A Gender-Based Theory of the Origin of the Caste System of India*

Chris Bidner[†] Mukesh Eswaran[‡]

December 11, 2012

Abstract

This paper proposes a theory of the origins of India's caste system by explicitly recognizing the productivity of women in complementing their husbands' skills. We explain the emergence of caste and also the core features of the caste system: its hereditary nature, its insistence on endogamy (marriage only within castes), and its hierarchical character. We demonstrate why the caste system requires the oppression of women to be viable: punishments for violations of endogamy are more severe for women than for men. When there are such violations, our theory explains why hypergamy (women marrying up) is more acceptable than hypogamy (women marrying down). Our model also speaks to other aspects of caste, such as notions of purity, pollution, commensality restrictions, and arranged/child marriages. We also suggest what made India's caste system so unique and durable. Finally, our theory shows that, contrary to claims made by the most dominant anthropological theory, economic considerations were of utmost importance in the emergence of the caste system.

Keywords: caste, endogamy, gender
JEL Codes: D02, J12, J16, Z13, Z12

*We thank Kumar Eswaran, Vinayak Eswaran, Patrick Francois, Tarun Jain, Amartya Lahiri, and, especially, Li Hao for thoughtful comments. We also thank participants of the ThReD Workshop at Budapest (June 2012), the Economic Theory workshop at the University of British Columbia, and seminar participants at the University of New South Wales. We wish to acknowledge the financial supported received from research grants from the ASB (Bidner) and the SSHRC (Eswaran).

†University of New South Wales, School of Economics. c.bidner@unsw.edu.au
‡University of British Columbia, Department of Economics. eswaran@econ.ubc.ca

1 Introduction

The caste system of India, having lasted for around 3,500 years, has proved to be one of the most enduring of Indian institutions. There have been many social organizations elsewhere in the world that have had a de facto caste system in antiquity and also in more recent periods.[1] However, it is in India alone that the caste system seems to have developed to a point that it still manages to retain importance in the contemporary social organization of India. This peculiar Hindu institution has withstood the influence of Buddhism (which eschewed caste), a religion that spread to the far East but could not take permanent root in its own country of origin. It has withstood the Muslim invasion of India from the 13^{th} Century CE and the subsequent establishment of the Mughal empire. It has survived the colonization of India by the British for two hundred years. And now, despite India's rapid economic development in recent decades, it still insinuates itself into the social, economic, and political fabric of the country. No researcher with any familiarity with India would claim that caste is a thing of the past. Why did the caste sytem of India arise, and what made it so unique and so resilient? This paper offers a novel theory that explains its origin and the essential role that the oppression of women has played in its establishment. The gender effects of caste have eluded the attention of analysts[2] because they have ignored the productive role of women. By explicitly recognizing this, we resolve many puzzling aspects of the caste system and also explain the asymmetric treatment afforded to women within this institution.

Since no theory can hope to explain every nuance of the elaborate caste system of India, we focus on the following core features at its incipience[3]: (i) caste was heriditary and largely based on occupation, (ii) it was strictly endogamous (marriage occured only within castes), and (iii) there was usually a well-established hierarchy between castes. Our theory explains all three features.

At a broad aggregate level, there are four castes (called *varnas*). These are the Brahmins (priests), the Kshatriyas (the warriors), Vaishyas (the merchants), and the Shudras (the menial workers). There is one other aggregate (now called Dalits) that fell outside the caste system and its people were deemed 'untouchable'. Nowadays, there are thousands of sub-castes (called *jatis*) within each *varna*, and the caste system essentially operates at the level of the *jati*. For the purpose of theorizing, however, we shall focus on the *varnas* first—referring to these as the castes—and then explain the rise of sub-castes.

Historians of caste since the 19^{th} Century had long argued that the caste system arose after an Aryan invasion from the north-west around 1,500 BCE after which the victors imposed an oppressive system on the vanquished. This was a conjecture based on references in the *Rig Veda*, the earliest of Hindu scriptures, to an Aryan race. However, this claim has been largely

[1]The classic article by Davis (1941) has a discussion of caste in a non-Indian context.
[2]Chakravarti (1993, 2003) is among the few exceptions.
[3]We follow Bougle (1971) in focusing on these features.

discredited in recent decades. There is no archeological evidence of any such invasion; the Vedic culture, which started after 1,500 BCE and which spawned the caste system, seems to have been an indigenous innovation of an earlier culture at Harappa [Shaffer (1984), Shafer and Lichtenstein (2005)]. Recently, genetic evidence has also confirmed that there could not have been any large scale infusion of genes into India since 3,500 BCE [e.g. Sahoo et al (2006)]. Since both archaeological and genetic evidence firmly imply that the caste system of India was an entirely *indigenous* development—not one foisted by foreign invaders—it therefore has to be explained in these terms.

In anthropology, the dominant theory of caste that prevails is that of Dumont (1970). In his highly influential *Homo Hierarchicus*, he claims that the twin concepts of 'purity' and 'pollution' are key to the caste system and he takes this as the premise of his theory. To an economist, however, this is somewhat unsatisfactory for it leaves unexplained why these concepts should be taken as given. Rather, it would seem that their importance needs to be explained in terms of more compelling primitives. Likewise, sociological theories start from the premise that each caste comes with a different status, designed by one or more groups.[4] Our treatment of caste endogamy sharply contrasts with these approaches that rationalize endogamy as an attempt to preserve the 'purity' or 'status' of a caste. The proposition that endogamy acts to strengthen the caste system is tautological if we take endogamy to be a central feature of the caste system. Our theory explains the genesis of endogamy and demonstrates that the purity/pollution dichotomy and the notions of status that figure so prominently in discussions of caste are the *means* employed to establish the caste system, and are not its cause.

Among economists, attempts to explain the caste system are few and far between. In the first formal treatment of caste, Akerlof (1976) envisages that a caste 'code' such as hiring only within caste is supported in equilibrium by the threat of being made an out-caste if the code is violated. While such a construction may be supported in equilibrium, it does not offer a plausible explanation for the various aspects of caste cited above. The key 'caste rule' that is liable to be broken here is that of hiring someone from a 'lower' caste. In a model dealing with class (but not caste), Cole, Mailath, and Postelwaite (1998) generate classes endogenously in equilibrium and in which upper-class members, concerned about loss of status, might behave differently than lower class members. The approach is reminiscent of Akerlof's in that violations do not harm the group, yet punishments are enacted out of a concern of being punished oneself. Freitas (2006) views caste as arising in a scenario with one-sided moral hazard: sellers exert effort but buyers of a product may renege on payment. Castes based on occupations are formed, and buyer payment is induced by the threat that all caste members will deny future service. Here the punishment is *across* castes, and the 'code' liable to violation again relates to interactions in the sphere of production and exchange. The

[4] See, for example, Milner (1994).

group punishment mechanism is similar to that stressed by Greif (1993) in a different context. In Choy (2011), agents search for partners and play a repeated Prisoners' Dilemma, in which partnerships occasionally break up. Small groups are optimal because the consequence of cheating is worse, so that higher cooperation can be sustained. This speaks to intra-caste punishment, but again the 'code' liable to violation relates to interactions in the sphere of production and exchange. Groups are not allowed to mix, for a previous 'mixer' is punished because others believe that he is myopic and will therefore cheat in his repeated interactions.

Existing theories do not speak to three important features of the caste system, features that are of central concern in our theory: hierarchy, gender, and endogamy.[5] The productivity of women takes center-stage in our analysis; just like men acquire their skills from their fathers, women acquire theirs from their mothers. Because occupational activities of families in ancient India, though determined by that of the husband's father, are jointly executed by the spouses, the complementary skills a wife brings are important to family income. A wife from a family plying a different trade has less appropriate skills and this reduces the family income. Likewise, if a person cannot find a spouse, family income is reduced. If a member of a group conceives a preference for an outsider and marries (a 'love marriage'), why should the group object? After all, this can happen only if the member values this preference more than the fall in income due to skill mismatch between the spouses. There is, however, an externality associated with out-group marriages. When a member marries outside the group, a group member of the opposite sex will not find a spouse within the group. This is a cost that is not internalized, and if endogamy is to obtain out-group marriages must be prevented. We model deterrence by having non-violators of endogamy norms punish violators in subsequent in-group interactions. Such punishments, however, are costly to both the punished and the punisher, and so endogamy may not be sustainable. When endogamy can be sustained, the occupational groups become castes.

The regulations of the caste system rationalize the various caste norms and the attendant punishments for violations. We show that as many groups will form castes as benefit from husband-wife skill complementarities even after incurring the cost of enforcing endogamy. Our theory has the advantage that it does not require one group (like the Brahmins) to impose the caste system on other groups. Although the Brahmins have undoubtedly had a strong influence in shaping the caste system, it is too much to believe that a small group could unilaterally impose such a system on the rest of the population. (This is the presumption of sociological and anthropological theories of caste.) We show, in fact, that it is only people in those occupations that have no overwhelming reason for endogamy that may be hurt by the formation of other castes. They would see a *de facto* restriction on their choice of marriage

[5]The importance of endogamy to caste was emphasized nearly a century ago in a prescient article by Ambedkar (1917).

partners without a compensating benefit.[6]

The benefit of endogamy is determined by the skill complementarity dictated by the technology of the husband's occupation. Groups that perceive the greatest spousal complementarities are the ones that are most zealous of forming castes and would be willing to impose the harshest punishments for violations of endogamy. These punishments, we show, are *gender dependent*: at least in the groups with strong complementarities, women have greater incentives to out-marry and, therefore, face greater punishments. This gender asymmetry would be enhanced if in our formal model we include (which we do not) the fact that, in patriarchal societies the perceived social cost of not finding a spouse is greater for men than for women. If violations to endogamy were to occur, we further show that the punishments inflicted on women are asymmetric in terms of which castes they out-marry into: marrying into a caste with greater skill complementarity between spouses invites a lower punishment than when marrying into castes that exhibit lower complementarities. Whether or not a caste would permit a woman to marry into another becomes the measure of the relative status of the two castes, and this status is endogenously determined here. When the degrees of skill complementarities can be ranked across castes, this translates into a hierarchy in castes—giving meaning to the terms 'marrying up' and 'marrying down'. In effect, we explain why women marrying 'up' (hypergamy) is considered more acceptable to a caste than its women marrying 'down' (hypogamy). To our knowledge, ours is the first paper in economics to explain caste endogamy and caste hierarchy from first principles. In the light of our model, we can also explain many other features of the caste system: arranged marriages, child marriages, the ban on widow (but not widower) remarriage, and the obssessive preoccupation with notions of purity and pollution.

This paper also contains a brief treatment of the role of social learning and the durability of skills in the intergenerational transmission of occupational human capital. For those groups to which social learning is important and skills could be passed on intact to their progeny, out-marrying is accompanied by another externality (an intergenerational one). By preventing the dilution of skills and the stock of knowledge over time, endogamy confers even greater benefits in this case. This observation offers a clue to why India was unique in the extent to which its social stratification got so rigidly embedded. Our discussion in this paper outlining the elaborate techniques developed by the Brahmins of India to transmit their stock of knowledge *without the benefit of written script* for more than a millennium is suggestive of why endogamy was so important to them. If this view is correct, it suggests that, whatever the uses and abuses of the caste system were over its long and infamous history, its *origin* may have been more benign—with no caste group having an ulterior motive of exploiting or suppressing other caste groups of society. In this sense, Dumont (1970) may have been right in his claim that power

[6] This is not to suggest that, *after* the caste system was established and a hierarchy has emerged, the upper castes did not exploit their position in the hierarchy.

was a secondary factor in the formation of castes. We do not, of course, agree with his more general claim that *all* political economy considerations are secondary; in our theory of caste, economics is fundamental. Our claim that the origin of caste may have been benign needs to be heavily qualified, however, for its implementation exacted a disproportionately high price from half the members within each caste, namely, women.

