How Do We Choose Our Identity? A Revealed Preference Approach Using Food Consumption*

David Atkin†, Eve Colson-Sihra‡ and Moses Shayo§

April 2019

Abstract

Are identities fungible? How do people come to identify with specific groups? This paper proposes a revealed preference approach, using food consumption to uncover ethnic and religious identity choices in India. We first show that consumption of identity goods (e.g. beef and pork) responds to forces suggested by social-identity research: group status and group salience, with the latter proxied by inter-group conflict. Moreover, identity choices respond to the cost of following the group's prescribed behaviors. We propose and estimate a modified demand system to quantify the identity changes that followed India's 1991 economic reforms. While social-identity research has focused on status and salience, economic costs appear to play a dominant role.

*We thank Ben Deaner, Omer Karaduman, and Sumit Shinde for excellent research assistance. David Atkin thanks the Stanford Economics Department and SIEPR for their hospitality while writing this paper. Moses Shayo thanks the I-Core Program at the Israel Science Foundation (grant no. 1821/12) and the Falk Institute for financial support. We thank Roland Benabou, Armin Falk, David Genesove, Marco Gonzalez-Navarro, Matthew Gentzkow, Ori Heffetz, Supreet Kaur, Edward Lazear, Shabana Mitra, Sharun Mukand, Salvatore Nunnari, Ran Shorrer, Katia Zhuravskaya, and many seminar participants for valuable comments.

†MIT and NBER

‡The Hebrew University of Jerusalem

§The Hebrew University of Jerusalem

1 Introduction

This paper asks whether people's social identities are fungible, and if they are, how do people come to identify with specific groups? In their seminal paper, Akerlof and Kranton (2000) detail how identities affect a host of outcomes of interest to economists, including human capital acquisition, labor market participation and poverty. The important role identity plays is also salient in the current political environment. Both commentators and scholars have linked the effects of globalization, immigration and rising inequality to a shift away from a now lower-status working class identity towards a nationalist one across both Europe and the US (Shayo 2009). These identity shifts have also been linked to changes in trade policies and opposition to globalization (Grossman and Helpman 2018), with the political repercussions seen in Brexit, the resurgence of the far right in Europe, and Donald Trump's election victory. Arguably, similar forces have played out in India with the ascent of Narendra Modi and the Bharatiya Janata Party (BJP) on the back of a surge in Hindu nationalism.

While it is straightforward to motivate why it is interesting to study identity choices and how they affect economic behavior, measuring identity choices is challenging. Broadly speaking, the existing literature has relied on ethnographic or historical case studies, on survey questions about identification, and on lab experiments. Case studies can provide important insights but are often very specific and evaluating causal arguments is hard. Surveys offer the possibility of broad representative samples, but it is less clear whether self-reported identities reflect day-to-day economic behavior. Finally, experiments can provide rich revealed-preference data but they are necessarily limited in scope to a particular time, place, and population (often students at American universities). A more recent strand of the literature uses naturally occurring data to understand how social identities can affect the behavior of judges (Shayo and Zussman 2011), team production (Hjort 2014), conflict (Depetris-Chauvin and Durante 2017), and grading decisions (Feld et al. 2016; Lavy et al. 2018). We build on these recent advances.

Our approach starts from the observation that consumption choices are both widely documented and are affected by the norms and taboos of groups people identify with. Since different groups have different norms, consumption choices have the potential to reveal the consumer's identity. For example, conditional on prices and income, a Muslim who consumes pork identifies less with his religion than a Muslim who abstains from pork. We explore whether this insight can help better understand identity choices. By drawing on standard and readily-available consumption data as well as well-established tools for analyzing them, we can investigate multiple determinants of identity among a large and representative population over a long period of

time. Turning the question around, we also ask whether ideas from social identity research can improve economists' understanding of consumer behavior.

Our setting is food consumption in India where, given the high levels of malnutrition, dietary choices driven by identity also have significant health implications. Two features make the Indian setting particularly suitable for implementing our approach. First, India is characterized by deep ethnic and linguistic divisions (Basu et al. 2016; Fearon 2003; Reich et al. 2009). At the same time, it is also religiously diverse, with members of the *same* ethnicity often distributed across different religions and castes. This provides well-defined sets of (potential) identities that individuals can choose from. Second, food consumption in India is associated with strong norms and taboos. In the words of noted anthropologist Arjun Appadurai (1988):

> *Food in lndia is closely tied to the moral and social status of individuals and groups. Food taboos and prescriptions divide men from women, gods from humans, upper from lower castes, one sect from another.*

The basic idea then is simple. A Hindu from Gujarat is born a member of multiple groups and hence has several possible identities. While she cannot choose to be Muslim or Tamil, she can choose whether to identify as Gujarati (one of India's many ethno-linguistic groups) or Hindu (one of India's major religions). Given the different norms and taboos across religious and ethnic groups, the food consumption bundle uncovers this identity choice. Our analysis examines how these revealed identities respond to key forces that economics and social psychology conjecture drive identity choices.

We base our analysis around a conceptual framework that we outline in Section 2. Following the identity economics literature (e.g. Akerlof and Kranton 2000), an individual h that identifies with group J suffers a utility loss if her consumption bundle is far from the prescribed or prototypical bundle of that group. However, extending Akerlof and Kranton (2000), we allow individuals to choose their identity among the multiple groups they are members of, depending on which provides the highest utility. We identify three major factors that affect this choice. From cognitive psychology and categorization theory (Nosofsky 1992; Turner et al. 1987) as well as experimental economics (Benjamin et al. 2010), we posit that if the salience of membership of a group rises, the utility from identifying with it rises.[1] From Social Identity Theory (Tajfel and Turner 1979), a high-status group provides someone identifying with it an enhanced self concept and hence greater utility. Finally, from standard price theory, the more costly it is to follow

[1] Shayo (2009) makes this force more precise, with individuals paying a higher cognitive cost for identifying with a group that is more different from them—and salience raising the attention weight they place on the distance between themselves and the group along a particular dimension, e.g. a religious or ethnic dimension.

the prescribed behaviors of a group, the less likely an individual is to identify with it.

This simple framework generates three sets of implications that we test using 350,000 detailed household consumption surveys collected as part of India's National Sample Survey (NSS) over the period 1987-2000. In Section 4, we focus on consumption patterns for the four most prominent food taboos associated with religious groups in India (the avoidance of beef, pork and alcohol, and vegetarianism), as well as two foods that display particularly strong regional preferences (rice and wheat). Then, in Section 5, we impose a particular functional form that incorporates identity choices in an Almost Ideal Demand System (Deaton and Muellbauer 1980) and estimate a demand system that covers all foodstuffs. Using these estimates, in Section 6 we quantify the relative importance of salience, status and economic costs in driving identity choices in India during the 1990s as well as the health and welfare effects of these choices.

We begin with outbreaks of Hindu-Muslim violence as shocks to the salience of Muslim and Hindu religious identities. We use a combination of cross-sectional and panel regressions exploiting variation in shocks to households of the same ethnicity-religion pair, either across districts or within districts over time. Since food prescriptions vary across groups, we can control for local supply and demand conditions by including good-district-time fixed effects in all our regressions. As our conceptual framework predicts, mounting inter-religious tensions—proxied by reports of conflict in the national media—are accompanied by increased adherence to religious taboos among Hindus and Muslims. For example, Hindus are more likely to abstain from beef and Muslims from pork. These findings hold conditional on prices and income and do not appear to be driven by availability issues or social desirability bias in reporting.

More surprisingly, Muslims consume more beef during times of inter-religious conflict, and Hindus more pork. This outcome, however, follows directly from our conceptual framework where households face a choice between religion and ethnicity. As households move away from their ethnic and towards their religious identity, they abandon the prototypical consumption bundles of their ethnicity—where, because of substantial numbers of Muslims and Hindus, both beef and pork consumption are low—in favor of more extreme religious identity bundles where other groups' taboos are ignored. We find further support for our dueling-identity framework when we consider the full demand system, with religious conflict shifting households towards the prototypical bundle of their religious group and away from that of their ethnic group. Further and more direct evidence for the relevance of ethnicity as an identity choice comes from state splits along ethno-linguistic boundaries. As these splits were preceded by demands for ethnic autonomy, we consider these events as shocks to the salience of ethnicity. Two of the three splits during our period of study lie along the geographic fault line that divides rice and wheat eaters

in India. Consistent with a shift towards their ethnic identities, as ethnic autonomy approached, rice and wheat consumption patterns diverged on either side of the future state border. And as predicted by our dueling-identity framework, religious-taboo adherence decreased.

Next, we borrow from the sociology literature (e.g. Parkin 1971; Weiss and Fershtman 1998) in proxying for group status with the returns to typical group occupations. As religious groups are over-represented in certain occupations, this proxy generates two shift-share identification strategies where we exploit either cross-district variation in local occupational returns or cross-time variation in national occupational returns. We find strong support for our theoretical prediction. A rise in the status of one's religion is associated with increased adherence to that religion's norms and taboos. Turning to the full demand system, we find that a rising relative status of one identity over another shifts consumption towards the prototypical bundle of that identity.

Finally, our framework suggests that when the local cost of adhering to a group's prescribed behavior is lower, households should be more likely to tilt their consumption towards the norms of that group. When considering prominent food taboos we explore a variant of this implication: that the endogeneity of identity is a force that generates complementarities between taboos. For example, if the price of pork rises, it is now relatively less costly for Muslim households to identify with their religion since they would forgo pork anyway. If this leads them to switch to their religious identity, their consumption of alcohol—another taboo in Islam—would decline, making pork and alcohol complements in demand. This prediction is borne out by the consumption data. We also find that demand for taboos is less own- and cross-price elastic compared to non-taboo goods. Estimation of the full demand system confirms the role of economic costs in driving identity choices more directly. We find that a household's consumption of a particular good diverges from the prototypical bundle of their religious group and shifts toward that of their ethnic group when the cost of their religious bundle rises relative to their ethnic bundle (conditioning on standard price and income effects).

