur the fixed costs of a tubewell boring and hand pump. Nevertheless, we ran estimations that check for such interaction effects, however, there is no evidence of a complementary relationship between access to groundwater and quality of land. This is particularly the case for water buyers, where the impact of quality of land on crop income is negative for them relative to others.

[16] The raw data also do not reveal any noteworthy correlations between credit and tenancy relations and water market activity or village level caste composition.

E. *Land Size*

Two further possibilities are worth discussing. One salient feature of the data is that land inequality is higher in the high-caste dominated villages. This follows because, on average, upper castes own larger plots than other castes. The efficient distribution of water tubewells given landholdings should not be affected by who owns which plot. Therefore, it should not be the case that the distance between the fields of buyers and the tubewells of sellers increases with the total plot size of sellers. However, it is likely to be the case that larger landholders have greater monopoly power. In this sense, water buyers could be worse off in high-caste dominated villages just because there are a larger number of monopolists, irrespective of caste status. However, this claim is not consistent with the findings of Table 6, which demonstrate that the total number of sellers and buyers per acre of cultivated land are not significantly different across the two types of villages. Moreover, the significant correlation between agricultural yields and residing in a low-caste dominated village is robust to including a Gini index of land inequality. This also holds true if we include interaction terms.

Suppose alternatively that upper castes curtail the supply of water simply because they have larger landholdings and perhaps deeper tubewells. Then we might expect to see a higher price of water just because of the presence of wealthier households with larger landholdings, not because they are upper caste per se. The problem with this explanation is that the cultivation intensity of the upper castes is lower than that of the BACs, and therefore, proportionally, they should have more water available to sell. It should only be in their interest to sell off their excess supply of water. It would seem that caste as a measure of social distance from potential buyers can better explain why they may not sell.

V. Conclusion

The central empirical finding here is that agricultural yields are systematically higher for low-caste households residing in villages dominated by lower castes (BACs), in terms of total land ownership, compared to villages dominated by upper castes. The interpretation most consistent with the data seems to be that yields are crucially determined by access to groundwater irrigation, which is distributed through private markets. These markets seem to work more efficiently in villages where the caste composition is more homogeneous. We conjecture that bilateral contracts, which are more easily enforced between members of the same caste, explain why trade could break down across caste groups. This conjecture is in accord with anecdotal evidence in the area under study, where it is found that water transactions are strongly interpersonal. Individuals tend to conduct such trade with members of their own caste or close relatives. The presence of the upper castes, who own a substantial proportion of the private tubewells and pumps, therefore causes an inefficiency in the distribution of groundwater. This inefficiency is large and seems to be an example of social distance (the caste system) directly affecting the functioning of markets. Relative to more complex goods or services, these water trading relationships are still quite simple. They are bilateral agreements between two individuals

who reside within close proximity and likely have done so for generations. That trade can break down under these circumstances is striking, particularly as the gains from trade are demonstrated to be enormous for these very poor households.

REFERENCES

Akerlof, George A. 1976. "The Economics of Caste and of the Rat Race and Other Woeful Tales." *Quarterly Journal of Economics*, 90(4): 599–617.

Alesina, Alberto, and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature*, 43(3): 762–800.

Anderson, Siwan. 2009. "Caste as an Impediment to Trade." Unpublished.

Angrist, Joshua D., and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Banerjee, Abhijit, and Lakshmi Iyer. 2005. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review*, 95(4): 1190–1213.

Banerjee, Abhijit, and Rohini Somanathan. 2007. "The Political Economy of Public Goods: Some Evidence from India." *Journal of Development Economics*, 82(2): 287–314.

Banerjee, Abhijit, Lakshmi Iyer, and Rohini Somanathan. 2005. "History, Social Divisions, and Public Goods in Rural India." *Journal of the European Economic Association*, 3(2–3): 639–47.

Besley, Timothy, and Robin Burgess. 2000. "Land Reform, Poverty Reduction, and Growth: Evidence from India." *Quarterly Journal of Economics*, 115(2): 389–430.

Béteille, André. 1996. *Caste, Class and Power: Changing Patterns of Stratification in a Tanjore Village*. Berkeley: University of California Press.

Drèze, Jean, Peter Lanjouw, and Naresh Sharma. 1999. "Economic Development in Palanpur, 1957–93." In *Economic Development in Palanpur Over Five Decades*, ed. Peter Lanjouw and Nicholas Stern, 114–242. Oxford, UK: Oxford University Press.