Recent literature reveals the continuing economic salience of caste in India. Anderson (2011) has shown that the incomes of low caste individuals are lower in villages dominated by high castes than those dominated by low castes because of a breakdown of trade in water for cultivation. Deshpande (2011) documents the pervasiveness of caste discrimination in contemporary India. Pande (2003) has shown theoretically and empirically how the contitutional mandate to help disadvantaged castes in India through assured representation can and has managed to increase transfers to these groups. Banerjee and Somanthan (2007) have shown that access to public goods are dependent on caste status, though political mobilization of some of the lower castes are improving matters. Recently, Hnatkovska, Lahiri, and Paul (2012) have shown that the conditions of the historically disadvantaged caste groups have been converging to that of others. Munshi and Rosenzweig (2006) show that, though low caste boys have been tardy, low caste girls in Mumbai have been switching to English language education in response to globalization. Munshi (2011) has shown both theoretically and empirically how, under suitable conditions, caste networks can actually help people move out of poverty. The work of Esteban, Mayoral, and Ray (2012) shows how caste may incite conflicts, especially when it contributes to fractionalization and polarization. In view of the pervasiveness of the economic effects of caste even today, our paper takes a step back to investigate why this institution came into being and how.

2 The Model

2.1 Fundamentals

Consider a population partitioned into C groups, where group $c \in \{1, ..., C\}$. The unit of analysis in our model is the household. Each household has one offspring, either male or female. The utility of each member of a household is the sum of the payoffs of the parents and offspring in their various interactions. We assume that group c contains N^c male offspring and N^c female. Presuming equal numbers of males and females is a reasonable assumption to invoke in a model, but the consequences of a biased sex-ratio can be intuitively discerned after we spell out the model. The households participate in the marriage market (unless the offspring have found a spouse before doing that) and then enter into a series of periodic interactions with other group households. These components of the setting are considered in turn.

2.1.1 Production and Marriage

The husband plies a trade inherited from his father; his skills are also inherited. The wife inherits her skills from her mother. After marriage, the wife joins her husband's household and the couple engages in joint production in the occupation of the husband. We should be explicit that caste is tied to occupation in our analysis because we are modeling a period of India's history when this was the case. (In a contemporary setting, of course, the link between caste and occupation is tenuous.) That caste-based work requires skill on the part of the husband has been always understood because the occupation of the husband determines the couple's livelihood. What has been missing in theoretical formulations of caste, however, is explicit recognition of the simple fact that wives also require suitable skills because they, too, contribute to the caste-activity. Dube (1996) brings home the importance of women in executing the traditional caste-based work in their husband's family.[7] David and Kramer (2001) show how women in the caste of potters are involved in every stage of pottery production helping their husbands. In fact, even when they do not engage in certain caste activities reserved for males, mothers give detailed instructions regarding these to their young sons and make corrections [Kramer (1997, pp. 47-49)]. The presence of Brahmin women—although usually forbidden from performing sacrificial rituals—perform many other rituals to aid their husbands and, indeed, their presence is essential [Jamison (1996)].[8]

To capture this economic contribution of wives, we posit that when both members of a couple hail from families that practice the same trade, the skills of the spouses are complementary. Marrying someone from one's own occupational group enables the production of an output Y that is jointly consumed by the couple.[9] However, when a male marries outside his group the couple's output falls to $\alpha_c Y$, where the parameter $\alpha_c \in [0, 1]$ is an inverse measure of the male-female complementarity in the activity of group c. This implies that when a female of group c' marries outside their group, the couple's output falls from Y to $\alpha_c Y$, where c corresponds to her husband's group. The parameter α_c determines how exacting the required skills are in the occupation of caste c, being smaller for occupations that require highly specific matches.[10] We are assuming Y to be the same for all occupations. Other than affording simplicity, this enables us to separate caste from class considerations in our theory;

[7] For instance, she writes "The necessity of continuing with occupational work is an important basis for marrying within caste. It is understandable then that a landowning cultivator family of the Kunbi caste in rural Maharashtra should be unhappy when one of its sons, after acquiring education, decides to marry an educated Brahmin girl. Even home-based work linked to cultivation is seen as outside the arena of her experience and below her status." [Dube (1996), pp. 4-5].

[8] "The wife is the back half of the ritual," she says, quoting from a classical exegetical Hindu text [Jamison (1996, p. 30)].

[9] Assuming Y to be a household public good in this manner is merely a simplification that sidesteps any bargaining issues between the couple, a consideration that is not the focus here.

[10] In principle, the parameter α could also depend on the group c' that the wife comes from, but we abstract from that possibility here to keep the model tractable.

our purpose is to isolate the former.

The choice of marriage partners is subject to a stochastic preference shock that generates a (cardinal) utility b, where b is a random variable distributed according to G. This allows for the possibility that individuals, perhaps by falling in love or by perception of a mutual temperamental compatibility, may choose to marry someone outside their group.[11] To model the arrival of this shock in the simplest possible way, suppose that the search for partners unfolds according to the following timing:

1. *Pre-market*: Denote by μ^{in} the probability of an individual encountering another of the opposite sex from the same group and by μ^{out} the probability, conditional on not meeting someone within the group, of encountering an individual of the opposite sex from a different group. We take these probabilities as exogenous here, but we shall discuss their endogeneity subsequently.
   (a) If someone is encountered, the pair then observes their preference shock, b. After observing b, both individuals decide whether to marry.
      - If the individuals belong to the same group and the couple marries, each member of the couple receives a marriage payoff (utility) of $Y + b$.
      - If the individuals belong to different groups and the couple marries, each member of the couple receives a payoff of $\alpha_c Y + b$, where α_c is the inverse complementarity associated with the husband's group (since he determines the occupation of the household).
      - If the individuals do not marry, each proceeds to the marriage market.

   (b) If no encounter occurs, the individual proceeds to the marriage market.
2. *Marriage Market*: The marriage market facilitates as many in-group marriages as possible.
   (a) An individual who finds an in-group partner will receive a payoff of Y.
   (b) Those who do not find an in-group partner receive a payoff y that is less than Y.[12] This may reflect their next best option in the marriage market (such as marrying someone from a different group who has met a similar fate, or perhaps the payoff from waiting until a future marriage market or a distant marriage market (neither of which is modeled here).

In summary, depending on the outcome, each individual gets a marriage-related payoff of $Y + b$ (for within group love marriage), or $\alpha_c Y + b$ (for out-group love marriage), or y (when

[11] We restrict ourselves to heterosexual alliances since endogamy is precisely about these.

[12] In general, y could depend on the gender $g \in \{m, f\}$ of the person. Furthermore, we might expect that y also depends on the extent of the complementarity in production, with occupations of high complementarity having lower y because the consequences to not having a spouse are worse. But we abstract from these ramifications in the formal model.

unmarried).[13] By construction, parents value these payoffs to their offspring precisely as the latter do. Let individual (and household) i's realization of this payoff be denoted U_i^{marriage}.

2.1.2 Interactions

The parents in our model have economic interactions amongst themselves, a phase into which they move after their offspring's matching stage. The setting of these interactions provides groups with the facility to enforce endogamy in the marriage market, if that is the agreed-upon social norm. Parents of violators of the norm are punished through these interactions, which we model as an infinitely repeated game. In this phase, households within a group are randomly paired each period and play the following stage game. Parents in household i (jointly) take action a_i at a strictly convex cost of $c(a_i)$, with $c'(0) < 1$. We posit that $\min\{a_i, a_{-i}\}$ units of a local public good are produced and jointly consumed by the pair of households, where $-i$ denotes the household other than i in the pair. (The public good aspect of the output is purely for algebraic convenience; nothing hinges on it.) The stage payoff to i is $u(a_i, a_{-i}) \equiv \min\{a_i, a_{-i}\} - c(a_i)$.

The efficient action in the stage game is $a_i^* = a_{-i}^* = a^* \in \arg\max_a 2(a - c(a))$. The advantage of this formulation is that any $a \in [0, a^*]$ is a Nash equilibrium of the stage game. Therefore we can support asymmetric equilibria in which there are two subgroups of players, the punishers and the punished, where the punishers are norm-respecters and the punished are norm-violators. The punishments, as we shall soon see, will depend on the gender of the violator. The punishers of group (or caste) c play $\hat{a}_{gc}$ if playing against a punished household with violator of gender g and play a^* otherwise, and the punished always play $\hat{a}_{gc} \in [0, a^*]$. So in every period the punished always confront players who play $\hat{a}_{gc}$, whereas punishers either play $\hat{a}_{gc}$ or a^* depending on whether they are confronting the punished or other punishers. The key is that it is costly to be classified as punished, but it is also costly for the punishers to inflict this punishment. The threat of acquiring "punished" status potentially dissuades deviations from endogamy if the punishment is severe enough. An advantage of this construction is that the extent of punishment can be continuously chosen with a trade-off in mind, namely, higher punishments are better able to dissuade norm violations but are also more costly to impose.

Let $u^* \equiv \sum_{t=1}^{\infty} \beta^t u(a^*, a^*)$ be the discounted payoff if the highest-payoff Nash equilibrium is played each period, where β is the common discount factor for all people. Let $\hat{u}_{gc} \equiv \sum_{t=1}^{\infty} \beta^t u(\hat{a}_{gc}, \hat{a}_{gc})$ be the discounted payoff to a punished household with violator of gender g in group c. The difference $\phi_{gc} \equiv u^* - \hat{u}_{gc}$ may be construed as the severity of punishment on households with violators of gender g in group c.

If $\hat{N}_g^c$ of the N^c members of gender g in group c are to be punished, then $z_{gc} \equiv \hat{N}_g^c/(N^c - 1)$ is the probability that a "punisher" encounters a "punished". Essentially, a household receives

[13] Note that we interpret Y and y in present value terms.

an interaction-related expected discounted payoff of either $u^{*} - z_{gc}\phi_{gc}$ (as a norm-conformer) or $u^{*} - \phi_{gc}$(as a norm-violator). (Note that even a conformer's payoff is determined by the proportion of violators because punishment is costly to inflict.) Let household i's realization of this payoff be denoted $U_i^{\text{interaction}}$, so that the total payoff to every member of the household is $U_i \equiv U_i^{\text{marriage}} + U_i^{\text{interaction}}$.

2.2 Discrete Shocks

We begin with the most transparent setting. We assume that preference shocks are discrete: $b \in \{0, B\}$, with p as the probability that $b = B$. We assume that B is small enough that we have

$$B + \alpha_c Y + y < 2Y, \tag{1}$$

for some group $c \in \{1, ..., C\}$. In other words, we assume the above inequality holds for the group with the highest degree of complementarity (the smallest α_c). The right hand side is the total payoff of an in-group couple's households in the absence of a preference shock. The left hand side is the sum of the payoff to person who has received (and has embraced) an out-group shock and a person in the group left without a spouse as a consequence. The assumption above implies that there is an externality associated with marrying outside the group when all others are marrying within the group, in the sense that the total payoff when a pair from the group marry is larger than the total payoff when one person marries outside the group. If this did not hold, it would never be optimal to interfere with marriage decisions (at least on the basis of expected payoff).

In this sub-section, we deal with the case in which the positive shock is large in the sense that all individuals who receive an out-group shock would prefer to marry outside the group when the punishment is zero. We assume here that

$$B \geq (1 - \alpha_c)\, Y, \tag{2}$$

for all $c \in \{1, ..., C\}$. The left hand side is the preference-related benefit to an individual of embracing an out-group spouse and the right hand side is the income penalty borne due to the mismatched spouses in production. This penalty is naturally the largest for groups with the most complementary spousal inputs. In effect, we are assuming that the above inequality holds for the group with the highest degree of complementarity (the smallest α_c).