Our demand system estimates allow us to quantify the relative importance of these three determinants over the tumultuous reform period 1987-2000. Our estimates suggest large shifts in identity choices. However, substantial heterogeneity across districts and religions led to only a small shift from religious to ethnic identities on net. While conflict-driven salience shocks have significant effects on identification decisions (consistent with previous literature and common narratives), quantitatively, economic cost and status were more important drivers of identity choice in this period. This is because conflict shocks are temporary—with the effects on identity fading out approximately nine months after the shock—and rather rare. In contrast, changes in prices and occupational returns are both more persistent and common.

The finding that the largest impacts arise from changes in the economic costs of identifying with a group is surprising given that this channel has received little attention in the social-identity literature. However, such a finding rationalizes recent and much-discussed efforts by the Hindu-nationalist BJP party to raise the effective price of beef through bans and legislation.[2] Our results suggest that to harden Hindu identities, such a strategy may be more effective than inciting religious violence—a tactic documented by Wilkinson (2004). This insight may extend beyond India. Since 2003, when Erdogan's Islamist AKP party came to power in Turkey, taxes on alcohol have increased steadily and now account for two-thirds of the cost of basic alcoholic drinks. Between 2003 and 2019, the price of raki—a traditional spirit—jumped by 1055 percent and the price of beer by 903 percent, while the CPI only rose by 299 percent.[3]

The paper contributes to two main (and largely disconnected) lines of inquiry: on consumption and on identity. We show that incorporating social identity into our analysis of consumer choice can enrich our understanding while maintaining well-established conceptual tools. For example, economists often assume that certain goods are complements or substitutes but lack a theory for why that is the case. It is therefore hard to understand why the *same* goods can be complements in one context and substitutes in another. Our analysis suggests that when certain goods come to be associated with certain social groups, this tends to generate complementarities. Furthermore, while several studies document how consumption is affected by individual status concerns (Charles et al. 2009; Heffetz 2011; Bursztyn et al. 2018; Bellet and Colson-Sihra 2018), we show that consumption patterns are shaped by *group* status as well. And, while conflict can obviously affect consumption patterns by destroying resources and changing relative prices, we show that it can also shift consumption by changing identity choices. A final point is that consumers may be more flexible than we usually assume. When the price of a particular group-related bundle increases, consumers can switch the norms they follow to better handle the new price environment. In this sense demand is more elastic, not with respect to any one price but rather to the price of the entire prescribed bundle.

We also make three contributions to the study of identity. First, The existing literature relies mainly on lab experiments, surveys, content analysis and ethnographic studies to measure identity (see Abdelal et al. 2009 for a review of methods). We propose a revealed preference approach to inferring identity, using naturally-occurring and widely-observed data (Bertrand and Kamenica 2018 develop an analogous approach for inferring cultural differences across groups).

[2] See, for example, https://www.nytimes.com/2017/07/11/world/asia/india-cows-slaughter-beef-leather-hindu-supreme-court-ban.html for coverage of the governments attempted ban on cow slaughter.

[3] Item level inflation between 2003 and February 2019 from the Turkish Statistics Institute (TUIK), https://biruni.tuik.gov.tr/medas/?kn=84&locale=en.

Second, while the "constructivist" literature in political science emphasizes the malleability of social identities (see Chandra 2012), most work in economics takes ethnic and religious identities as given (see e.g. Alesina and La Ferrara 2005's review of the literature on ethnic diversity). Our goal in this paper is to understand the endogenous determination of identities (see also Eifert et al. 2010; Shayo and Zussman 2011; Hjort 2014). Furthermore, due to the wide and long coverage of the data we are able to simultaneously study several of the main determinants of identification suggested by the experimental and theoretical literatures in psychology and economics. Third, previous studies of identity have not, to the best of our knowledge, estimated the effect of prices on identification decisions, which prove to be substantial.

The paper proceeds as follows. Section 2 introduces the conceptual framework and hypotheses. Section 3 describes the data. Section 4 provides reduced form evidence that salience, status and cost drive identity choices as revealed by the consumption of prominent identity goods. Section 5, derives and estimates a demand system incorporating identity using all food consumption. Section 6 quantifies the importance of the three determinants. Section 7 concludes.

2 Conceptual framework

Consider a society composed of several groups, where a member of society can belong to multiple groups at once. We use h to denote an individual consumer, J to denote a group, and i to denote a good. $x_{ih} \in [0,1]$ is h's consumption of good i, expressed as a budget share, and x_h is h's consumption bundle.

Let G_h be the set of groups to which individual h belongs, e.g., her religious group and her ethnic group. G_h is exogenously given. However, being born to a particular ethnic group does not necessarily imply identification with that group. Our focus is to understand the individual's choice of which group within G_h to identify with. Following the social identity literature, "identification" with a group has two main features. The first is conformity to group norms and behavior (Akerlof and Kranton 2000; Benjamin et al. 2010). In our case, denote by $\overline{x}_J$ the *prescribed* consumption bundle of social group J. For example, if group J has a taboo on the consumption of good i, then $\overline{x}_{iJ} = 0$. Identifying with group J then implies seeking to reduce the distance $d(x_h, \overline{x}_J)$ between own consumption and group J's prescribed behavior. We also allow for exogenous shocks to the perceived distance between individual h and each of the groups in G_h. Specifically, we use κ_{hJ} to denote shocks to the salience of h's membership in group J.[4] The

[4]This is made more precise in Shayo (2009) and Sambanis and Shayo (2013), where salience is the (potentially endogenous) change in the attention weight that individuals place on a particular dimension (e.g. their religion or ethnicity) when forming perceptions of the distance between themselves and the various groups in society. In the present paper we are interested in exogenous shifts in the salience of h's membership in different groups.

second feature of identification with a group is caring about its status (Tajfel and Turner 1979). We denote group J's status by y_J. This lends itself to the following simple definition.

Definition 1. Individual h *identifies* with group $J \in G_h$ if her preferences can be represented by:

$$U_{hJ} = U(x_h, y_J, \kappa_{hJ}; \overline{x}_J), \tag{1}$$

where U is decreasing in the distance $d(x_h, \overline{x}_J)$, and increasing in y_J, κ_{hJ}.

Let $x^*_{hJ}(p, m_h, y_J, \kappa_{hJ}; \overline{x}_J)$ be the optimal consumption bundle chosen when h identifies with group J, given a vector of prices p and income m_h. Since the consumption bundle is a function of the group one identifies with, consumption choices can in principle reveal one's identity.

As noted by Shayo (2009), the above two features—distance and status—also capture the major determinants of individuals' identification choices documented in the literature. That is, other things equal, individuals are more likely to identify with groups with higher status and with groups that are perceived as more similar to them, where perceived similarity can be affected by both the distance in consumption space and by the salience of other dimensions such as ethnicity and religiosity. This leads to the following assumption.

Assumption 1. *Identity is endogenous. Individual h identifies with group J if and only if*

$$J = \underset{J \in G_h}{argmax} V_{hJ}(p, m_h, y_J, \kappa_{hJ}; \overline{x}_J),$$

where V_{hJ} denotes h's indirect utility if he/she identifies with J:

$$V_{hJ}(p, m_h, y_J, \kappa_{hJ}; \overline{x}_J) = U(x^*_{hJ}, y_J, \kappa_{hJ}; \overline{x}_J). \tag{2}$$

Thus, the choice of identity itself responds systematically to the social and economic environment.[5] In Section 5 we derive consumer demand using a specific functional form that nests the Almost Ideal Demand System (Deaton and Muellbauer 1980).

Given the endogenous nature of identity, we expect individual h to be more likely to tilt consumption towards $\overline{x}_J$ (e.g., respect group J's taboos) when:

1. Her membership in group J is more salient (κ_{hJ} is higher).
2. The status of group J is higher (y_J is higher).
3. The cost of adhering to J's norms/taboos is lower ($p \cdot \overline{x}_J$ is lower).

[5] The sharp choice between identities—and restricting attention to cases where the set of optimal identities is a singleton—is for conceptual clarity. Empirically, we will not distinguish between a model where households make a binary choice between identities and one where they choose the relative weights they place on each identity.

A particularly interesting implication of endogenous identity formation is that it tends to generate complementarities between goods that are encouraged (or taboo) in a particular group. To clarify this mechanism, it is instructive to consider a simple example where we shut down standard substitution effects (i.e. cross-price elasticities are zero under any specific identity).

Example. There are two groups, $G_h = \{A, B\}$, and three goods $i = 1, 2, 3$. Consumer h's utility when identifying with group J is Cobb-Douglas: $U_{hJ} = \sum_i \beta_i^J \ln x_i$, where $\sum_i \beta_i^J = 1$ for all $J \in \{A, B\}$. Group A encourages (or equivalently B discourages) the consumption of both goods 1 and 2, hence $\beta_1^A > \beta_1^B$ and $\beta_2^A > \beta_2^B$. Suppose an individual initially identifies with group A, then consider an increase in the price of good 1, p_1. If identity is fixed, budget shares do not change ($x_2 = \beta_2^A$). However, it is easy to see that $\frac{\partial}{\partial p_1}(V_{hA} - V_{hB}) < 0$.[6] Thus, if identity is endogenous (Assumption 1), the individual may shift to identify with group B as A's prescriptions are too expensive. This would imply $x_2 = \beta_2^B < \beta_2^A$ and $\frac{dx_2}{dp_1} < 0$. Endogenous identity makes the two goods encouraged by group A complements.

This example provides a second prediction. If good 1 is a taboo for group B, then identifying with B can imply $\beta_1^B = 0$. In this case, as long as p_1 is sufficiently high so that the individual continues to identify with group B, the individual would be at a corner ($x_1 = 0$), and demand for the taboo good would not respond to prices.

In the framework above, the norms of the groups you identify with enter your preferences directly and so there is conformity even though there is no material incentive to do so. Such a modeling choice finds strong support in the experimental literature. For example, priming the salience of ethnic, professional or criminal group membership increases conformity although material incentives or image concerns are held constant (Benjamin et al. 2010; Cohn et al. 2014, 2015), and conformity is also observed under anonymous conditions (Burnkrant and Cousineau 1975). However, there may also be non-psychological benefits associated with a low distance $d(x_h, \overline{x}_J)$ from the prescribed behavior of group J. For example, employment and business opportunities may be affected by the degree of closeness to other members of the community. Such benefits may be larger if group J is high status. In the present paper we will not be able to distinguish between there being psychological or material benefits from respecting the group's prototypical behavior. This distinction only matters for our analysis in the welfare calculations in Section 6.3.