Dubash, Navroz K. 2000. "Ecologically and Socially Embedded Exchange: 'Gujarat Model' of Water Markets." *Economic and Political Weekly*, 35(16): 1376–85.

Dumont, Louis. 1970. *Homo Hierarchicus: The Caste System and its Implications*. Chicago: University of Chicago Press.

Jacoby, Hanan G., and Ghazala Mansuri. 2008. "Land Tenancy and Non-contractible Investment in Rural Pakistan." *Review of Economic Studies*, 75(3): 763–88.

Jacoby, Hanan G., Rinku Murgai, and Saeed ur Rehman. 2004. "Monopoly Power and Distribution in Fragmented Markets: The Case of Groundwater." *Review of Economic Studies*, 71(3): 783–808.

Jha, Hetukar. 1991. *Social Structures of Indian Villages: A Study of Rural Bihar*. New Delhi: Sage Publications.

Kajisa, Kei, and Takeshi Sakurai. 2000. "Price Determination under Bilateral Bargaining with Multiple Modes of Contracts: A Study of Groundwater Markets in India." Foundation for Advanced Studies on International Development (FASID) Discussion Paper 2000-002.

Kishore, Avinash. 2004. "Understanding Agrarian Impasse in Bihar." *Economic and Political Weekly*, 39(31): 3484–91.

Kozel, Valerie, and Barbara Parker. 1999. "Poverty in Rural India: The Contribution of Qualitative Research in Poverty Analysis." http://siteresources.worldbank.org/INTPOVERTY/Resources/WDR/stiglitz/Kozel2.pdf.

Metcalf, Thomas R. 1979. *Land, Landlords, and the British Raj: Northern India in the Nineteenth Century*. Berkeley: University of California Press.

Mukherji, Aditi. 2004. "Groundwater Markets in Ganga-Meghna-Brahmaputra Basin: Theory and Evidence." *Economic and Political Weekly*, 39(31): 3514–20.

Munshi, Kaivan, and Mark Rosenzweig. 2005. "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth." Bureau for Research and Economic Analysis of Development (BREAD) Working Paper 097.

Munshi, Kaivan, and Mark Rosenzweig. 2008. "The Efficacy of Parochial Politics: Caste, Commitment, and Competence in Indian Local Governments." Bureau for Research and Economic Analysis of Development (BREAD) Working Paper 182.

Neale, Walter C. 1962. *Economic Change in Rural India; Land Tenure and Reform in Uttar Pradesh, 1800–1955*. New Haven: Yale University Press.

Pant, Niranjan. 2004. "Trends in Groundwater Irrigation in Eastern and Western UP." *Economic and Political Weekly*, 39(31): 3463–68.

Prasad, T. 1993. "Groundwater Development for Economic Emancipation in the Lower Ganges Basin: Problems, Prospects, and Strategies." In *Groundwater Irrigation and the Rural Poor: Options for Development in the Gangetic Basin*, ed. Friedrich Kahnert and Gilbert Levine, 139–46. Washington, DC: The World Bank.

Rawal, Vikas. 2002. "Non-Market Interventions in Water-Sharing: Case Studies from West Bengal, India." *Journal of Agrarian Change*, 2(4): 545–69.

Shah, Tushaar. 1993. *Water Markets and Irrigation Development: Political Economy and Practical Policy*. Bombay: Oxford University Press.

Shah, Tushaar, and Vishwa Ballabh. 1997. "Water Markets in North Bihar: Six Village Studies in Muzaffarpur District." *Economic and Political Weekly*, 32(52): A183–90.

Shankar, K. 1992. "Water Markets in Eastern UP." *Economic and Political Weekly*, 27(18): 931–33.

Singh, D. R., and R. P. Singh. 2003. "Groundwater Markets and the Issues of Equity and Reliability to Water Access: A Case of Western Uttar Pradesh." *Indian Journal of Agricultural Economics*, 58(1): 115–27.

Srinivas, M. N. 1955. "The Social System of a Mysore Village." In *Village India*, ed. M. Marriott. Chicago: University of Chicago Press.

Srinivas, M. N. 1987. *The Dominant Caste and Other Essays*. New Delhi: Oxford University Press.

Wood, G. 1999. "Private Provision after Public Neglect: Bending Irrigation Markets in North Bihar." *Development and Change*, 30(4): 775–94.