It would never be optimal, of course, to punish people if it did not alter behavior (in terms of dissuading marriage outside the group). Given the binary nature of the shock, and the above assumption on B, it is therefore either optimal to have zero punishments or to have punishments sufficiently high to discourage deviation by all (an endogamy norm).[14] Notice

[14]To be sure, our assumption rules out the possibility that it is optimal to punish the household of one

that punishments will not actually have to be inflicted in either case—the benefit of positive punishments is to ensure more even marriage market participation, while the cost is the forgone positive shocks. We now describe the relative payoffs.

For notational simplicity, we focus on one group. Note that marriage will always occur between in-group members who receive the positive shock, regardless of punishment levels. With probability $\tilde{\pi}(N) = \text{Bin}(N \mid N^c, 1 - \mu^{\text{in}} \cdot p)$ exactly N of the N^c males and N of the N^c females do not receive the positive shock with an in-group member (the equality across genders follows from the fact that in-group marriages occur in pairs). We now analyze the payoffs arising when there are exactly N males and females within the group that are liable to encounter a preference shock with an out-group member prior to the marriage market.

We begin with the case in which there is no punishment. Suppose that n_m of the N males and n_f of the N females do not marry in the first stage and therefore end up in the marriage market. The number of in-group alliances that will be formed is the smaller of n_m and n_f. The aggregate payoff to households with offspring of gender $g \in \{m, f\}$ for a particular realization of the pair (n_m, n_f) is:

$$\begin{aligned} U_g(n_m, n_f \mid N) = (N^c - N)(Y + B + u^*) + (N - n_g)\left(\alpha_g Y + B + u^*\right) \\ + \min(n_m, n_f)(Y + u^*) + \left[n_g - \min(n_m, n_f)\right](y + u^*), \end{aligned} \tag{3}$$

where $\alpha_m = \alpha_c$ and $\alpha_f = \bar{\alpha}_{-c}$ (the average α among groups that females out-marry into). The first term on the right side is the total payoff to those who receive in-group preference shocks (which they embrace); the second term is that to those with an out-group alliance in the pre-market stage; the third term is that to people who find matches in the (same-group) marriage market; and the last term is the total payoff to those who remain single.

The aggregate Benthamite payoff of all members in the group is

$$U(n_m, n_f \mid N) \equiv U_m(n_m, n_f \mid N) + U_f(n_m, n_f \mid N). \tag{4}$$

The above payoffs, of course, are conditional on the realization of (n_m, n_f). The probability of drawing the pair (n_m, n_f) from N males and N females is given by the product of two binomial random variables:

$$\pi(n_m, n_f \mid N) = \text{Bin}(n_m \mid N, 1 - q) \cdot \text{Bin}(n_f \mid N, 1 - q),$$

where

$$q \equiv p\, \mu^{\text{out}}$$

is the probability of receiving the positive preference shock with an out-group member condi-

gender sufficiently highly while not punishing the other.

tional on not receiving the shock with an in-group member.

The expected aggregate payoff, $U^{NP}(N)$, of the N males and N females in the group in the *no punishment* case is then

$$U^{NP}(N) \equiv \sum_{m=0}^{N} \sum_{f=0}^{N} \pi(n_m, n_f)\, U(n_m, n_f). \tag{5}$$

The aggregate payoff, $U^E(N)$, of the N males and N females in the group when an *endogamy* norm is enforced with full punishment[15] is simply:

$$U^E(N) = 2(N^c - N)(Y + B + u^*) + 2N\,(Y + u^*). \tag{6}$$

Thus, conditional on there being N males and N females liable to encounter preference shocks with out-group members, the net payoff to the group of endogamy, as perceived by a Benthamite planner, is the difference $U^E(N) - U^{NP}(N)$. The unconditional net benefit to endogamy is therefore $\sum_{N=1}^{N^c} \tilde{\pi}(N) \cdot [U^E(N) - U^{NP}(N)]$.

Proposition 1: The net benefit to endogamy is (a) increasing in Y, (b) decreasing in y, (c) decreasing in B, and (d) decreasing in α_m and α_f.

This proposition gives some sense of the groups with the largest incentives to adopt endogamy norms. In understanding the proposition, it is sufficient to focus on the net benefit to endogamy for a given value of N - i.e. the difference between $U^E(N)$ from (6) and $U^{NP}(N)$ from (5) - since the parameters do not affect $\tilde{\pi}(N)$. We begin by noting that out-marrying carries an income penalty due to skill mismatch and that this is internalized by the couple. However, there is also the externality associated with marrying an outsider in that it inflicts an income penalty of $(Y - y)$ on a group member who cannot find a spouse as a result. So the net benefit of endogamy increases in Y and decreases in y, explaining parts (a) and (b) of the proposition. This net benefit is naturally decreasing in B, by part (c), because endogamy forces people to abandon their out-group preference shock. Part (d) brings home the importance of spousal complementarity in production. Groups with a high degree of spousal complementarity in production (low α_m) will tend to favor endogamy norms because of the greater income loss due to mismatch. Individuals who out-marry internalize the attendant income penalty but they do so because they are more than compensated by the preference shock B. Nevertheless, since the surplus from a given B is increasing in α_m and α_f, the net benefit to endogamy decreases in the latter. This is part (d) of the above proposition.

It may seem at first blush that endogamy may not necessarily be most preferred in groups

[15] Note that, as long as the punishment is severe enough, it never has to be inflicted in equilibrium because all violations are prevented.

of high spousal complementarity because groups with a low α_m will tend to have $\alpha_f > \alpha_m$ (since females from relatively low α groups will tend to encounter males from relatively high α groups on average). In other words, even if endogamy is preferred by the (households of) males of these groups it may not be preferred by the (households of) females, and a planner who weights them equally (as does the Benthamite planner we are invoking while aggregating payoffs) may not deem endogamy optimal. However, if groups are of the same size, then since males and females enter symmetrically in this case, it is the *average* value of α that matters. And the average value of $\alpha_m + \alpha_f$ will be increasing in α_m.[16] We can then be sure that groups with a high degree of gender complementarity are the most likely to favor endogamy. In the case with just two equal-sized groups the value of $\alpha_m + \alpha_f$ must be the same across the groups, and both are equally likely (or unlikely) to prefer endogamy. If the income y of being without a spouse is lower in the high-complementarity group (which would make sense precisely because of the greater need for a spouse in production), then this group will have a stronger preference for endogamy even when there are only two groups.

The following result, the proof of which is in the Appendix, provides a sufficient condition for when endogamy is definitely preferred.

Proposition 2: Endogamy with full punishment is preferred to no punishment if the probability of an individual receiving an out-group shock conditional on not having obtained an in-group one, μ^{out}, is sufficiently small (but positive).

The case for endogamy hinges on externalities. The externality associated with an individual embracing an out-group shock is that there is the possibility of a group member of the opposite sex not finding a spouse as a result and thereby receiving a reduced payoff. However, if equal numbers of males and females embrace out-group shocks, there is no such externality and insisting on endogamy would needlessly force some people to reliquish their preferred out-group partners of choice. Thus the cost of the externality associated with out-group marriages depends on there being a *difference* in the numbers of males and females opting to marry outside the group. Even though the expected value of this difference is zero (given that q is the same for males and females), this is not the relevant entity of interest. Rather, it is the expected value of the *absolute* difference that matters because the externality is incurred regardless of whether it is males or females who are on the long side of the marriage market.

To begin, consider the extreme case whereby out-group members are never encountered: $\mu^{\text{out}} = 0$. In this case all those that do not receive a positive preference shock with an in-group member are certain to find an in-group member in the marriage market nonetheless. There

[16]Proof: When groups are of equal size, the average value of α_f for females from group c who outmarry is $\alpha_f = \left(\sum_{k\neq 1}^{C} \alpha_k\right)/(C-1)$. Therefore $\alpha_m + \alpha_f = \alpha_c + \left(\sum_{k=1}^{C}(\alpha_k - \alpha_c)\right)/(C-1)$
$= \alpha_c(C-2)/(C-1) + \left(\sum_{k=1}^{C} \alpha_k\right)/(C-1)$, which is increasing in α_c.

is no particular benefit to endogamy in this case - in fact, the net benefit to endogamy is exactly zero since out-group marriages never occur. As μ^{out} is increased, the possibility of their being a mismatch in the aggregate numbers of males and females receiving out-group shocks emerges. Specifically, for small μ^{out} the set of possible outcomes is dominated by three events: an out-group shock is received by (i) no male and no female, (ii) exactly one male but no female, and (iii) no male but exactly one female. As discussed above, the net benefit to endogamy is zero in case (i). However, in cases (ii) and (iii) there is a positive net benefit to endogamy due to the externality identified above.

We may note that endogamy would never be preferred if all groups in society had identical skills; heterogeneity is essential for this ensures $y < Y$. If all groups were perfectly symmetrical in abilities, upon not finding a within-group spouse a person could marry, with no income loss, an out-group individual who has met a similar fate. Then there would be no gains to enforcing endogamy. For our model to be meaningful, we require hereogeneity. This is not an issue, however, in any real-world setting where different occupations require different skills.

There are at least two interesting and important implications that follow from Proposition 2. First, endogamous groups can form *sequentially.* As one group adopts endogamy norms, this effectively decreases μ^{out} since encountering a member of an endogamous group is equivalent to not encountering them at all (since they will never agree to marriage, even if the positive shock is received). This is especially so when the endogamous group puts in place various restrictions on the movements and actions of its members (we shall discuss some of these later in the paper). In accordance with Proposition 2, the subsequent fall in μ^{out} for other groups would encourage them to adopt endogamy norms, and so the process of caste formation can cascade. This is consistent with the evidence, such as there is, on the formation of castes. The oldest reference to caste is in the Rig Veda, a Hindu scripture. This Veda mentions only two castes (Brahmins and Kshatriyas) at the beginning (composed roughly at 1,500 BCE) and by the end (composed some 500 years or so later), some hymns mention three castes (Brahmins, Khatriyas, and Vaishyas). The fourth group of Shudras was a later addition to the caste system after 1,000 BCE. It seems to be the case that the Brahmins were the first endogamous group [Le Bon (2005)]; they separated out from the warriors (Kshatriyas), who subsequently separated from the agriculturists.

The second important implication of the above proposition is that the caste system may be incomplete in the sense that there may exist a group for which there may be no serious spousal complementarity in production and it would rather not circumscribe personal choice of marriage partners. Nevertheless, this group may be hurt by the formation of other castes because endogamy is *de facto* imposed on them (as no other group will out-marry with them). The group without a caste (the 'untouchables', now referred to as Dalits) is a case in point. The Dalits do not accept the validity of the caste system and they see it as a monstrous institution imposed on them without their consent.[17] The fact that individual groups found it beneficial to cordon themselves off from the rest of the population in terms of marriage certainly does not suggest that the collective benefitted as a whole even during the inception of the caste system.

It may be noted in passing that another necessary condition for endogamy to be preferred is that the size of the preference shock not be too large. If the shock is very large, a group's expected aggregate payoff will clearly be maximized if no contraints are put on personal choices.

Asymmetries by Gender and Caste

If endogamy is optimal, how do punishments vary across genders and across caste groups? These are the issues we now turn to. To dissuade a deviation we need the payoff from conformity to be at least as large as that from violation: $Y + u^* \geq \alpha_g Y + B + u^* - \phi_g$. That is, the punishment must be at least as large as ϕ_g^*, the minimal required punishment:

$$\phi_g \geq \phi_g^* \equiv B - (1 - \alpha_g)Y,$$

where we must remember that when $g = f$, the parameter α_f denotes the (inverse) degree of complementarity of the group the woman has married into. The term $B - (1 - \alpha_g)Y$ is the benefit conferred by the preference shock less the income penalty suffered due to mismatch of the spouses in production. Since this entity is a measure of the temptation to violate endogamy norms, any punishment that would dissuade violations must be at least as large as this.