A final remark. While in a more general framework, group status y_J, salience κ_{hJ}, and prices p are endogenous to economic policy and to (aggregate) individual choices, in the present paper we seek to understand individual responses to exogenous shifts in status, salience and prices.

[6] The indirect utility function in this case is $V_{hJ} = B_J + \ln m_h - \sum_i \beta_i^J \ln p_i$, where B_J is a function of the β^J's.

3 Data and Context

To explore the hypotheses outlined in Section 2, we require data on the consumption patterns of households, their religious and ethnic groups, and shocks to prices, status, and salience.

3.1 Household Data

Our primary source of data is the Indian NSS Consumer Expenditure survey. Each survey provides household expenditures and quantities consumed for more than 300 goods and services, as well as economic, demographic and social characteristics. We use all three thick survey rounds—covering around 120,000 households per round—that contain both district identifiers and overlap with the conflict data we shall introduce shortly: the 43rd round (1987-1988), the 50th round (1993-1994) and the 55th round (1999-2000). We use the 1987 district boundaries to follow 419 districts (contained in 77 regions) over time.[7] The 50th round contains the exact date each household was surveyed, while the other rounds only provide the quarter. We use household survey weights throughout to make our results nationally representative.

Food consumption The three rounds of surveys allow us to track 124 consistently-categorized food products over rounds for which surveyors record both expenditure and quantity. Appendix Table A.1 lists these items by food group.

Prices From expenditures and quantities, we calculate unit values that serve as price proxies. Following Atkin (2013), to guard against outliers and quality issues, we replace household-level unit values with the median unit value in each village/urban block in that survey round.[8]

Group status As we detail in Section 4.3, our measures of group status are based on the returns to occupations disproportionately held by group members. We calculate these occupational returns from the NSS surveys using detailed occupation codes for the household's primary occupation, as well as total household expenditure which serves as a proxy for household income.

3.2 Religious and Ethnic Groups in India

Religious groups The NSS surveys record the religion of the household as well as caste (two categories, scheduled castes and others). We focus our analysis on the three largest religious groups in India—Hindus, Muslims and Christians. As Hindu castes have different taboos and prescribed behaviors, we further divide Hindus into those from lower castes ("scheduled caste"

[7] As only region and no district identifiers are available for urban households in round 50, we exclude these households from any district-level analysis.

[8] The village price is robust to outliers and not contaminated by quality effects or measurement errors which affect the price response at the household level. If no consumption of the good is reported in the village/urban block, we compute the median price at an incrementally higher level of aggregation.

in the survey) and those from upper castes ("others" in the survey), resulting in a total of four religious groups. These religious groups are present in every region of India, as shown in Appendix Figure A.1, and represent 90 percent of the total population. In contrast, the omitted groups either have very few members (Jains, Buddhists, Zoroastrians), are geographically concentrated (Sikhs in Punjab) or are not a single religious group (scheduled tribes).

Ethnic groups India is characterized by enormous ethnic diversity. This can be seen in the diversity of language groups spoken (e.g. Tamil is Dravidian, Gujarati is Indo-Aryan, Meghalayans speak Austro-Asiatic languages, Sikkimese speak Sino-Tibetan languages). It is also apparent from more recent genetic evidence which finds India to be four times more genetically diverse than Europe (Reich et al. 2009), and to comprise five separate ancestral migrations (Basu et al. 2016). This diversity is coupled with high ethnic endogamy and limited migration such that ethnicities are spatially segregated across India. In response to the demands of ethnic groups following Independence, India chose to draw its state borders along ethno-linguistic lines.[9] Thus, the state of residence of the household in the NSS surveys serves as a good proxy for ethnicity and is the measure we use in our analysis.

3.3 Religious and Ethnic Goods

Central to our approach is the idea that the consumption choices of Indian consumers can reveal their chosen identity. In the first part of our analysis (Section 4) we examine goods that are strongly associated with specific religious and ethnic identities.

Religious goods For religious identities in India, there are four particularly prominent examples: beef, pork, meat and alcohol. The avoidance of beef consumption is deeply ingrained in the Hindu population and widely seen as one of the purest practices a Hindu could accomplish.[10] Pork consumption is equally if not more taboo for Muslims.[11] Non-vegetarian food is more generally a taboo for practicing upper-caste Hindus (especially Brahmins), who see this as an act of violence that goes against their religious beliefs.[12] Lower-caste Hindus do not widely share

[9] The first ethno-linguistic State, Andhra (presently Andhra Pradesh), separated Telugu-speaking people from the State of Madras (presently Tamil Nadu) in 1953. It was formed following Sreeramulu's fatal hunger strike and subsequent public protests demanding a state to preserve the culture of the Andhra people. In 1956, the State Reorganisation Act redrew state boundaries along ethno-linguistic lines.

[10] Gandhi writes: *"Hindus will be judged not by their tilaks, not by the correct chanting of mantras, not by their pilgrimages, not by their most punctilious observances of caste rules, but their ability to protect the cow"* (*Young India*, 6 October 1921, p. 36). A few Scheduled Castes (some Dalits, formerly known as untouchables) consume beef, especially those who traditionally worked as scavengers or leather workers. Because of the enormous stigma, these groups typically stop this practice when they rise in social status, a process called *sanskritization* by Srinivas (1956).

[11] The Qur'an, Surah Al-Baqarah 2:173 says: *"He has forbidden you only carrion, blood, the flesh of the swine, and that which has been offered to other than Allah."*

[12] Several sacred texts mention eating meat as an impure and sinful act, among them the *Manusmriti* (5.48-5.52): *"One can never obtain meat without causing injury to living beings... There is no greater sinner than a man who,*

Table 1: Taboos by Religious Group

Religious Identity	Beef	Pork	Meat	Alcohol
Hindu Upper Caste	x	x	x	x
Hindu Scheduled Caste	x			
Muslim		x		x

this vegetarian norm, and in fact have developed a non-vegetarian cuisine particular to their identity.[13] Our primary definition of vegetarianism includes avoidance of eggs, poultry and fish. Appendix G repeats our analysis with a less strict definition that does not incorporate eggs and fish. The results are qualitatively similar. Finally, alcohol is a taboo shared by both Muslims and upper-caste Hindus.[14] Table 1 provides a schematic summary of these taboos.

Direct evidence for these four taboos comes from Figure 1, which reports the share of each religious group that abstains from consuming each of the four taboos, for each quarter of each NSS round. The top left panel shows that upper-caste Hindus essentially report no beef consumption, and less than 10 percent of Scheduled Castes report consuming beef, while around 40 percent of Christians and Muslims consume some. Similarly, almost no Muslims report consuming pork (top right panel) or alcohol (bottom right). Finally, Hindus are much more likely to be vegetarian than Christians or Muslims, and this is more pronounced for Upper Castes (bottom left). These differences are remarkably persistent over our sample period.

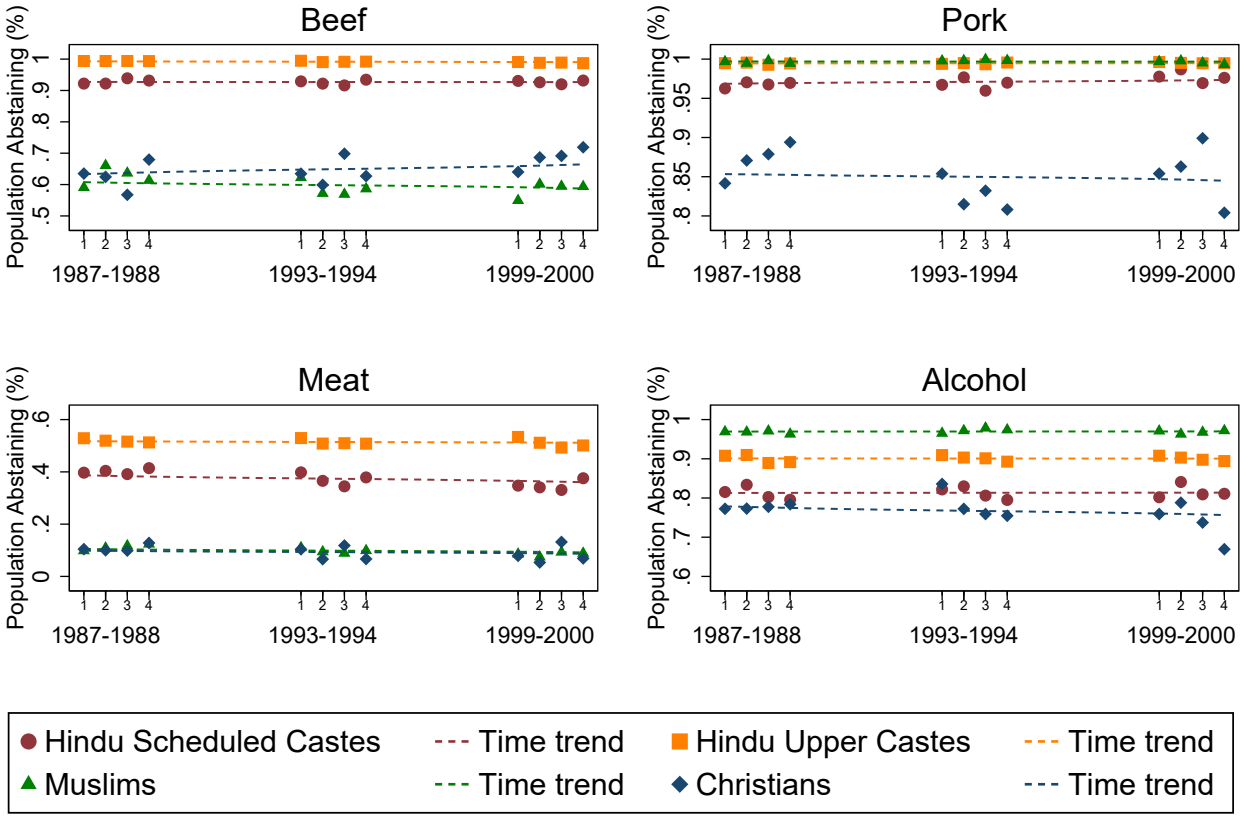
Ethnic goods Identifying ethnic goods is less straightforward, as the examples cited in the literature are less stark—in part because of the absence of formal prohibitions like those present within religions. That is not to say that there are no regional foods associated with different ethnicities. Perhaps the most prominent example is the strong regional preferences for the staple cereals rice and wheat (Chakravarti 1974). Despite the fact that these two cereals provide similar nutrition per rupee, there is dramatic regional variation in their consumption and cooking practices. For example, the relative price of rice and wheat is similar in the states of Kerala and Punjab, yet Keralites consume thirteen times more rice than wheat and Punjabis ten times more wheat than rice (Atkin 2016). In Section 4.2, we exploit the fault line bisecting India that divides rice eaters in the South and East from wheat eaters in the North and West (clearly seen in Appendix Figure A).

outside of an offering to gods or ancestors, wants to make his own flesh thrive at the expense of someone else's."

[13] Examples of such dishes can be found in a book on Dalit food, *Anna He Apoornabrahma*, written by Shahu Patole in 2015 to serve as a counterpoint to the many Brahmin vegetarian cookbooks.

[14] This is seen from the Qur'an's prohibition (Surat 5:91)—*"Satan only wants to cause between you animosity and hatred through intoxicants and gambling and to avert you from the remembrance of Allah and from prayer. So will you not desist?"*—and from the Mahabharata (Adi Parva, 76): *"a wretched Brahmin unable to resist the temptation to drink shall be considered one who's lost all virtue and considered guilty of murdering one if his own caste. Yes, he shall be hated in this and other worlds."*

Figure 1: Fraction of Population Abstaining from each Taboo, by Religion-Round-Quarter



3.4 Conflict Data

For data on episodes of Hindu-Muslim violence—a familiar feature of post-partition India—we draw on the 1950-1995 Varshney-Wilkinson Dataset, extended to 2000 by Mitra and Ray (2014). The dataset collects occurrences of Hindu-Muslim conflict that are sufficiently intense to be reported by the national press (specifically *The Times of India*, a leading national newspaper). For every episode, the dataset provides the date of incidence, the city/village, and the number of people killed, injured, or arrested.[15] The dataset records 505 Hindu-Muslim riots between 1987 and 2000, with about 4100 fatalities. Appendix Table A.2 presents descriptive statistics of recorded conflicts by state during the three rounds of consumption surveys. Our baseline measure of conflict is the occurrence of at least one outbreak of violence between Hindus and Muslims in the household's district close to the date of the survey.

[15]The dataset also records the duration and the reported proximate cause of the riot, but no information on which side initiated the violence.

4 Demand for Identity Goods

As described in Section 3.3, certain goods have long been associated with specific religious and ethnic identities. A natural first step is to explore how the consumption of these goods responds to the three forces that our conceptual framework suggests drive identity choices. Focusing on prominent identity goods—i.e. goods where external sources can corroborate their strong association with specific religions or ethnicities—prior to turning to a fully-specified demand system analysis, serves two main purposes. First, the reduced form analysis serves as a proof of concept that consumption data can be used to help understand identity choices. Second, the identification arguments for the analysis using prominent identity goods will be the same as when considering the full vector of food consumption in Section 5, and so we discuss identification in this section.

We begin by analyzing salience as a determinant of group identity. We consider two shocks to salience, conflict between religious groups and changes in political autonomy for ethnic groups.

4.1 Religious Conflict and Identity

A long line of scholarship argues that inter-group conflict hardens identities.[16] In fact, in the Indian context, it has been argued that politicians use religious violence precisely to that end (Wilkinson 2004). Through the lens of our conceptual framework, conflict along religious lines raises the salience of membership in one's religious group (κ_{hr}) as in Sambanis and Shayo (2013). Thus, religious conflict increases identification with one's religious group and tilts consumption x_h towards the prescribed behaviors of that religion $\overline{x}_r$.

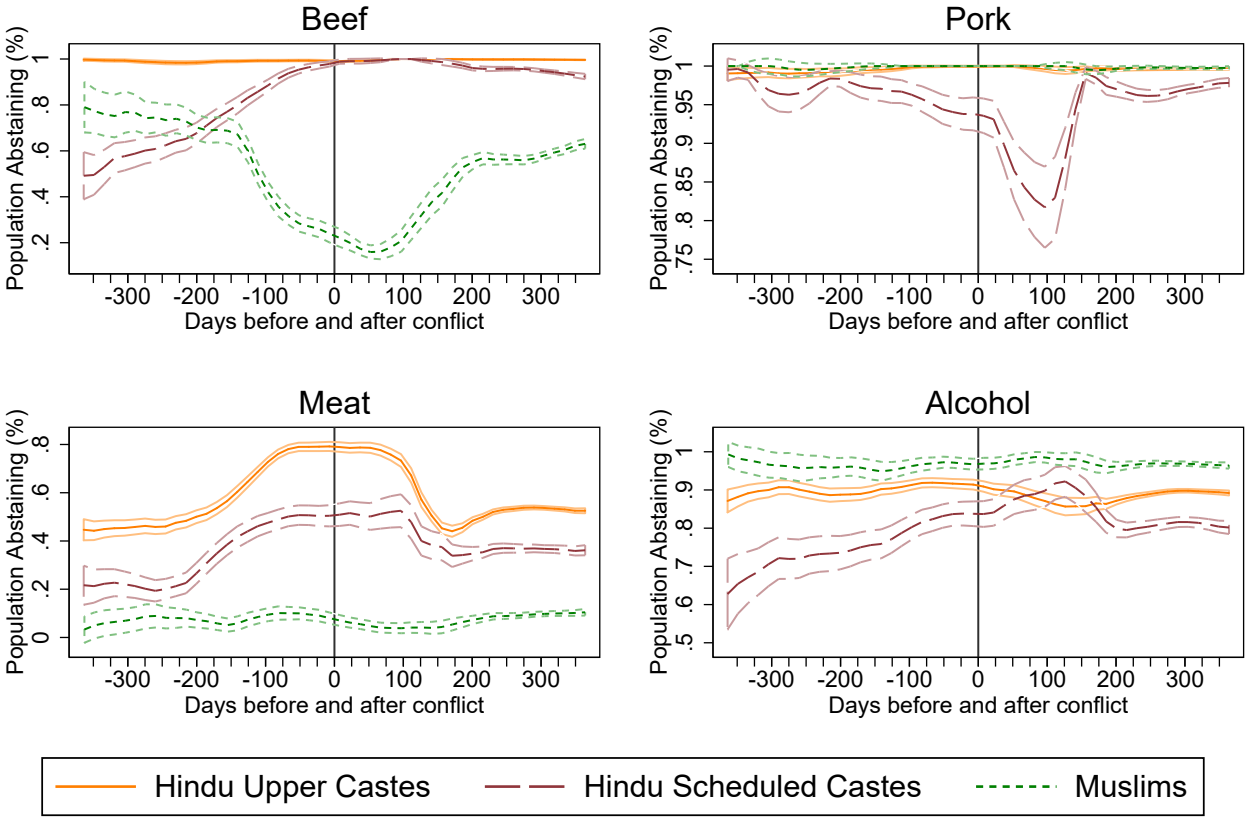
From a standard economic perspective the effects of violence on identity are not so clear-cut; while conflict can affect prices and incomes, conditional on those it should have no effect on consumption choices. Further, conflict along a particular dimension, say religious, could conceivably lead individuals to want to hide or downplay their religious identity to avoid harm (see e.g. Kalyvas 2008). To separate these different hypotheses, we turn to the consumption and conflict data described above to explore whether households abstain more from religiously-taboo goods during times of Hindu-Muslim conflict.

High-Frequency Plots from the 1993-94 Survey Round

Before turning to our district-level regression analysis—where we condition on prices and incomes, as well as on a battery of fixed effects—it is instructive to look at the raw data. Specifically,

[16] See Shayo and Zussman (2017) for a recent discussion. *Security Studie*s (13:4) contains a collection of essays debating Kaufmann (1996)'s thesis that "in ethnic wars both hyper-nationalist mobilization rhetoric and real atrocities harden ethnic identities to the point that cross-ethnic political appeals are unlikely to be made and even less likely to be heard" (p. 137).

Figure 2: Recorded Conflict and Taboo Avoidance, NSS 50th Round (1993-1994)



Notes: Kernel-weighted local polynomial regression of abstention on days before/after conflict. 95 percent confidence intervals shown with lighter shading.

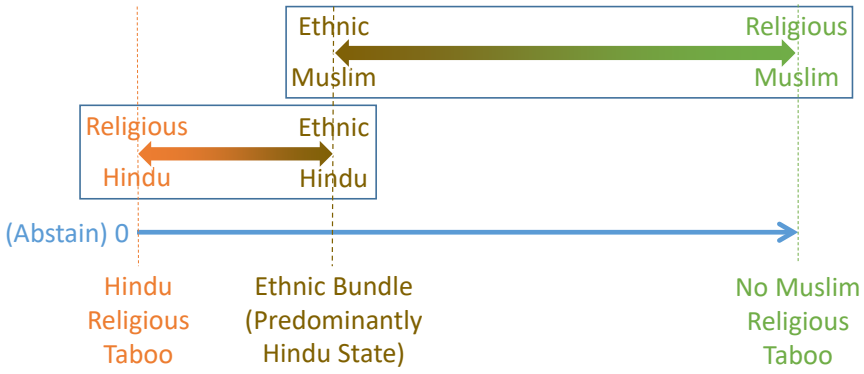
we present plots of the raw consumption data against the time since a reported religious conflict in the household's region. These reports serve as a proxy for religious tensions, and hence the salience of Hindus' and Muslims' religious identity. Figure 2 shows non-parametric regressions of a dummy for whether the household abstains from consuming a good on the number of days before or after a conflict. For this specific exercise we restrict attention to NSS round 50 (July 1993 to June 1994), since this is the only round that reports the actual day of the survey. Given the sparsity of conflicts in any particular round, and the fact that round 50 does not contain district identifiers for urban areas where most conflicts occurred, these plots focus on conflicts at the level of 77 regions in India.[17] The patterns are qualitatively similar if we condition on household per capita expenditure, local prices, and region-month fixed effects (Appendix Figure B.1.1) or restrict attention to a balanced sample of locations surveyed both in the 6 months before and after a conflict (Appendix Figure B.2.1).