Proposition 3: The minimal required punishment to deter violations of endogamy by
(i) males, ϕ_m^*, is independent of which group their wives come from, and
(ii) by females, ϕ_f^*, is higher when they marry into a group with lower gender complementarity (higher α).

Since a couple's occupation is determined by the husband's, no matter which group his wife is drawn from it is the technology of his occupation that dictates the degree of complementarity. If the occupations are sufficiently different (as we are assuming), the training of women drawn from all possible outgroups is equally irrelevant. The fall in output due to spousal mismatch in production is the same for all out-group wives and is determined only by α_c, where c is the husband's group. This explains part (i) of the above proposition. The lower the cost of mismatch between spousal capacities, the greater is the incentive for a woman to violate endogamy norms when she receives a out-group preference shock. She has less incentive to

[17] While there is a vast literature on this, Ambedkar (1917) offers one of the most compelling critiques of the caste system from the standpoint of the Dalits.

embrace an out-group shock with a man whose occupation requires a high level of complementarity. So the punishment required to dissuade violations by women have to be increasing in α_c, where c is the intended husband's group. This explains part (*ii*) of the proposition. Since it is the groups with the highest complementarities that are most interested endogamy and because it is women from precisely these groups who are most tempted to out-marry, these are the groups that inflict the severerest punishments for violations.

It might be objected that, in an inter-caste marriage, the wife's parents' occupation may be different but perhaps not very different from the husband's. So how can we assume that wives drawn from all occupations will lower output by the same amount? The distinction being made here (of slight differences) is between the occupations of *sub-castes.* The generation of sub-castes and when they arise will be discussed subsequently. Here we are addressing the formation of the *varna* system, and the activities of the different *varnas* are substantially different in nature.

A very important implication of the above proposition is that it proposes a hierarchy of groups based on the degree of spousal complementarity in production. It is well recognized that women are punished more when marrying 'down' (called 'hypogamy') than for marrying 'up' (called 'hypergamy'). The terms 'up' and 'down' suggest a hierarchy. We claim that 'down' corresponds to lower complementarity (higher α) and 'up' to greater complementarity (lower α). Consider four broad castes labeled $1, 2, ...4$, with $\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$. The logic of Proposition 3 informs us that members of Caste 2 will punish more severely its women who marry into Caste 3 than those who marry into Caste 1. Likewise, members of Caste 3 will punish more severely its women who marry into Caste 4 than those who marry into Castes 1 and 2. So, if norm violation is to occur for women, Castes 2 and 3 find marriage into Caste 1 less objectionable than the alternatives. Caste 3 finds marriage of their women into Castes 1 and 2 less objectionable than the alternative (which is marriage into Caste 4). Thus, based on the treatment of women who violate endogamy, a hierarchy will be established. In this hierarchy, Caste 1 will be at the top, Caste 2 will follow, then Caste 3, and finally Caste 4. If the four castes are identified as Brahmins, Kshatriyas, Vaishyas, and Shudras, this hierarchy uncannily resembles the dictum laid down by law-giver Manu [Buhler (1886, IV, 12-13)]:

> A Shudra woman alone (can be) the wife of a Shudra, she and one of his caste (the wives) of a Vaishya, those two and one of his own caste (the wives) of a Kshatriya, those three and one of his own caste (the wives) of a Brahmana.

In other words, women can marry 'up' but not 'down', and the dictum lays down the precise hierarchy of the castes. The point to note here is that it is not caste hierarchy that determines whom women can marry; the causation is the reverse.

Our theory of hierarchy in castes hinges on the the degree of complementarity of the spousal inputs into production: Caste 1 exhibits the greatest complementarity and Caste 4 the least.

Arguably, this complementarity is reflected in the extent of the specialization in the tasks of the husband and the wife. This specialization is rather extreme among the Brahmins and the Kshatriyas. Among the Brahmins, only men can be priests who specialize in exacting rituals but women aid them in essential ways (see Section 3 for details). Among the Kshatriyas, usually only men are warriors but since there is no scriptural injunction against women taking up arms, there is a little more substitutability possible than amongst Brahmins. The greatest substitutability is in tasks entailing the least amount of human capital—those that are menial, occupations that comprise the Shudra caste. It is important to realize that, though we have opted to model the husband-wife complementarity in production only, this complementarity is also important in the *raising of children*. Raising children to follow in a particular tradition requires intimate knowledge of the procedures followed in that tradition and this is imbibed from childhood.

Apart from gender asymmetry exhibited in hypergamy and hypogamy, there is a more pernicious gender asymmetry in that women are punished more harshly than men in the group with the highest complementarity (i.e. the group that is most likely to favor endogamy).

Proposition 4: In the group with the most spousal complementarity, females face higher punishments than males.

Following the labeling used above, the group with the highest complementarity is the one that is uppermost in the caste hierarchy and has a complementarity parameter α_1. Thus the minimal punishment required to dissuade its males from intercaste marriages is $B-(1-\alpha_1)Y$. For its females, however, the minimal punishment needed to thwart deviation is the smallest value of $B-(1-\alpha_f)Y$, for $\alpha_f \in \{\alpha_2, \alpha_3, \alpha_4\}$. Since $\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$, it follows that the smallest punishment that women face is $B-(1-\alpha_2)Y$, which is larger than $B-(1-\alpha_1)Y$.

This proposition provides one rationale for gender asymmetry in punishments that is frequently observed, at least for women of the uppermost castes, and also ordained in the sacred Hindu texts [Chakravarti (1993)]. There is another reason that is at least as compelling which we have not incorporated here but the validity of which can be easily discerned in the light of our model. The key reason for endogamy norms of the caste system here is the externality that out-group alliances inflict on the group, which results in a mismatch between the number of marriageable males and females in the group. Now in a patriarchal society, the perceived cost to a male member of not leaving behind progeny is deemed much higher than that of a female remaining without children. There is evidence to suggest that patriarchy was in place in the pastoral and agricultural economy of India well before the caste system evolved [Chakravarti (2003, Ch. 4)].[18] Therefore, a social planner would deem the external effects of a woman's

[18] Holden and Mace (2003) have shown that the move from horticulture to pastoralism led to the switch from matrilineality to patrilineality among the Bantu speaking regions in Africa. Recently, Alesina, Giuliano,

defection to be much more deleterious to the group than a man's, and that the punishment needed to thwart women's violations have to be even more severe. The severity of the punishments meted out to women perhaps has at least as much to do with patriarchal values as with the incentives confronting women, but the two effects work in the same direction, namely, making punishments harsher on women than on men.

Note that we cannot assert that the result in Proposition 4 holds for all other castes. For example, to dissuade a Caste 3 female from marrying into Caste 1 it will require a minimal punishment of $B-(1-\alpha_1)Y$, which is smaller than the minimal punishment of $B-(1-\alpha_3)Y$ that would dissuade a male of Caste 3 from an inter-caste marriage. However, if we had added a relatively greater cost to males for not leaving behind progeny that we have alluded to above, the result in Proposition 4 would likely hold for the lower castes, too.

The advantage of endogamy identified here is a lower bound if we think of the planner as also caring about dynamic skill spill-overs. That is, the planner would prefer endogamy even more if they are concerned about avoiding skill dilution as a result of out-group marriages. We formalize this argument in Sub-section 2.4 below for it is very relevant to understanding the Brahmin insistence on endogamy.

The benefits of endogamy in the simplified discrete-shock model disappear as the population becomes infinitely large. Intuitively, a "law of large numbers" comes into play whereby individuals are almost surely able to find an in-group partner in the marriage market as long as the *proportion* of males and females out-marrying is the same (as it is in this case, given by q). There is thus a cost but no benefit to endogamy in the limit. This occurs because we have assumed the discrete shock is large enough that it is always embraced in the absence of punishments; there is no incentive-dependent margin of the preference shock below which an individual will not violate endogamy. This is not so, however, when the preference shock is a continuous variable. In this case, groups with high complementarity find that their males are less likely to agree to an out-marriage than the females (who face lower complementarity partners on average). In this case, when punishment levels are equal, there is an expected mismatch in the marriage market even as the population size grows to infinity. We now turn to this version of the model.

2.3 Continuous Shocks

In this sub-section we allow the preference shock b to be continuously distributed according to a distribution function $G(.)$. We assume a large population comprising two groups, labeled 1 and 2, with the former exhibiting greater complementarity in production (that is, $\alpha_1 < \alpha_2$). As before, recall that a group member encounters another group member with probability μ^{in}

and Nunn (2011) have shown that geographical regions around the world which historically used the plough in agriculture promoted patriarchal values that are visible even today.

and, conditional on not having had such an encounter, an out-group member with probability μ^{out}.

Encounters between group members always end in marriage. Encounters with out-group members, however, end in marriage if the shock is large enough: $b \geq \bar{b}_g$, where the critical minimum is depends on the gender g. (This is because, as we have already seen, the incentives to marry outside the group are gender-dependent and so are the punishments required to enforce endogamy.) The event $b \geq \bar{b}_g$ occurs with probability $1 - G(\bar{b}_g)$. Let q_g be the probability that a group member of gender g marries outside the group conditional on not encountering an in-group member. This is given by expression

$$q_g \equiv \mu^{\text{out}}(1 - G(\bar{b}_g)).$$

Simple rearrangement of the above definition allows us to express the magnitude of the cut-off shock $\bar{b}_g$ as a function of the probability q_g:

$$\bar{b}_g(q_g) = G^{-1}\left(1 - q_g/\mu^{\text{out}}\right).$$

If the proportion of the group population without in-group shocks who marry outsiders is to be fixed at q_g, then the planner would wish to allow only those with realizations $b \geq \bar{b}_g(q_g)$ to out-marry if b were observable. The cut-off shock, $\bar{b}_g(q_g)$, by definition, satisfies the condition that a person of gender g who receives such a preference shock is indifferent between marrying the outsider and relinquishing this option and entering the marriage market. The payoff to marrying the outsider is $\alpha_g Y + \bar{b}_g(q_g) - \phi_g(q)$, where $\phi_g(q_g)$ denotes the punishment that is inflicted on the household of the person for violating the endogamy norm. If the person reliquishes the option of marrying the outsider, it not clear that they will find a spouse within the group, for they may find themselves on the long side of the market if too many people of the opposite gender have out-married. Denote by $\rho_g(q)$ the probability of finding an in-group partner upon entering the marriage market, where q denotes the vector (q_m, q_f). This probability is given by

$$\rho_g(q) \equiv \min\left\{\frac{1 - q_{-g}}{1 - q_g}, 1\right\},$$

where $-g$ denotes the gender other than g. The probability $\rho_g(q)$ is determined by the relative scarcities of genders g and $-g$ in the market. The expected income of this marginal individual in conforming would then be $\rho_g(q)\,Y + [1 - \rho_g(q)\,]\,y$, if the household were not required to punish violators. With probability $\mu^{\text{out}}(1 - \mu^{\text{in}})(1 - G(\bar{b}_g)$ the household would be paired up with a violator in the stage game in any period. Since this probability may be written $q_g\,(1 - \mu^{\text{in}})$, the expected cost of punishing violators is $q_g\,(1 - \mu^{\text{in}})\,\phi_g(q)$. Thus the expected payoff of an individual of gender g who enters the in-group marriage market is $\rho_g(q)\,Y + [1 -$

$\rho_g(q)]\, y - q_g\,(1-\mu^{\text{in}})\,\phi_g(q)$.

Since the individual who has received an out-group preference shock $\bar{b}_g(q_g)$ is indifferent, by definition, between embracing and relinquishing the out-group member, we must have

$$\alpha_g\, Y + \bar{b}_g(q_g) - \phi_g(q) = \rho_g(q)\, Y + [1-\rho_g(q)]y - q_g\,(1-\mu^{\text{in}})\,\phi_g(q),$$

where $g \in \{m, f\}$. This gives a relationship between punishment $\phi_g(q)$ and the probability q_g:

$$\phi_g(q) = \frac{\alpha_g\, Y + \bar{b}_g(q_g) - \rho_g(q)\, Y - [1-\rho_g(q)]y}{1-(1-\mu^{\text{in}})\, q_g}.$$

Then, $\phi_g(q)$ is that punishment which ensures that only a proportion q_g of the population will violate the endogamy norm. (There may be a constraint that the size of the punishment is limited by the value of interaction; we cannot have $\phi_g(q) > u^*$.) Instead of formulating the problem as one in which the planner chooses punishments, we see that we can equivalently cast it as one in which the planner chooses the probability q_g. Substituting the expression for $\bar{b}_g(q_g)$ in terms of q_g derived earlier, the implied punishment for gender g is given by

$$\phi_g(q) = \frac{\alpha_g\, Y + G^{-1}\left(1 - q_g/\mu^{\text{out}}\right) - \rho_g(q)\, Y - [1-\rho_g(q)]y}{1-(1-\mu^{\text{in}})\, q_g}.$$

From the above expression, the following result readily obtains:

Proposition 5: For a given q_g, the required punishment for gender g is (a) decreasing in the income Y if $\rho_g(q) > \alpha_g$, and (b) increasing in the probability, μ^{out}, of an individual receiving out-group preference shocks.

Part (a) of Proposition 5 follows from the fact that, when Y is higher, individuals are faced with two opposing effects. On the one hand, if they opt to conform, they can look forward to a higher income only if the probability $\rho_g(q)$ of finding a spouse in the marriage market is high. On the other hand, there is a greater absolute income decline due to skill mismatch between spouses in an inter-caste marriage and so they require less inducement by way of punishment on their parents to conform with endogamy. As long as the probability $\rho_g(q)$ is sufficiently high ($> \alpha_g$), the second effect dominates. As a result, when Y increases the required punishment to induce conformity declines. When the probability of out-group encounters μ^{out} increases, in order to keep q_g fixed the cut-off shock $\bar{b}_g$ must be increased. The indifferent deviator, who now is a person with a larger preference shock for an outsider, therefore requires higher punishment to be dissuaded from marrying the outsider. This explains part (b) of the proposition.

Suppose that, for some $\bar{q}$, the planner were to choose $q_m = q_f \equiv \bar{q}$, that is, the planner equates in expectation the proportion of males and females who embrace out-group preferences.

This is a reasonable condition to impose since the planner would like to ensure that the marriage market is balanced, with equal numbers of males and females. In this case, $\rho_g(\bar{q}) = 1$, since when the marriage market is balanced anyone who enters the market is assured of finding a spouse, thereby minimizing the externality associated with mismatches in the number of male and female participants in the marriage market. We do not investigate the conditions under which it would be optimal to set $q_m = q_f$, but it is reasonable to presume that this would be the case when $y/Y << 1$. In this case, the cost to not finding a spouse from within the group is quite large and the planner would seek to minimize it. Furthermore, if y/Y is lower for the high complementarity groups, as is likely, these are precisely the groups who would wish to equate q_m and q_f. Then, substituting the expression for $\bar{b}_g(q_g)$ in terms of q_g derived earlier, the implied punishment for gender g is given by

$$\phi_g(\bar{q}) = \frac{G^{-1}\left(1 - \bar{q}/\mu^{\text{out}}\right) - (1 - \alpha_g)Y}{1 - (1 - \mu^{\text{in}})\,\bar{q}}.$$

Recall that for males α_g is the (inverse) complementarity dictated by their own occupational technology, whereas for females it is that for their spouses' technology. This makes the punishment gender-dependent. Since the above expression is increasing in α_g, we have the following.

Proposition 6: Consider the high complementarity group. If their planner were to balance the marriage market, then in enforcing endogamy (the households of) their violating females are punished more than (the households of) their violating males.

The intuition is straightforward. For the higher caste in this two-caste scenario, $\alpha_m = \alpha_1 < \alpha_2 = \alpha_f$. Since the income penalty due to spousal mismatch is smaller for females, they have a greater incentive to violate endogamy and hence the punishment on the household has to be more severe. If the punishments are the same, a higher proportion of women would violate endogamy (see Figure 1). For the lower caste, however, $\alpha_m = \alpha_2 > \alpha_1 = \alpha_f$, and so this model would suggest a higher punishment for males. The reason this is unlikely in practice is that, as before, we have not incorporated in our model the perceived differential cost to males and females to leaving behind progeny in a patriarchal system. Were this to be incorporated, the result in Proposition 6 would obtain for the lower caste, too.

In the above analysis we have assumed that the sex-ratio (ratio of females-to-males) in the group is unity. This assumption, however, can be violated—as Sen (1990) and Coale (1991) initially revealed, there are very significant numbers of missing women in India.[19] This is not necessarily only a recent phenomenon. In an analysis of data from 1901, Chakraborty and Kim

[19]See Anderson and Ray (2010) for the most reliable figures for missing women.

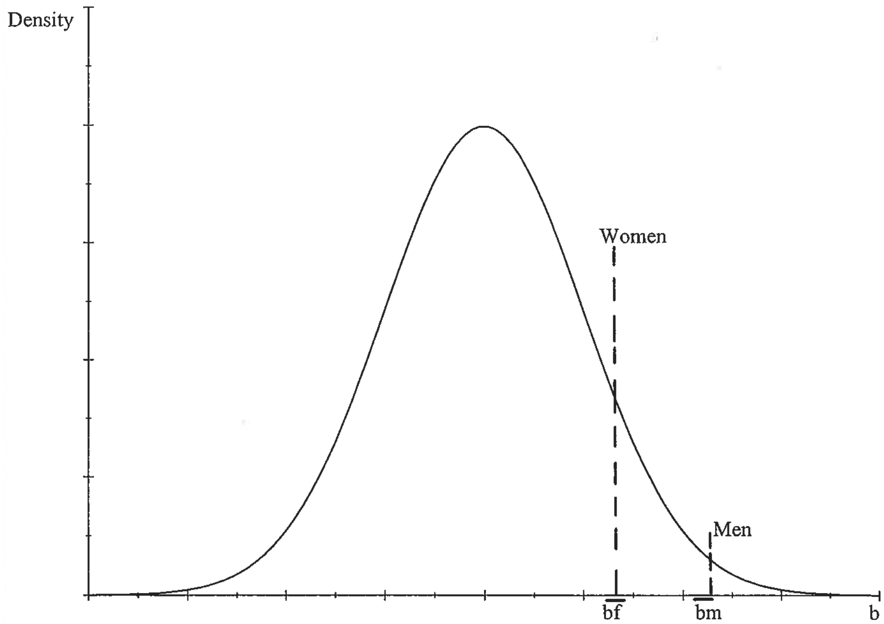

Figure 1: Critical shocks for a given punishment for both genders

(2009) have shown that the sex ratios were significantly lower than unity particularly in the upper castes. As is becoming increasingly clear in recent research, it is in patriarchal societies that there is a strong son preference and sex ratios are biased against females. In the context of our analysis, this suggests that a violation of endogamy by upper caste women would invite an even greater punishment from the group because it is already short of marriageable women. In fact, in such a scenario a social planner would require that a higher proportion respect endogamy among females than among males in order to match their absolute numbers in the marriage market. In other words, to compensate for the women who are missing because of patriarchal values, caste endogamy would require measures that exacerbate the asymmetric treatment of women.

2.4 The Importance of Skill and its Externalities

The advantage of endogamy identified here represents a lower bound if we think of incorporating inter-generational skill transmission. We briefly sketch out a model of this phenomenon in this section because, we argue, it has been important for caste formation in India.

In the previous sections, we interpreted Y as the exogenous output of a couple comprising in-group members. We now endogenize Y as the steady state value derived from a learning process in which human capital is generated and transmitted. To get at this in the simplest manner possible, suppose that the maximal output level, Y, of generation t individuals depends on (i) a constant, Y_0, (ii) a depreciated value of the maximal output level of the previous generation, δY_{t-1}, (with $0 < \delta \leq 1$), and (iii) the average actual output produced in the preceding generation, $\overline{Y}_{t-1}$. Aspect (iii) captures the learning feature of skills; what was actually produced in the previous period contributes this period's maximal output. Note that δ here is a durability parameter, with $\delta = 1$ if the human capital is transmitted intact from one generation to the next even when the output is at its maximal rate. Specifically, we suppose that

$$Y_t = Y_0 + \sigma \overline{Y}_{t-1} + \delta Y_{t-1},$$

where σ measures the extent of social learning. It should be mentioned for future reference that we expect the parameters σ and δ to be increasing in the group size. Larger groups would facilitate greater learning by increasing the number of within-group interactions. Furthermore, larger groups would also be better equipped to ensure durability of the group's skills.[20] So we can write $\sigma = \sigma(N^c)$ and $\delta = \delta(N^c)$, with $\sigma'(.) > 0$ and $\delta'(.) > 0$.

We shall focus on males because it is their occupations that are inherited by the next generation. If the expected proportion of males who do not get in-group partners is z_1 and the expected proportion of males who marry outside the group is z_2, then

$$\overline{Y}_{t-1} = (1 - z_1 - z_2)Y_{t-1} + z_1 y_{t-1} + z_2 \alpha Y_{t-1}.$$

The first term on the right hand side is the average output from couples comprising same-caste members, the second is the output (now time-dependent) from unmarried members, and the last term is the output from inter-caste couples. We let $\tau \equiv y_t/Y_t$ denote the output of a single-member household relative to the maximal output and we take this to be a constant independent of the generation. Then $\overline{Y}_{t-1} = [(1 - z_1 - z_2) + z_1\tau + z_2\alpha]Y_{t-1}$. Substituting this expression into that for Y_t, we obtain the long-run steady state maximal output level that, abusing notation, we denote by Y:

$$Y = \frac{Y_0}{1 - \delta - \sigma[1 - (1 - \tau)z_1 - (1 - \alpha)z_2]}.$$

As expected, this output is increasing in σ and δ, and therefore Y is increasing in the group size, N^c. This output is decreasing in z_1 and z_2, thereby providing an added reason why the group may benefit from endogamy. Out-marriages (an increase in z_2) reduce social learning by reducing the size of the group. And when there is an increases in the number of people who cannot find spouses in the marriage market (an increase in z_1), caste skills are again lost. Note that z_1 is decreased by raising female punishment, whereas z_2 is decreased by raising male punishment. Thus, in general, the punishments a planner would implement would reflect the skill consequences of endogamy violations.

[20] Languages routinely die out, for example, when the group size gets too small.

We may reasonably interpret the effect on Y of a small reduction in z_1 and z_2 as the marginal benefit to greater endogamy. The following result obtains from the partial differentiations of the above expression.

Proposition 7: The marginal benefit to greater endogamy is higher for groups with occupational human capital that is more subject to social learning (captured by σ) and/or is more durable (captured by δ).

Since we have posited σ and δ to be increasing in group size, the benefits of endogamy would be higher for larger groups. Proposition 7 suggests that planners in groups where the occupations lend themselves to group learning and where the techniques of reproducing the output of previous generations have been mastered would be more inclined to adopt endogamy and maximize the gains from them. This observation takes us into our discussion of why the Brahmins insisted on endogamy.