[17] For each household we calculate the days before or after the first conflict occurred in their region. As we plot days since conflict over the range -365 days to +365 days, to code the date of first conflict we only consider conflicts occurring between July 1992 and June 1995.

The upper left panel of Figure 2 plots the fraction of the population that abstain from beef consumption—the most prominent religious taboo good for Hindus. While essentially no upper-caste Hindus report consuming beef, in peaceful times a considerable proportion of scheduled-caste Hindu households do not respect this taboo. However, there is a marked increase in scheduled caste Hindu households reporting zero beef consumption in the run up to a reported episode of conflict, followed by a modest decline in the following months.

Notice that these consumption patterns may, to some extent, serve as leading indicators for the eruption of conflicts. Recorded conflict—as manifested by reports that reach the national press in India—is typically preceded by mounting inter-religious tensions. It is as much these tensions that make religious identity salient as it is discrete conflicts that result from these tensions boiling over. We explore this timing further when addressing reverse causality and other identification concerns in the regression analysis below.

The timepath of beef consumption for scheduled-caste Hindus suggests that religious identity is indeed strengthened by conflict. However, our framework—in which households choose between *two* possible identities (religious and ethnic)—has additional implications, compared to a model where households simply choose how religious to be and choose to become more religious at times of conflict. For example, beef is not a taboo for Muslims, and so the prescribed behavior for someone who identifies as Muslim allows for high beef consumption. In contrast, the average household in most States of India consumes little or no beef and so the typical *ethnic* cuisine is characterized by little beef consumption. These two prototypical behaviors are indicated as points on the horizontal axis in the graphic below. If religious tensions lead Muslims to identify more with their religious identity, this means they identify less with their ethnic identity, with the implication that beef consumption should rise as they move between the two groups' prescribed behaviors.



This prediction is strongly supported by the timepath of Muslim beef consumption (upper left panel of Figure 2). Beef avoidance drops markedly around the time of conflict before rising

back up in subsequent months.

The upper right panel of Figure 2 shows pork avoidance (a strong Muslim taboo). We see the mirror image of the above pattern. Muslims essentially never consume pork, nor do upper-caste Hindus, presumably due to the meat taboo we will turn to next. However, scheduled-caste Hindus, for whom pork is not a taboo, markedly increase their pork consumption around the time of conflict, while mostly avoiding it in peaceful times.

Finally, the bottom two panels of Figure 2 show timepaths for the two less-adhered to taboos, abstention from animal protein (vegetarianism), and alcohol avoidance. There are pronounced increases in vegetarianism among both lower- and upper-caste Hindu groups around the time of conflict, while patterns are more murky for alcohol avoidance.

In the regression analysis we confront concerns that changes in availability or social desirability bias may be behind our findings. Here we simply note that such stories do not explain why beef consumption rises for Muslims and pork consumption rises for lower-caste Hindus in times of conflict. And as shown in Appendix Figure B.2.3, we do not see a similar divergence in Muslim-Hindu consumption patterns for chicken and mutton that would typically be purchased at the same butchers.[18]

Combining Taboos and Survey Rounds

We now turn to a more formal econometric analysis at the quarterly level that combines all four taboo goods and all three survey rounds. The longer sample period and larger sample size allows us to focus on district-level conflict (the most disaggregated geographic unit that can be followed across rounds), and to include Christians in our analysis—the smallest of our religious groups and one that has no taboos with respect to the four goods.

We run the following difference-in-differences specification, variants of which will also be used to assess the effects of changes in costs and status. For household h, of religion r , in district d, ethnicity (state) s, in round-quarter t, consuming good i:

$$\begin{aligned}\text{Abstain}_{ihrdst} = \alpha_1 \text{Taboo}_{ir} + \alpha_2 \text{Conflict}_{rdt} + \alpha_3 \text{Taboo}_{ir} \times \text{Conflict}_{rdt} \\ + \sum_j \gamma_{1ij} \ln price_{jht} + \gamma_{2i} \ln realfoodexp_{ht} + \delta_{idt} + Additional_FE + \epsilon_{iht}, \quad (3)\end{aligned}$$

where Abstain_{ihrdst} is an indicator variable that takes the value 1 if the household does not consume good i; Taboo_{ir} is an indicator for good i being a taboo for religion r (from Table 1); and

[18] Muslims are massively over-represented in the butcher trade, comprising 45 percent of butchers but only 12 percent of the Indian population (Appendix Table B.3.6). Thus, if availability or fear of travel were driving our results, we would expect differential responses for Muslims and lower-caste Hindus. Relatedly, Appendix Figure B.2.2 finds similar patterns splitting the sample into locations with high and low levels of Religious Fractionalization, a robustness test the logic of which we discuss in more detail below.

Conflict$_{rdt}$ is an indicator for Hindu/Muslim conflict in the district at time t (or potentially a vector of lags and leads of conflict). We code Conflict$_{rdt} = 0$ for Christian households.

The key coefficient of interest is α_3, which captures how consumption differences across taboo and non-taboo goods vary with conflict. To take a concrete example, we expect Hindus to be more likely to abstain from beef than from pork ($\alpha_1 > 0$). Conflict may also reduce the likelihood of abstaining from goods that are taboo for *other* religions—but not one's own (captured by a negative α_2). Indeed our model makes such a prediction, discussed above in the context of Muslim beef consumption. Alternatively, forces outside our model may drive a relationship between consumption and conflict that we wish to control for through including a main effect of conflict. Finally, a positive α_3 implies that conflict further increases abstention from the taboo good relative to non taboo goods.

As we are estimating the determinants of demand, we also include standard controls for prices and incomes: $\ln price_{jht}$ is the village median price of good j that controls for own- and cross-price effects; and $\ln realfoodexp_{ht}$ is the log of per-capita food expenditure deflated by a Stone price index that controls for income effects. To control for local supply and demand conditions not adequately captured by prices, we further include good-district-time fixed effects δ_{idt}. These fixed effects also absorb other factors that affect consumption of good i in district d at time t and may correlate with conflict.[19] Thus, we explore how taboo/non-taboo consumption differences between religions *within* a location vary with conflict. We cluster standard errors at the rdt level.[20]

The remaining worry is that the inter-religious tensions that conflict is proxying for are correlated with omitted variables that affect abstinence and vary at the religion-ethnicity level or below. To further alleviate these identification concerns, we pursue two additional fixed-effect strategies that we will use throughout the paper. Both strategies control for religion-ethnicity fixed effects, so that the coefficients of interest are identified only from variation within households that are choosing between the same two identities (e.g. Muslim Gujaratis). Thus, the fixed effects absorb the taboo indicator in both strategies.

Cross-Sectional Identification ($Additional_FE = \delta_{irst}$): here we include good-religion-ethnicity-time fixed effects to control for temporal shocks to adherence at the religion-ethnicity level. For example, there may be a health campaign in Gujarat to reduce alcohol consumption among lower-caste Hindus. The fixed effects ensure that we are identifying effects from variation across

[19] We do not include controls for household characteristics such as demographics and primary occupation as these are potentially endogenous to identity choices. For completeness, Appendix F replicates all the key tables in this and subsequent sections including controls for household characteristics, and finds qualitatively similar results.

[20] Results are robust to higher spatial clustering at the religion-region-time level (see Appendix Table B.3.1).

Table 2: Religious Conflict and Taboo Adherence

	LHS Variable: Abstain from Consuming Good i					
	Baseline	Cross-section	Panel	Panel		
	(1) All	(2) All	(3) All	(4) All	(5) Urban	(6) Rural
taboo=1	0.155***					
	(0.00217)					
conflict +/- 6 months	-0.102***	-0.0215	-0.0445*			
	(0.0227)	(0.0222)	(0.0268)			
taboo=1 × conflict +/- 6 months	0.0920***	0.0275***	0.0358***			
	(0.0127)	(0.00722)	(0.00714)			
conflict past (6 months)				-0.0243	0.0330	-0.145***
				(0.0226)	(0.0291)	(0.0293)
conflict present/future (6 months)				-0.0403	-0.0177	-0.0919
				(0.0381)	(0.0264)	(0.0690)
taboo=1 × conflict past (6 months)				0.0379***	0.0337**	0.0324***
				(0.00852)	(0.0172)	(0.0103)
taboo=1 × conflict present/future (6 months)				0.0251**	0.0466**	0.00873
				(0.0100)	(0.0186)	(0.0127)
Observations	1,115,640	1,115,292	1,114,116	1,114,116	347,556	764,344
Adjusted R^2	0.540	0.576	0.594	0.594	0.615	0.602
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No	No	No	No
religion*state*product*district*quarter	No	No	Yes	Yes	Yes	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. Conflict is an indicator for at least one occurrence of Hindu-Muslim conflict in the district. Columns 1-3 consider a conflict occurrence in the six months before or after the household is surveyed. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Columns 4-6 differentiate the effect of a conflict occurrence in the six months preceding the quarter of the survey, and the six months covering the quarter of the survey and the subsequent quarter. Column 5 restricts the analysis to the urban population, and column 6 to the rural population. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * p < 0.10, ** p < 0.05, *** p < 0.01.

districts with and without conflict within a particular state, religious group, and time period.

Panel Identification ($Additional_FE = \delta_{irsdq}$): here we include good-religion-ethnicity-district-quarter of year fixed effects to control for omitted variables that generate persistent deviations in abstinence across religious groups at the district level. For example, Muslims in Uttar Pradesh may be more observant of the alcohol taboo in districts containing important mosques, and religious tensions may also be more common in those locations. The fixed effects ensure that we identify effects from variation in conflict across survey rounds within the same religion-district-quarter (with the quarter controlling for seasonality in consumption). As this omitted-variable concern is the more serious of the two, we consider this our preferred specification.

Table 2 reports the results of estimating equation (3) without any of the additional fixed effects (column 1), with the cross-sectional fixed effects (column 2), and with the panel fixed effects (column 3). Recall from our discussion of Figure 2 that conflict serves as a proxy for reli-

gious tensions, and that tensions rise before conflict reports reach the national press (something we explore further below). Accordingly, in Columns 1-3 we code Conflict$_{rdt} = 1$ if any conflict incident occurred in the 6 months before or after the quarter in which the survey was administered. Appendix Table B.3.3 justifies this choice by including all four quarterly leads and all four quarterly lags in our preferred panel specification. The coefficients on the taboo interactions are only significant for conflict shocks in this +/- 6-month period.

Our main interest is the interaction term in the third row. Across all three specifications, religious conflict increases the adherence to the taboos prescribed by the household's religion. The effect sizes are substantial. An incident of conflict increases the proportion of households adhering to their religious taboos by 3.6 percentage points in our preferred panel specification (column 3) relative to the effect on non taboo goods. The estimate from the cross-sectional fixed-effects specification is similar in both size and significance, despite drawing on very different variation.[21] Effect sizes are larger in the baseline specification which does not include the additional fixed effects (column 1). In that specification, the main taboo effect is not absorbed, which shows that abstinence is 15.5 percentage points higher if the good is a taboo for the household's religion. This difference increases by two-thirds during times of conflict. Finally, and consistent with the patterns in Figure 2, the coefficients on conflict in the second row suggest that households are less likely to respect the taboos of other groups in times of conflict.[22]

Appendix Table B.3.2 repeats the analysis using a non-binary measure of conflict, log fatalities, with similar results. There is also evidence that the effect of conflict spills over to nearby districts, at least when exploiting cross-sectional variation (see Appendix Table B.3.5).

Reverse Causality, Availability Issues and Social Desirability Bias

There are three potential confounds to interpreting these results as evidence for our framework. The first worry is reverse causality. Shocks to determinants of identity at the $rsdt$ level other than salience—such as changes in group status or the economic costs of identifying with a group that we explore later—may lead to identity changes which both change taboo-good consumption and affect conflict incidence. We partially deal with this issue in Section 5 by jointly examining the effects of all three forces in our conceptual model. Here, although we do not have random shocks to conflict that would explicitly rule out reverse causality in the case of Hindu-Muslim violence, we can dig deeper into the time structure of the conflict effect to provide some reassurance. The basic idea follows Mitra and Ray (2014)'s point that "Hindu-Muslim riots are

[21] Figure 2 showed larger unconditional effects for beef abstention using the date of survey that is only available in NSS 50. Appendix Table B.3.4 finds similarly large effects using the regression analysis restricted to beef in NSS 50.

[22] Note that, given the δ_{idt} fixed effects, this coefficient is identified from differences between Hindus and Muslims (whose group experiences religious conflict) and Christians (whose group is not part of the conflicts in our data).

primarily an urban phenomenon". Thus, for *rural* households, a conflict recorded in their district is less likely to reflect the culmination of local tensions that have been simmering for some time (where taboo adherence should precede conflict if local tensions raise salience) and more likely to reflect exogenous urban forces (where taboo adherence should follow conflict if urban conflict raises salience in adjacent rural areas).

Column 4 of Table 2 breaks the conflict variable into two: conflicts that occurred in the six months preceding the quarter of the survey, and conflicts that occurred the six months that include the quarter of the survey and the subsequent quarter. Overall, both interactions show a positive association with taboo adherence. This is also the case when focusing on urban households (column 5). However, this is not the case when focusing on rural households (column 6), where records of religious conflict are associated with increased religious identification in the months *following* the recorded conflict, but not before. This finding for rural areas is inconsistent with the simple reverse-causation worry outlined above where increasing religiosity—and the associated adherence to religious taboos—drives conflict and thus precede it.

The second potential confound is that taboo abstention, such as less beef consumption by lower-caste Hindus during times of conflict, could simply be a matter of availability. For example, Hindus may no longer feel safe traveling to Muslim areas to purchase beef. Such availability issues should at least partly be captured by our price controls and the good-district-time fixed effects. However, to address this concern more directly, we use information from the NSS surveys on the religion and location of butchers in India in the three survey rounds.[23] If Hindu/Muslim conflict prevents Hindus from patronizing Muslim butchers and vice versa, we should expect this availability bias to be stronger in locations with few butchers of alternative religions. To explore this hypothesis, Table 3 adds interactions with the share of non-Hindu/Muslim butchers in that region to the specification in equation (3). Column 1 uses the cross-sectional identification strategy and column 4 shows our preferred panel approach. In both, the triple interaction is insignificant, and flips sign across identification strategies.

A second way to explore the availability issue is more indirect but exploits more localized variation. Specifically we now add to equation (3) interactions with religious heterogeneity, calculated at the village or urban neighborhood level.[24] If availability is driving the conflict effect, we would expect magnitudes to be largest in places where people of different religions live in separate neighborhoods. In these places, purchasing goods taboo to your religion during times

[23] We define a butcher as a household where the primary industry is "retail trade in meat, fish and poultry" or "slaughtering, preservation or preparation of meat", and/or the primary occupation is "butchers and meat preparers". There are 1342 butchers in our data. Appendix Table B.3.6 shows butcher numbers and shares by religion.

[24] We use First-stage Sampling Unit identifiers that correspond to villages in rural areas and urban blocks in urban areas. Ten households are surveyed within each FSU.

Table 3: Conflict and Religious Composition

	LHS Variable: Abstain from Consuming Good i					
	Cross-section			Panel		
	(1)	(2)	(3)	(4)	(5)	(6)
taboo × conflict	0.0290***	0.0335***	0.0334***	0.0354***	0.0375***	0.0402***
	(0.00739)	(0.00858)	(0.00797)	(0.00727)	(0.00962)	(0.00896)
taboo × conflict × non hindu/muslim butcher share	-0.0434			0.0360		
	(0.167)			(0.0509)		
taboo × conflict × religious fractionalization		-0.0275			-0.00342	
		(0.0330)			(0.0336)	
taboo × conflict × hindu/muslim fractionalization			-0.0312			-0.0213
			(0.0370)			(0.0374)
Observations	1,107,484	1,115,292	1,115,292	1,106,292	1,114,116	1,114,116
Adjusted R^2	0.577	0.576	0.576	0.594	0.594	0.594
main effects and double interactions	Yes	Yes	Yes	Yes	Yes	Yes
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	Yes	Yes	Yes	No	No	No
religion*state*product*district*quarter	No	No	No	Yes	Yes	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. Conflict is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed. Columns 1 and 4 interact the share of non-Hindu/Muslim butchers in the region. Columns 2 and 5 interact the religious fractionalization in the neighborhood. Columns 3 and 6 interact the Hindu-Muslim fractionalization in the neighborhood. See text for a detailed description of each measure. Columns 1-6 include the baseline fixed effects. Columns 1-3 add the fixed effects for cross-sectional identification. Columns 4-6 add the fixed effects for panel identification. All regressions include the main effects and interactions of taboo, conflict and religious composition. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * p $<$ 0.10, ** p $<$ 0.05, *** p $<$ 0.01.

of conflict may necessitate traveling to hostile neighborhoods.

We use two variants of the widely-used fractionalization index, which is the probability that two randomly-selected households from a given neighborhood do not belong to the same religion.[25] "Religious Fractionalization" (third row in Table 3) is calculated using all seven religions available in the NSS data (treating all Hindus as one religion). "Hindu-Muslim Fractionalization" (fourth row) is calculated using only the Hindu and Muslim populations. In none of the four cases is the triple interaction with $\text{Taboo}_{ir} \times \text{Conflict}_{rdt}$ close to significant, although the sign is negative. This suggests that if any availability issues are present they are likely small.

The final worry is that households under-report taboo behaviors and do so more during conflict. One possibility is fear of retribution from the local community if someone finds out. Such fears may be elevated in times of conflict if one thinks that *other* members of the community care more about respecting religious taboos at such times (that is, others identify more religiously). Like the availability issues above, households may be more worried about social sanctions if they truthfully report taboo consumption in local communities that are more homogeneous. We reject this hypothesis above when discussing availability. Another possible reason for under-reporting would be a heightened desire to please the surveyor—essentially elevated so-

[25] $Fractionalization = 1 - \sum \pi_r^2$ where π_r is the proportion of people in religion r in the neighborhood.

cial desirability bias—in times of conflict. This could generate the type of behavior we observe if the NSSO sent Muslim surveyors to Muslim households and Hindu surveyors to Hindu households. The NSSO is careful to avoid such an outcome,[26] but we have no information about the interviewer to directly explore this possibility (and we note that such heightened desire to please might itself be a manifestation of stronger religious identification). This explanation would also not account for the increase in consumption of non-taboo goods during conflict (e.g. more Muslims report eating beef during conflict).

As emphasized above, accounting for local supply and demand conditions—using both good-district-time fixed effects as well as prices and incomes—is crucial to identify the effect of conflict on the consumption of taboo goods. Our analysis, however, can also shed light on the effect of conflict on local markets. Specifically, our results suggest a shift towards religious identity at times of religious conflict, hence decreasing the demand for taboo goods among those whose religion forbids those goods, but possibly increasing it among those whose religion does not. As long as supply is not perfectly elastic, any change in net demand should affect prices. As we show in Appendix Table B.3.7, a 1-percentage-point increase in conflict-induced abstention rates leads to a 0.84 percent drop in prices.[27]

4.2 Ethnic Struggles and Identity

We now explore the ethnic dimension more directly by using shocks to ethnic salience due to changes in political autonomy. As noted in Section 3, in 1956 Indian state borders were drawn along ethno-linguistic boundaries under the States Reorganization Act. However, in some cases similar ethnicities were grouped together within the same state. In the subsequent years, many of the states that still contained substantial ethnic divisions split into smaller units. Such splits were often disputed and preceded by political campaigns and demands for "azaadi" (liberty) that emphasized ethnic distinctions.