3 Source of Uniqueness of the Indian Caste System[21]

Many societies have had institutions that shared some characteristics of caste. For example, European guilds that existed in the Middle Ages have sometimes been compared to caste. But as Senart (1930) and Weber (1946) have persuasively argued, the similarities are minimal and misleading. (For example, endogamy is not a feature of guilds.) In the 5^{th} Century CE emperor Theodosius, to shore up the Roman empire, passed a law making all public functions hereditary but this did not last after the breakup of the empire.[22] Ancient Egypt had something like a caste system. However, specialization was not necessarily hereditary and a strong monarchy prevented the castes from remaining distinct.[23] In the light of such historical examples, the endurance of India's caste system is very puzzling. Furthermore, in the Hindu caste system there was an unparalled obssession with ritual purity and the 'status' that apparently derived from it. What was the source of this? What historical circumstance led to this peculiar cultural innovation?

It is agreed by scholars of the caste system such as Dumont (1970) that the Brahmins were the lynch-pin of the caste system and that they initiated the first separation into endogamous groups (between the Brahmins and the Kshtriyas). Why was this? The argument in the previous section suggests that groups that possess the most human capital would be the most dedicated to enforcing endogamy. Without a doubt, the Brahmins have had by far the greatest amount of human capital in terms of education (which is the reason why only Brahmins were

[21]This section owes much to a conversation with Vinayak Eswaran.
[22]See Risley (1915, pp. 270-272) for a discussion of this.
[23]See Bougle (1971, pp. 14-16).

permitted to teach). But this was not all. Their stock of knowledge was of the religious, non-secular kind. The Brahmins were largely responsible for transmitting the Hindu scriptures intact over the last three millennia. Of this period, for around a millennium this transmission was done *orally*, without the benefit of written script.[24] Even after script became available in India, the transmission proceeded orally, it is estimated, for another millennium.[25] To appreciate the enormity of the task, note that the *Rig Veda*, which is the first among four of the Vedas, has around 11,000 verses (about 1.38 times the size of the *New Testament*). If to the four Vedas, we add the *Upanishads*, the *Brahamanas*, the *Puranas*, etc., we begin to see that the task of transmitting these error-free generation after generation is monumental. That the transmission was indeed largely error-free may be inferred from the fact that there is only a single version of each of these scriptures extant; there are no variants, which surely would have arisen if there were deviations from the originals. In view of this, an enormous burden was placed on the memory of those who were of the priestly occupation.

An entire educational system was invented and perfected in Vedic times (starting around 1,500 BCE) to accomplish the task of oral transmission because script was unknown in India until 300 BCE [Scharfe (2002, Ch. 2)]. The task was accomplished by children (usually boys) who start memorizing and constantly repeating the scriptures from a very early age. If a boy were to memorize 10 verses a day, it would take him three years to memorize the entire *Rig Veda*. And, no doubt, constant repetition is required to continue retaining it in memory. Under the careful supervision of a teacher for many months of the year, Vedic Brahmin students recited the scriptures in pairs, so that one boy may step in and correct the other were he to falter.[26] The benefit from learning through interaction here (as captured by the parameter σ in the previous section) should be clear enough.

Various techniques were devised by the Brahmins to ensure the authenticity of the oral reproduction. There were upto eleven different variants of memorization techniques developed [Filliozat (2004)]. One was the ordinary continuous memorization; another was a halting word-by-word recitation; a third was repetition of successive pairs of words; a fourth followed the correct order of these pairs with a reverse order recitation and then followed again by the correct order repetition; and so on. The purpose of these many techniques was to familiarize the student with viewing the scripture from every possible angle so as to embed the verses in the mind permanently. These various methods also facilitated accurate transmission by enabling comparisons across the versions from different techniques and an immediate correction of errors. In this manner, the scriptures were transmitted with precision from teachers to

[24] Scharfe (2002) claims, "The Brahmin reciters of the Veda have done a stupendous service to India and indeed to humanity by preserving in immaculate poetry the religious purity of the poetry of their ancestors who lived more than three millennia ago—without the aid of writing or other mechanical devices" (Ch. 13). See also Avari (2007, pp. 76-80).

[25] See Witzel (2003, p. 69).

[26] See Scharfe (2002, p. 26).

students across generations without break.[27] In the concepts of the previous section, this degree of accuracy corresponded to a high degree of durability in reproduction (the parameter δ was high).

Transmission of the scriptures across generations was further facilitated by the division of labor: different scriptures were relegated to different family lines. Some specialized in the *Vedas*, others in the *Upanishads*, etc. Even among these, there was further sub-specialization. (For example, there is a distinction between those who specialized in two of the four *Vedas*, in three of them, or in all four. They formed separate subcastes among the Brahmins, with the rules of endogamy separating them.) This division of labor would have made it essential that the group size not diminish, for that would have meant a loss of scriptures. This underlines the importance of ensuring progeny who will carry forward the tradition, which of course requires the availability of women for marriage. The preferred women would clearly be those with skills complementary to those of their husbands and with an upbringing conducive to the Brahmin's way of life. The choice of wives would also have been thought very important because mothers spend more time with the children than do fathers. There is considerable evidence in contemporary settings suggesting that religious beliefs and practices are most effectively transmitted across generations when the parents have similar beliefs [see Clark and Worthington (1990) for a review]. Myers (1996) finds that the family context is extremely important to belief transmission; the absence of conflict between parents increases the efficacy of religious socialization of children.

The theoretical work of Bisin and Verdier (2000) has shown that those in society who subscribe to certain cultural and religious beliefs which they would like to pass on to their children are more likely to have homogamous marriages (that is, both partners share the same beliefs). They also put more effort into socializing their children, and in the long-run minorities manage to retain their identities despite being embedded in a "melting pot". These insights are quite relevant to the situation under consideration here. It is not surprising that the Brahmins of India, who had very compelling reasons starting three millennia ago for transmitting their knowledge and beliefs to future generations, would have opted to be homogamous.

A household environment that was not conducive to this peculiar form of learning invented by the Brahmins would have resulted in the loss of the capital the Brahmins deemed most precious. Any distraction or incompatibility would have proved detrimental to a Brahmin's task. We submit that the self-perceived mandate of the Brahmins to preserve the scriptures *at a time when there was no script* was the unique circumstance in India that led to the exacting features of the Brahmanical caste system. The measures that were subsequently implemented to enforce endogamy in this system and what followed as a consequence are discussed in the

[27] "This ensured an imppeccable textual transmission superior to the classical texts of other cultures; it is, in fact, something like *tape-recording* of ca. 1,500–500 BCE. Not just the actual words, but even the long-lost musical (tonal) accent (as in old Greek or in Japanese) has been preserved up to the present." [Witzel (2003, pp. 68-69). Emphasis in the original]

next section.

4 Endogamy and the Origin of Sub-Castes (*Jatis*)

Caste endogamy is enforced by a litany of highly specific regulations and punishments for violations, many of which are catalogued in the *Code of Manu* (compiled around 300 CE). The punishments are imposed and overseen by caste councils comprising village elders from the caste. The punishments generally seem to fit the perceived crime: minor infractions are punished with fines and penances of various sorts, and major ones can result in excommunication from the caste—which meant that all caste members shun their company—and sometimes even death. The most important punishments, from our point of view, are those pertaining to endogamy. Actions that are likely to lead to violations of endogamy are proscribed and actual violations of endogamy severely punished (often with excommunication).

Punishments for offences depend on the castes of the offender and the victim [Kane (1953, v. 4, pp. 80-81), Tambiah (1973)]. A distinction is made between two sorts of offenses. For one kind, those in the lower castes were given higher punishments. In particular, a lower caste member (deemed 'impure') received a harsher punishment for injuring a Brahmin (deemed 'pure') than did a Brahmin who inflicted the same injury on a lower caste person. The putative principle here is that the higher the purity (or status) of the victim relative to the offender, the greater the punishment. In fact, this is true even today: experimental evidence by Hoff et al (2011) shows that members of the lower caste are less willing to punish higher caste members who pepetrate harm on a fellow caste member than are higher caste members in a reciprocal scenario. This self-serving preferential treatment of Brahmins does suggest the use (or abuse) of power, but it is also consistent with the goal of maximizing the group size of a caste that is persuaded it is carrying precious intergenerational capital. For offenses of another kind, punishments are harsher for the higher castes. Those actions involving pollution by contact with inferiors, for example, fall in this category. The principle here is that the higher the status of the offender, the greater is the fall in purity though contact with inferiors and the penance must be commensurately greater [Tambiah (1973)]. The punishments for both sorts of offenses can be explained quite simply as means through which higher castes maintained group size and endogamy—by preventing other groups from harming the upper castes on the one hand and, on the other, by dissuading the upper castes from mixing with the lower. In either case, we see the notions of purity and pollution as being merely instrumental in serving the goal of endogamy and maintaining group size (for its intergenerational spillover effects).

Among the actions that could lead to pollution, an important one is commensality and the eating of lower-caste food. Eating together necessitates interaction and leads to familiarity, and there are strict restrictions on inter-caste commensality. Again, rules like these are rationalized in terms of the purity/pollution dichotomy: the allegedly purer higher castes cannot afford to pollute themselves through contact or by eating lower caste food. We argue that these restrictions simply arise from the need to minimize the probability (μ^{out}) of an individual conceiving a preference for out-group members. Once again, notions of purity and pollution are merely vehicles for minimizing violations of endogamy.

Other than murder, the most serious offences were deemed to be inter-caste marriages—naturally enough, because these threatened the very existence of the caste system. Once again, such marriages were deemed offensive (especially hypogamy) for reasons pertaining to purity and pollution.[28] The implementation of strict caste endogamy came with very oppressive measures on women. To prevent women from establishing contact from out-group males, their autonomy was severely curtailed.[29] Here is what Manu said: "By a girl, by a young woman, or even by an aged one, nothing must be done independently, even in her own house. In childhood a female must be subject to her father, in youth to her husband, when her lord is dead to her sons; a woman must never be independent." [Buhler (1886, V, 147-148)]

Women were repeatedly derided for ostensibly having an excessive sexual appetite in the *Code of Manu* and other *Dharma-Shastras* (Rules of Right Conduct) [Wadley (1977), Chakravarti (1993)]. "Through their passion for men, through their mutable temper, through their natural heartlessness, they become disloyal towards their husbands, however carefully they may be guarded in this (world)." [Buhler (1886, IX, 15)] The passions of women were deemed inately impure, unless channeled into virtuous avenues through unconditional devotion to a husband.[30] "Though destitute of virtue, or seeking pleasure (elsewhere), or devoid of good qualities, (yet) a husband must be constantly worshipped as a god by a faithful wife," says Manu [Buhler (1886, V, 154)]. The frequently expressed misogyny in Brahmanical literature very likely served several related purposes. First, it attempted to thwart men from falling prey to their own sexual urges (conveniently projected onto women) and from dissipating themselves in sexual escapades instead of pursuing their caste preoccupation of faithfully reproducing the scriptures. Secondly, impugning women's character would have provided men with the justification to monitor their sexuality in order to enforce endogamy and ensure paternity. Finally, as forcefully argued by Chakravarti (2003), religious approval of the view that women are allegedly inferior to men would have aided the process of socialization whereby women come to internalize these views and thus become unwitting accomplices in their own subjugation. To be sure, oppression of women through the attempted control of their sexuality is a common feature of all patriarchal societies. Our model quite unambiguously suggests, however, that in the Indian context the caste system has exacerbated the oppression of women

[28] Yalman (1963, p.42) reports that in Sri Lanka (which shares the caste system among the Hindus) if a child was conceived through hypogamy, both mother and child used to be drowned as a deterrent.

[29] Even the *Bhagavad Gita*, a very important scripture for Hindus that was composed around 200 BCE, echoes the idea that the chastity of women is essential to the maintenance of castes (Ch. 1, verse 41).