While no state splits occurred within our sample period, three splits took place in November 2000, five months after the end of the 55th survey round. Predominantly Chhattisgarhi-speaking areas of the state of Madhya Pradesh were carved out to form the new state of Chhattisgarh; what was once the kingdom of Jharkand was separated from Bihar; and the former kingdoms of Garwhal and Kumaon were split from Uttar Pradesh to became the new state of Uttaranchal (now called Uttarakhand). Again, these splits were preceded by political activism and agitation,

[26] Enumerators are sent to villages selected at random and survey 10 randomly-selected households in that village. To further ensure impartiality, the NSSO randomizes between two sets of survey teams, one hired by the NSSO headquarters and one by the state NSSO office, and then checks for discrepancies.

[27] We regress log prices at the good-district-time level on the fraction of population abstaining, instrumented by predicted abstention from estimating equation (3). Our preferred specification (column 4) also includes conflict as an independent variable but we find no direct effect beyond its effect through abstinence. See Table B.3.7 for details.

and were the outcome of drawn-out political processes. For example, the 1990s saw region-wide strikes and rallies in Chhattisgarh, and in several election cycles the local BJP party in Jharkand ran on a platform of statehood. Thus, it is reasonable to hypothesize that for people living in these states, the salience of their ethnicities rose sharply during the 1990s.

As shown in Appendix Figure A, two of these splits fall along the fault line documented in Section 3.3 that divides wheat and rice eaters in India: Chhattisgarh and Jharkand are predominantly rice consumers, while Madhya Pradesh and Bihar tilt towards wheat. Thus, for these two state splits, we can observe whether the increasing salience of their ethnic identity in the run up to the state splitting led to greater consumption of their ethnic staple.

We run the following specification for household h in location d (either one of 80 districts or 10 regions, see below), ethnicity (i.e. future state) s, in round-quarter t, consuming good i:

$$\begin{aligned}\text{CerealShare}_{ihdst} = \theta_1 \text{EthnicCereal}_{is} \times \text{Round}_{93-94} + \theta_2 \text{EthnicCereal}_{is} \times \text{Round}_{99-00} \\ + \sum_j \gamma_{1ij} \ln price_{jht} + \gamma_{2i} \ln realfoodexp_{ht} + \delta_{its_o} + \delta_{idq} + \epsilon_{iht}, \quad (4)\end{aligned}$$

where $\text{CerealShare}_{ihdst}$ is h's share of expenditure spent on $i \in \{\text{rice, wheat, other cereals}\}$, EthnicCereal_{is} is an indicator variable that takes the value 1 if cereal i is the ethnic cereal in future state s (based on which side of the fault line the future state lies on), and $\text{Round}_{\text{xx-yy}}$ are survey round dummies. Thus, we explore how differences in consumption patterns on either side of the future border evolve as the year of the state split approaches. As before we include price and real-expenditure controls, as well as good-time-original state fixed effects δ_{its_o} to control for local supply and demand conditions in each period at the level of the existing state. Finally, we include good-location-quarter of year fixed effects δ_{idq} which serve the same purpose as the panel fixed effects above by controlling for persistent consumption differences across different locations. Standard errors are clustered at the dt level.

Note that the δ_{idq} fixed effects absorb an $\text{EthnicCereal}_{is} \times \text{Round}_{87-88}$ dummy. Hence, θ_1 reveals the increase in consumption of the ethnic cereal in the future state between the 43rd and 50th round, and θ_2 reveals the increase between the 43rd and 55th round. If the increasing salience of ethnicity in the run up to statehood leads to an increase in households identifying with their ethnic identity and hence consuming their ethnic cereal, we would expect $\theta_2 > \theta_1 > 0$. Similarly, if the increase in the salience of ethnicity is more pronounced for people living close to a new ethnic border, we would expect θ_2 to be higher in those areas.

In order to explore the evolution of consumption across survey rounds, Table 4 first presents a region-level analysis (recall that the round 50 surveys do not contain identifiers for urban

Table 4: Ethnic Goods and State Splits Along Ethnic Lines

	LHS Variable: Share Spent on Cereal i				
	(1) All Regions	(2) Border Regions	(3) All Districts	(4) Border+Neighbor Districts	(5) Border Districts
Ethnic Cereal × 1987-1988	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
Ethnic Cereal × 1993-1994	0.0300**	0.0374***			
	(0.0130)	(0.0135)			
Ethnic Cereal × 1999-2000	0.0691***	0.0760***	0.0553***	0.0623***	0.0929***
	(0.0124)	(0.0122)	(0.00858)	(0.0129)	(0.0180)
Observations	128,023	70,379	93,114	39,710	23,730
Adjusted R^2	0.732	0.772	0.793	0.830	0.836
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes
oldstate*round*quarter*product	Yes	Yes	Yes	Yes	Yes
region*quarter*product	Yes	Yes	No	No	No
district*quarter*product	No	No	Yes	Yes	Yes

Notes: Dependent variable is the share of cereal i (rice, wheat or other cereals) in total cereal expenditure. Ethnic Cereal is an indicator variable that takes the value 1 if cereal i is the ethnic cereal in future State. 1987-1988, 1993-1994 and 1999-2000 are round dummies with the initial round 1987-1988 as the reference group. Columns 1-2 are region-level regressions: column 1 includes all regions and column 2 restricts to regions along the future state border. Columns 3-5 are district-level regressions: column 3 includes all districts, column 4 restricts to border and border-adjacent districts, and column 5 to border districts. All regressions include the baseline fixed effects controlling for local supply and demand conditions (original state-time-product) and the fixed effects for panel identification (region- or district-quarter-product). Robust standard errors clustered at region-round-quarter (columns 1-2) or district-round-quarter (columns 3-5) in parentheses. Regressions weighted by survey population weights. * p < 0.10, ** p < 0.05, *** p < 0.01.

districts so the $\text{EthnicCereal}_{is} \times \text{Round}_{93-94}$ coefficient from a district-level analysis conflates changes over time with differences between urban and rural households). Consistent with the hypothesis above, column 1 shows that the share of the ethnic cereal rose by 3 percentage points between the 1987 and 1994 (from a baseline of 52 percent), and by a further 3.9 percentage points between 1994 and 2000. Appendix Table C.1 shows that a stronger result also holds—households on the Chhattisgarh and Jharkand (southeast) sides of the future state borders increased their rice consumption as the November 2000 state splits approached, and at the same time households on the northwest side of the border increased their wheat consumption.

We further analyze these changes by focusing on households living close to the future border. We expect districts either side of the border to be more similar and so there are fewer concerns with omitted variables in this specification. In addition, we would expect larger effect sizes in these locations. First, because, as mentioned above, the state split is likely to be even more salient for households living close to a new ethnic border. Also, rice/wheat consumption was more mixed in these locations (see Figure A), hence there was more scope to move towards the ethnic cereal. We find support for these predictions using either the region-level regression or a more disaggregated district-level specification (with the trade-off that the district-level analysis excludes round 50 households for the reason discussed above). The increase in ethnic cereal consumption between 1987 and 2000 grows by 0.7 percentage points between column 3 to column 4 of Table 4 as we restrict attention to households living in border districts and those

districts adjacent to border districts (and similarly comparing all regions to border regions in columns 1 and 2). The effect size further increases by a sizeable 3 percentage points when we focus only on border districts in column 5. For border districts, the ethnic cereal share rose by 9.3 percentage points over this period of rising ethnic salience.

There are two potential confounds. First, such results may be driven by the imposition of trade barriers that would naturally lead to a greater availability and a lower price for the locally-produced cereal (which is likely to correspond to the ethnic cereal, as discussed in Atkin (2013)). Here we are helped by the fact that we explicitly control for prices and, more importantly, that our sample period *precedes* the actual state splits and thus predates the imposition of any trade barriers. We can also rule out anticipatory supply responses, i.e. farmers shifting production to the ethnic good in expectation of a future border thus increasing supply and lowering prices. Appendix Table C.2 shows that the relative price of the ethnic good *rose* rather than fell in the run up to the state split.[28]

Second, there may be substantial migration during this period, as households move to ensure they are on the same side of the new border as their co-ethnics, and these movements alone may be behind the changing consumption patterns. The NSS data sheds light on such movements by asking households whether they moved location and, if so, whether they moved district within the same (original) state. As shown in Appendix Figure C.1, while cross-district movements did increase between 1987 and 2000, these changes were actually less pronounced in the border regions, making such an explanation for our results implausible.

Returning to our conceptual framework, we can also explore the effect of an increase in ethnic salience on the consumption of *religious* taboo goods. As households move towards their ethnic identity they should move away from their religious identity. To explore this prediction we adapt equation (4) to the demand for taboo goods (see Appendix Table C.3).[29] Consistent with our framework, the results suggest that households adhere less strongly to their religious taboos as a state split approaches, especially in border districts (although coefficients are only significant in the latter case).

Taken together with the results from section 4.1, the evidence strongly suggests that shocks to the salience of one's membership of a religion or an ethnicity—proxied by religious violence and demands for ethnic autonomy—drive identity choices as revealed through consumption choices over prominent identity goods.

[28] We run the same regression specification as in equation (4) but now replace the dependent variable with log prices at the good-district-quarter-round level (and removing the price and income controls): $\ln\text{Price}_{idt} = \theta_1 \text{EthnicCereal}_{is} \times \text{Round}_{93-94} + \theta_2 \text{EthnicCereal}_{is} \times \text{Round}_{99-00} + \delta_{its_o} + \delta_{idq} + \epsilon_{idt}$.

[29] Our dependent variable is now abstention from one of the four religious taboos introduced in section 3.3. On the right-hand side, we interact an indicator for the good being a taboo with the round dummies.