[30] A telling example is found in the Hindu epic *Ramayana* in which Sita, the wife of god Rama, embodies the highest ideals of the Hindu wife. After she was abducted by Rama's archrival and Rama rescues her by waging a war, he asked Sita to prove that she had remained chaste during her absence.

over and above what they would have suffered in an established patriarchy.

As mentioned earlier, the view of anthropologist Dumont (1970) is that the entire caste system follows from religious notions of purity and pollution. He argues that political economy considerations are secondary; secular notions of power had no play in the the establishment of the caste system. Although he has been criticised for this position [e.g. see Marglin (1977)], we agree with Dumont that secular notions of power and exploitation probably had at best a small role in the emergence of the caste system. Our theory certainly requires no such premise. However, it is incumbent on us to inquire where such notions of purity and pollution came from, for invoking them comes close to *assuming* what needs to be explained. In our view, the caste system arose to exploit the complementarities in family production. The degree of complementarity, as we have shown, determines the rank in the caste hierarchy. Once hierarchy has been established, however, there is no doubt that it was subsequently utilitzed down the millennia for economic and political advantage by the upper castes. Our model only speaks to the issue of the *origin* of of the caste system.

In our view, the motivation of endogamy in the caste system was fundamentally socioeconomic and so we reject Dumont's claim that political-economy considerations were secondary. Notions of purity and pollution that Dumont takes as given were, as we have seen above, are merely the means to an end. One thing that stands out in caste interactions is that pollution always trumps purity [Mandelbaum (1970, Vol. 1, Ch. 11)]. If the two come into contact, it is pollution that wins out; it is almost never that purity neutralizes the pollution. The reason for this immediately becomes clear if we assess the notions of purity and pollution in terms of their functions. If endogamy is to be sustained through these twin concepts, pollution *must* always dominate purity. Were the reverse true, the maintained distinctions between the castes would be erased by those who are allegedly pure. A theory that begins with the purity/pollution dichotomy as a given confuses ends and means. The end was to sustain endogamy for socioeconomic reasons; notions of purity and pollution were the means.

Our model can explain other features of the caste system like arranged marriages. Quite simply, this expedient reduces the probability of offspring conceiving a preference for an out-group member. In fact, it even reduces the probability of contact with in-group members of the opposite sex, thereby granting greater paternity assurance to potential within-caste grooms. Similarly, child marriage—a practice initiated by the upper castes—commits children to marital alliances before they reached puberty and become open to entertaining preferences for out-group members. Our model also suggests a reason why widow remarriages were banned (especially in the upper castes), but not widower remarriages. Widows are typically older and so further along in their reproductive cycle. If they remarry, the chances are higher that they will find younger grooms because older men would mostly be married and widowers would be few and far between. Marriages of widows with younger men would result in fewer descendants, something that castes would have wanted to avoid because group size mattered.

This also explains the exacting requirements of 'purity' imposed on widows. The need to ban widower remarriages is less pressing, however, because males can have children until much later in their lives and the reproductive costs to the community of remarriage would be correspondingly lower. Once again we see gender asymmetries arising from the need to maintain caste integrity.

We now explain the emergence of sub-castes (*jatis*). Once the practice of endogamy came to be established through the means described above, our model suggests that these very means can be employed to sustain endogamy between subgroups (sub-castes). Thus subgroups with slightly differentiated occcupations requiring complementary spousal skills came into existence. So within the varna of Brahmins, there were many different specializations and each specialized branch would have liked to enforce endogamy. Likewise for specializations within the Ksatriya, Vaishya, and the Shudra varnas. The demand for subcastes within each varna would depend upon the extent of *differentiation* in occupations within the varna and how transferrable or untransferrable skills were across these occupations. Differentiation in occupations would depend on the specialization warranted by the economy. Thus economic development, which would increase the number of activities and occupations, would lead to an increase in the number of sub-castes. Even within the occupations relating to textiles, for example, during the mid-18th through mid-19th centuries in South India, weavers, bleachers, dyers, printers, etc. all formed separate sub-castes. [Kumar (1982, Vol. II, p. 355)]. As long as the size of each subgroup is sufficiently large (which, again, depends on the extent of the market), knowledge-spillovers across generations would motivate endogamy within these subgroups and *jatis* would obtain. Thus we could say, paraphrasing Adam Smith, that division into *jatis* is dependent on the extent of the market. We would expect the number of *jatis* to also increase with population size. It is in this manner, we claim, that thousands of sub-castes have come into existence over the millennia and endogamy came to operate at the level of these sub-castes. Our theory, then, clearly predicts that *jatis* emerged from *varnas*.

There are some examples of how some sub-castes have sought to move higher up in the hierarchy by relinquishing practices that are eschewed by the higher castes and adopting some of their practices [Stevenson (1954)]. Srinivas (1956) has referred to this process as 'Sankritization', whereby the lower castes seek to emulate the higher. But this process is usually also accompanied by changes in economic activities and occupations. Blunt (1969, pp. 50-57) gives many examples of the fissioning of castes as a result of slight changes in occupations. He also gives numerous examples of sub-castes rising in the social ranks by giving up certain customs. However, he points out that "*the change of custom is the result of fission and not its cause*" (p. 53, emphasis added). This observation fits in well with our theory: shifts in economic activities change the requirements of spousal complementarity and rationalize a separate endogamous sub-caste that rejects potential spouses from the original sub-caste. But this would would occur only when the size of the subgroup is large enough because only then can we expect intergenerational spillovers to be significant.[31] The change in ritual practices then are merely signals that separate in-group members from former group members who are no longer deemed appropriate for alliances.

5 Conclusions

In this paper we have proposed a theory of how the caste system originated in India, explaining all its essential features. This Hindu institution is embedded so inescapably in the country's social fabric that even Christians and Muslims living in India are known to adhere to the caste system.[32] Our theory posits that women's economic contributions to household production are essential to the income from their husband's occupations. When both spouses hail from families in the same occupations, spousal complementarities in production enable incomes to be at maximal levels. When there is a mismatch in spousal skills, income falls. When a member of an occupational group out-marries, there is an externality in that some group member of the opposite sex does not find a match within the group. To minimize this externality the group enforces endogamy with the threat of punishments, and a caste is born. We showed that the punishments for violations are harsher for women than for men. We also explained why there is less resistance to women marrying up (hypergamy) than to them marrying down (hypogamy), where 'up' and 'down' are determined by the relative degrees of complementarities of the two castes. If the castes can be ranked by these measures of complementarities, we showed that a hierarchy in castes obtains. Our theory also brings out the fact that caste is integral to the subjugation of women in the Indian context, a claim that has been made by some feminist writers [especially Chakravarti (2003)].

Our theory of the origin of the caste sytem suggests that an *inter-group* exercise of power may not have been a necessary part of the formation. That exercise of power (and the attendant inequity) seems to have arisen after the establishment of hierarchy. Our theory has shown that to implement the caste system, there had to be mostly only a *within-group* exercise of power. Women had to be treated with greater severity than men to enforce endogamy. Ironically, it is women of the upper castes that have borne the brunt of this. This aspect of caste is very visible in India even today: the actions of upper caste women are monitored and circumscribed much more than those of lower caste women. Patriarchy—possibly originating at least partially from the need for paternity assurance in a world where only maternity is certain—is itself oppressive to women because of the male need to control their sexuality [Smuts (1995), Lerner (1986)]. When the exigencies of caste are added to this, the oppression of women is that much worse. Caste is not patriarchy, but the elaborate caste system of India as it has come down certainly requires it.

[31] Large subgroup sizes would also be required for the obvious reason of avoiding too much in-breeding.

[32] For Christians, see Tharamangalam (1996); for Muslims, see Bhatty (1996).

In recent decades, the empowerment of women has become an important component of the agenda in developing countries. Since the theory of this paper demonstrates that the caste system depends on the oppression of women, it follows that the empowerment of women will weaken the caste system. Caste-based punishments for violations of endogamy will have little bite if women have lucrative outside options. This link between the empowering of women and the undermining of the caste system is worth recognizing.

References

[1] Alesina, A.F., P. Giuliano, and N. Nunn (2011), "On the Origins of Gender Roles: Women and the Plough", Centre for Economic Policy Research, Discussion Paper No. 8418, June.

[2] Akerlof, G. (1976), "The Economics of Caste and of the Rat Race and Other Woeful Tales", *Quarterly Journal of Economics*, 90, pp. 599-617.

[3] Ambedkar, B.R. (1917), "Castes in India: Their Mechanism, Genesis and Development", *Indian Antiquary*, XLI, pp. 81-95.

[4] Anderson, S. (2011), "Caste as an Impediment to Trade", *American Economic Journal: Applied Economics*, 3, pp. 239–263.

[5] Anderson, S. and D. Ray (2010), "Missing Women: Age and Disease", *Review of Economic Studies*, 77, pp. 1262-1300.

[6] Avari, B. (2007), *India: The Ancient Past*, Routledge, London.

[7] Banerjee, A.V. and R. Somanthan (2007), "The Political Economy of Public Goods: Some Evidence from India", *Journal of Development Economics*, 82, pp. 287–314.

[8] Bhatty, Z. (1996), "Social Stratification Among Muslims in India", in *Caste: Its Twentieth Century Avatar*, (ed.) M.N. Srinivas, Viking, New Delhi, pp. 244-262.

[9] Bisin, A. and T. Verdier (2000), " 'Beyond the Melting Pot': Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits", *Quarterly Journal of Economics* , 115, pp. 955-988.

[10] Blunt, E.A.H. (1969), *The Caste System of Northern India*, S. Chand & Co., Delhi.

[11] Bougle, C. (1971), *Essays on the Caste System*, translated by D.F. Pocock, Cambridge University Press, Cambridge.

[12] Buhler, G. (1886), *The Laws of Manu*, Clarendon Press, Oxford.

[13] Chakravarti, U. (1993), "Conceptualising Brahmanical Patriarchy in Early India: Gender, Caste, Class and State", *Economic and Political Weekly*, 28: 14, (April 3, 1993), pp. 579-585.

[14] Chakravarti, U. (2003), *Gendering Caste Through a Feminist Lens*, Stree, Calcutta.

[15] Choy, J.P. (2011), "A Theory of Cooperation in Groups, with Evidence from Nepal", Department of Economics, Yale University, unpublished manuscript.

[16] Clarke, C.A. and E.L. Worthington (1990), "Family Variables Affecting the Transmission of Religious Values from Parents to Adolescents: A Review", in *Parent-Adolescent Relationships*, (eds.) B.K. Barber and B.C. Rollins, University Press of America, Boston.

[17] Coale, A.J. (1991), "Excess Female Mortality and the Balance of the Sexes in the Population: An Estimate of the Number of 'Missing Female"', *Population and Development Review*, 17, pp. 517-523.

[18] Cole, H.L., G.J. Mailath, and A, Postelwaite (1998), "Class Systems and the Enforcement of Social Norms", *Journal of Public Economics*, 70, pp. 5–35.

[19] David, N. and C. Kramer (2001), *Ethnoarchaeology in Action*, Cambridge University Press, Cambridge.

[20] Davis, K. (1941), "Intermarriage in Caste Societies", *American Anthropologist*, New Series, 43, pp. 376-395.

[21] Deshpande, A. (2011), *The Grammar of Caste*, Oxford University Press, New Delhi.

[22] Dube, L. (1996), "Women and Caste", in *Caste: Its Twentieth Century Avatar*, (ed.) M.N. Srinivas, Viking, New Delhi, pp. 1-27.

[23] Dumont, L. (1970), *Homo Hierarchicus: The Caste System and Its Implications*, translated by M. Sainsbury, Chicago University Press, Chicago.

[24] Esteban,J., L. Mayoral, and D. Ray (2012), "Ethnicity and Conflict: An Empirical Study", *American Economic Review*, 102, pp. 1310–1342.

[25] Filliozat, P.-S. (2004), "Ancient Sanskrit Mathematics: An Oral Tradition and a Written Literature", in *History of Science, History of Text*, (ed.) K. Chemla, Springer, Netherlands.

[26] Freitas, K. (2006), "The Indian Caste System as a Means of Contract Enforcement", Nortwestern University, unpublished manuscript.

[27] Greif, A. (1993), "Contract Enforceability and Economic Institutions in Early Trade: the Maghribi Traders' Coalition", *American Economic Review*, 83, pp. 525-548.

[28] Hoff, K., M. Kshetramade, and E. Fehr (2011), "Caste and Punishment: The Legacy of Caste Culture in Norm Enforcement", *Economic Journal*, 121, pp. F449-F475.

[29] Holden, C.J. and R. Mace (2003), "Spread of Cattle Led to the Loss of Matrilineal Descent in Africa: a Coevolutionary Analysis", *Proceedings of the Royal Society of London B*, 270, pp. 2425–2433.

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

[31] Jamison, S.W. (1996), *Sacrificed Wife/Sacrificer's Wife: Women, Ritual, and Hospitality in Ancient India*, Oxford University Press, New York.

[32] Kane (1953), *History of Dharmasastra*, Bhandarkar Oriental Research Institute, Poona.

[33] Kramer, C. (1997), *Pottery in Rajasthan: Ethnoarchaeology in Two Indian Cities*, Smithsonian Institution Press, Washington.

[34] Kumar, D. (1982), "Regional Economy (1757-1857): South India" in *The Cambridge Economic History of India*, Vol. II, (ed.) D. Kumar, Cambridge University Press, Cambridge.

[35] Le Bon (2005), "The Formation of the Castes in India", in *The Great Events by Famous Historians*, Vol. 1, (eds.) R. Johnson, C. Horne, and J. Rudd, Project Gutenberg.

[36] Lerner, G. (1986), *The Creation of Patriarchy*, Oxford University Press, New York.

[37] Mandelbaum, D.G. (1970), *Society in India: Continuity and Change*, University of California Press, Berkeley.

[38] Marglin, F.A. (1977), "Power, Purity and Pollution: Aspects of the Caste System Reconsidered", *Contributions to Indian Sociology*, 11, pp. 245-270.

[39] Milner, M. (1994), *Status and Sacredness: A General Theory of Status Relations and an Analysis of Indian Culture*, Oxford University Press, Oxford.

[40] Munshi, K. (2011), "Strength in Numbers: Networks as a Solution to Occupational Traps", *Review of Economic Studies*, 78, pp. 1069–1101.

[41] Munshi, K. and M. Rosenzweig (2006), "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy", *American Economic Review*, 96, pp. 1225–1252.

[42] Myers, S.M. (1996), "An Interactive Model of Religiosity Inheritance: The Importance of Family Context", *American Sociological Review*, 61, pp. 858-866.

[43] Pande, R. (2003), "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India", American Economic Review, 93, pp. 1132-1151.

[44] Risley, H. (1915/1969), *The People of India*, Oriental Books Reprint Corporation, Delhi.

[45] Sahoo, S. et al (2006), "A Prehistory of Indian Y Chromosomes: Evaluating Demic Diffusion Scenarios", *Proceedings of the National Academy of Sciences*, 104, pp. 843–848.

[46] Scharfe, H. (2002), *Education in Ancient India*, Brill Academic Publishers, Amsterdam.

[47] Sen, A.K. (1990), "More than 100 Women are Missing", *New York Review of Books*, 37, December 20th.

[48] Senart, E. (1930), *Caste in India: The Facts and the System*, Methuen & Co., London.

[49] Shaffer, J.G. (1984), "The Indo-Aryan Invasions: Cultural Myth and Archaeological Reality", in *The People of South Asia: the Biological Anthropology of India, Pakistan, and Nepal*, (ed.) J.R. Lukacs, Plenum Press, New York.

[50] Shafer, J.G. and D.A. Lichtenstein (2005), "South Asian Archaeology and the Myth of Indo-Aryan Invasions", in *The Indo-Aryan Controversy: Evidence and Inference in Indian History* , (eds.) E.F. Bryant and L.L. Patton, Routledge,, London.

[51] Smuts, B. (1995), "The Evolutionary Origins of Patriarchy", *Human Nature*, 6, pp. 1-32.

[52] Srinivas, M. N. 1956, "A Note on Sanskritization and Westernization", *Far Eastern Quarterly*, 15, 4:481-496. Reprinted in *Collected Essays*, M.N. Srinivas, Oxford University Press, New Delhi 2002.

[53] Stevenson, H.N.C. (1954), "Status Evaluation in the Hindu Caste System", *Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 84, pp. 45-65.

[54] Tambiah, S.J. (1973), "From Varna to Caste Through Mixed Unions", in *The Character of Kinship*, (ed.) J. Goody, Cambridge University Press, Cambridge.

[55] Tharamangalam, J. (1996), "Caste Among Christian in India", in *Caste: Its Twentieth Century Avatar*, (ed.) M.N. Srinivas, Viking, New Delhi, pp. 263-291.

[56] Wadley, S.S. (1977), "Women and the Hindu Tradition", *Signs*, 3, pp. 113-125.

[57] Weber, M. (1946), "India: The Brahman and the Castes", *From Max Weber: Essays in Sociology*, (trans. & ed.) H. H. Gerth and C. W. Mills, Oxford University Press.

[58] Witzel, M. (2003), "Vedas and Upanishads", in *The Blackwell Companion to Hinduism*, (ed.) G. Flood, Blackwell Publishing, Malden, USA.

[59] Yalman, N. (1963), "On the Purity of Women in the Castes of Ceylon and Malabar", *Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 93, pp. 25-58

Appendix

Proof of Proposition 1: Since $\tilde{\pi}(N)$ is independent of the parameters stated in the proposition, it is sufficient to analyze their effect on $U^E(N) - U^{NP}(N)$. To this end, after some simplification, the expression for $U_g(n_m, n_f)$ in the absence of endogamy norms may be rewritten as

$$\begin{aligned} U_g(n_m, n_f) = N\,(Y + u^*) \\ + (N - n_g)\,[B - (1 - \alpha_g)Y] - [n_g - \min(n_m, n_f)]\,(Y - y_g), \end{aligned} \quad (7)$$

while that for the aggregate payoff may be rewritten as

$$\begin{aligned} U(n_m, n_f) = 2(N^c - N)(Y + B + u^*) \\ + 2N\,(Y + u^*) \\ + (N - n_m)\,[B - (1 - \alpha_m)\,Y] + (N - n_f)\,[B - (1 - \alpha_f)\,Y] \\ - [\max\{n_m, n_f\} - \min\{n_m, n_f\}]\,(Y - y). \end{aligned} \quad (8)$$

The net payoff to endogamy, $U^E(N) - U^{NP}(N)$, may be written as

$$U^E(N) - U^{NP}(N) = \sum_{m=0}^{N} \sum_{f=0}^{N} \pi(n_m, n_f)\, w(n_m, n_f), \quad (9)$$

where (after some simplification)

$$\begin{aligned} w(n_m, n_f) \equiv [\max(n_m, n_f) - \min(n_m, n_f)]\,(Y - y) - \\ \{(N - n_m)\,[B - (1 - \alpha_m)\,Y] \\ +(N - n_f)\,[B - (1 - \alpha_f)\,Y]\}\,. \end{aligned} \quad (10)$$

The results for Y, y and B, parts (a), (b) and (c) of the proposition respectively, immediately follow from noting that w defined above is increasing in Y and decreasing in y and B.

The fact that q is the same across genders implies that $\pi(n_m, n_f) = \pi(n_f, n_m)$. Therefore, we can write

$$U^E - U^{NP} = \sum_{n=0}^{N} \pi(n, n)\, w(n, n) + \sum_{n_m=1}^{N} \sum_{n_f=0}^{n_m-1} \pi(n_m, n_f)\, \widetilde{w}(n_m, n_f), \quad (11)$$

where

$$\widetilde{w}(n_m, n_f) = w(n_m, n_f) + w(n_f, n_m)$$
$$= 2[\max\{n_m, n_f\} - \min\{n_m, n_f\}](Y - y) - (N - n_m + N - n_f)[2(B - Y) + (\alpha_m + \alpha_f)Y]. \quad (12)$$

We see that $\widetilde{w}$ is clearly decreasing in $\alpha_m + \alpha_f$. The results that the net benefit to endogamy, too, is decreasing in $\alpha_m + \alpha_f$, part (d) of the proposition, follow since

$$w(n, n) = -(N - n)[2(B - Y) + (\alpha_m + \alpha_f)Y] \quad (13)$$

is decreasing in $\alpha_m + \alpha_f$. □

Proof of Proposition 2: Since $\tilde{\pi}(N)$ is independent of μ^{out}, it is sufficient to analyze the effect of μ^{out} on $U^E(N) - U^{NP}(N)$.

When $\mu^{\text{out}} = 0$, there are no instances of out-marriage and therefore $U^E - U^{NP} = 0$. Since $p > 0$ and $q = \mu^{\text{out}} \cdot p$, the result follows once we show that $d(U^E - U^{NP})/dq > 0$ when evaluated at $q = 0$. Using (11), we have

$$\frac{d}{dq}(U^E - U^{NP}) = \sum_{n=0}^{N-1} \frac{d}{dq}\pi(n, n)\, w(n, n) + \sum_{n_m=1}^{N} \sum_{n_f=0}^{n_m-1} \frac{d}{dq}\pi(n_m, n_f)\, \widetilde{w}(n_m, n_f), \quad (14)$$

where we have dropped $\pi(N, N)\, w(N, N)$ in the first term because it is zero.

Letting $\xi_1 = m + f$ and $\xi_2 = N - m + N - f$, we have

$$\frac{d}{dq}\pi(n_m, n_f) = \begin{pmatrix} N \\ m \end{pmatrix} \begin{pmatrix} N \\ f \end{pmatrix} [\xi_2 q^{\xi_2 - 1}(1 - q)^{\xi 1} - \xi_1 (1 - q)^{\xi 1 - 1} q^{\xi_2}] \quad (15)$$

$$= \begin{pmatrix} N \\ m \end{pmatrix} \begin{pmatrix} N \\ f \end{pmatrix} q^{\xi_2 - 1}[\xi_2 (1 - q)^{\xi 1} - \xi_1 q (1 - q)^{\xi 1 - 1}]. \quad (16)$$

Notice that the limit of the bracketed term as $q \to 0$ is ξ_2. Therefore, we have

$$\lim_{q \to 0} \frac{d}{dq}\pi(m, f) = \lim_{q \to 0} \frac{d}{dq} \begin{pmatrix} N \\ m \end{pmatrix} \begin{pmatrix} N \\ f \end{pmatrix} \xi_2 q^{\xi_2 - 1} = 0 \text{ if } \xi_2 > 1 \text{ and } 1 \text{ if } \xi_2 = 1. \quad (17)$$

Since $\xi_2 > 1$ if and only if $(m, f) \notin \{(N, N-1), (N-1, N), (N, N)\}$, and $\xi_2 = 1$ if and only if $(m, f) \in \{(N, N-1), (N-1, N)\}$, it follows that

$$\lim_{q \to 0} \frac{d}{dq}(U^E - U^{NP}) = \lim_{q \to 0} \frac{d}{dq}\pi(N, n-1)\, \widetilde{w}(N, N-1)$$
$$= 2(Y - y) - [2(B - Y) + (\alpha_m + \alpha_f)Y],$$

which is positive by (1). Since the net benefit $(U^E - U^{NP})$ is zero at $q = 0$ but is increasing at this point, endogamy is beneficial for sufficiently small q and therefore for small μ^{out}. $\square$