4.3 Status and Identity

The second widely-discussed determinant of group identity is the status of the groups. Status is a central group characteristic in both theoretical and empirical research on social identity and inter-group relations. The basic argument is that low group status results in unfavorable comparisons between the ingroup and relevant outgroups, leading people to identify more with high status groups (see e.g. Ellemers et al., 1999 for discussion and Bettencourt et al., 2001 for a meta analysis). To test whether group status affects identification decisions as revealed through consumption choices, we return to the four religious taboos considered above. We run the same specification as in equation (3), but now replace conflict with Status_{rdt} which measures the status of religious group r in district d at time t:

$$\begin{aligned}\text{Abstain}_{ihrdst} = \alpha_1 \text{Taboo}_{ir} + \alpha_2 \text{Status}_{rdt} + \alpha_3 \text{Taboo}_{ir} \times \text{Status}_{rdt} \\ + \sum_j \gamma_{1ij} \ln price_{jht} + \gamma_{2i} \ln realfoodexp_{ht} + \delta_{idt} + Additional_FE + \epsilon_{iht}. \quad (5)\end{aligned}$$

As in the conflict regressions, we control for price and income effects, as well as for local supply and demand conditions via δ_{idt}. This also addresses the possibility that the local wealth distribution may directly affect supply and hence consumption (for example, by affecting the number of shops selling a particular identity good). And as before, we present results using both the additional cross-sectional fixed effects δ_{irst} that deal with temporal shocks to adherence at the religion-ethnicity level, and the panel fixed effects δ_{irsdq} that deal with persistent deviations in abstinence across religious groups at the district level.

We follow the sociology literature (e.g. Parkin 1971; Weiss and Fershtman 1998) by proxying for group status with the returns to the occupations that are typical of the group.[30] We consider a group as relatively high status in a location if its members in that location are in relatively highly-paid occupations. However, simply running equation (5) with status measured as the group's local occupational returns raises several endogeneity problems not fully addressed by the battery of fixed effects. First, identity choices may drive the local occupational mix, resulting in reverse causation (for example, if I identify as an upper-caste Hindu I may choose not to work as a butcher). Second, identification patterns may directly affect local occupational returns through, for example, productivity benefits from stronger ethnic-, caste- or religion-based business networks, or through ingroup bias and discrimination. This again can lead to reverse causation. We therefore employ two measures of status that address these concerns.

[30] In a recent review article, Connelly et al. (2016) state that "within sociology, there is a long-standing recognition that in industrialised societies, occupations are the most powerful single indicator of levels of material reward, social standing and life chances".

Status$_{rdt}^{national_occ(r)}$: This measure draws on the fact that in India, different religions and castes are over-represented in certain occupations. Appendix Figure D.1 documents the substantial (and statistically significant) heterogeneity in occupational shares across religions. Thus, our first status measure combines the national occupation shares by religion with cross-district variation in local occupational returns:

$$\text{Status}_{rdt}^{national_occ(r)} = \sum_{o} \log w_{odt} \theta_{od^-rt},$$

where w_{odt} are real returns in occupation o in district d at time t. We use the 107 2-digit occupation codes o consistently recorded across survey rounds.[31] θ_{od^-rt} is the national occupation share in religious group r, where the occupation shares are calculated leaving out own district, which we denote by d^-. When coupled with the cross-sectional fixed effects δ_{irst}, variation in local occupational returns drives status differences across districts for a particular religious group and state. For example, if demand in one district is particularly strong for shoes and leather goods—an occupation dominated by scheduled-caste Hindus at the national level—this raises the status of identifying as a scheduled-caste Hindu in that district. This approach directly tackles the first endogeneity concern, that the local occupational mix is driven by identity choices.

Status$_{rdt}^{national_w(o)}$: The second status measure exploits changes in national returns to different occupations. Appendix Figure D.2 documents the substantial heterogeneity in the growth of occupational returns that provides the key variation over our sample period. We use a standard Bartik shift-share that combines cross-round variation in national occupational returns with initial local occupation shares by religion:

$$\text{Status}_{rdt}^{national_w(o)} = \sum_{o} \log w_{od^-t} \theta_{odrt_o},$$

where w_{od^-t} are national occupational returns leaving out own district, and θ_{odrt_o} are occupation shares by religion-district-quarter in the initial survey round t_0. Coupled with the panel fixed effects δ_{irsdq}, identification comes from status changes across rounds within a religion-district-quarter driven by national changes in occupational returns. For example, if there is an increase in the national returns to weaving due to increased global demand for Indian handicrafts, this will raise the status from identifying as a Muslim within districts that initially had many Muslim weavers. This approach deals with the second endogeneity concern above, that identity choices drive local occupational returns.

Table 5 presents the results. Columns 1-3 use the Status$_{rdt}^{national_occ(r)}$ measure for the three

[31] Absent reliable wage data, we use the total per capita expenditure of households with primary occupation o deflated by the all-India CPI.

Table 5: Status and Choice of Identity

	LHS Variable: Abstain from Consuming Good i					
	Baseline	Cross-section	Panel	Baseline	Cross-section	Panel
	(1)	(2)	(3)	(4)	(5)	(6)
taboo=1	-0.0380*			-0.688***		
	(0.0231)			(0.0497)		
$\text{status}_{rdt}^{national_occ(r)}$	-0.299***	-0.0132	-0.0579***			
	(0.0168)	(0.0173)	(0.0155)			
taboo=1 $\times$ $\text{status}_{rdt}^{national_occ(r)}$	0.0710***	0.0466***	0.0363***			
	(0.00752)	(0.00776)	(0.00714)			
$\text{status}_{rdt}^{national_w(o)}$				-0.235***	-0.00231	-0.0264
				(0.0138)	(0.0122)	(0.0192)
taboo=1 $\times$ $\text{status}_{rdt}^{national_w(o)}$				0.273***	0.0812***	0.0359**
				(0.0160)	(0.0148)	(0.0152)
Observations	1,111,072	1,110,724	1,109,544	1,089,132	1,088,876	1,088,280
Adjusted R^2	0.540	0.576	0.593	0.541	0.576	0.591
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No	No	Yes	No
religion*state*product*district*quarter	No	No	Yes	No	No	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. In columns 1-3, status is measured by local returns to the national occupational mix of each religion. In columns 4-6, status is measured by national returns to the initial local occupational mix of each religion. Columns 1 and 4 include the baseline fixed effects, columns 2 and 5 add the fixed effects for cross-sectional identification and columns 3 and 6 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * p $<$ 0.10, ** p $<$ 0.05, *** p $<$ 0.01.

fixed-effect specifications (with the cross-sectional fixed effects in column 2 most appropriate), and columns 4-6 use the $\text{Status}_{rdt}^{national_w(o)}$ measure (with the panel fixed effects in column 6 most appropriate). Consistent with our conceptual framework, for all six columns, α_3 is positive and significant: households are more likely to adhere to the norms and taboos of their religious group (relative to those of other groups) when its status is higher. Notice that increases in group relative status tend to *increase* consumption of goods that are taboo to other religions but not your own (second and fourth rows). This is again consistent with a shift away from a shared ethnic cuisine that partially adheres to the taboos of other religious groups. However, these negative coefficients are not significant in our preferred specifications (column 2 and 6).

In terms of magnitudes, we focus on the most conservative estimates that come from the more-standard Bartik shift-share with panel fixed effects in column 6. A twenty-percent increase in the real returns to occupations that one's religious group initially specializes in—the 90th percentile of status changes between 1987 and 2000—increases abstinence of that member's religious taboos by 0.72 percentage points relative to non taboo goods.

4.4 Costs and Identity

The final determinant of group identity is the most economic in nature (and the least studied in the psychology literature): the market price of identifying with a group. Economists have long argued that religiosity—and associated prescribed behaviors such as church attendance or fertility rates—responds to economic incentives (Smith 1776; Manski and Mayshar 2003; Gruber and Hungerman 2008). Extending this literature, in the next section we study how identities respond to the cost of the (entire) prescribed bundle, while in this section we focus on own- and cross-price effects on the consumption of taboo goods.

As noted in Section 2, identity choice has several implications in this context. First, identifying with a group that has a strong taboo against consuming a particular good tends to make demand for that good less sensitive to prices, as the consumer could be at a corner solution. Second, and more interestingly, endogenous identity formation tends to generate complementarities between goods that are taboo for a particular group (absent strong countervailing income or substitution effects). Suppose the price of pork rises. Since Muslims are now less likely to consume pork anyway, their religious identity becomes more attractive relative to their ethnic identity. If this price effect leads to a change in identity, they would then consume less alcohol. Thus, our framework offers a theoretical foundation for why some goods are complements or substitutes. As with the effects of conflict, this implication is not obvious. If identity were fixed, one might expect the violation of one taboo to increase the propensity to *respect* other taboos, e.g due to "conscience accounting" (Gneezy et al., 2014).

We run a similar specification as in the previous taboo regressions (equations 3 and 5) but now, to provide a direct test of our hypotheses, we restrict the own- and cross-price elasticities (previously included in the controls) to depend only on whether the own or other good is a taboo:

$$\begin{aligned}\text{Abstain}_{ihrdst} =& \alpha_1 \text{Taboo}_{ir} + \alpha_2 \ln p_{iht} + \alpha_3 \sum_{j \neq i} \ln p_{jht} + \alpha_4 \text{Taboo}_{ir} \times \ln p_{iht} + \alpha_5 \text{Taboo}_{ir} \sum_{j \neq i} \ln p_{jht} \\ &+ \alpha_6 \sum_{j \neq i} \text{Taboo}_{jr} \ln p_{jht} + \alpha_7 \text{Taboo}_{ir} \sum_{j \neq i} \text{Taboo}_{jr} \ln p_{jht} \\ &+ \gamma_{2i} \ln realfoodexp_{ht} + \delta_{idt} + Additional_FE + \epsilon_{iht} \end{aligned} \quad (6)$$

We test two sets of hypotheses. The first concerns the demand for taboo goods. If a good is taboo for a particular group, we would expect that for households identifying with that group, the own- and cross-price elasticities would be relatively small. As the dependent variable is abstinence, and we expect $\alpha_2 > 0$, this hypothesis corresponds to $\alpha_4 < 0$ and sign$(\alpha_3) \neq$ sign(α_5).

The second set of hypotheses directly relates to identity changes. The mere fact that two

goods are both taboo does not have direct implications for whether goods are complements or substitutes. However, if identities change in response to prices, this generates a force that makes taboos complements with each other (beyond any standard substitution effects holding identity fixed): a change in the price of one taboo good changes the relative attractiveness of different identities and hence adherence to other taboos. In terms of equation (6), this implies that $\alpha_7 > 0$, i.e. that cross-price elasticities are more negative when both goods are taboos for the same identity group.

Table 6 reports the results. Columns 1-3 include only the baseline set of fixed effects, and the subsequent columns report results using the cross-section and panel fixed effects. For each set, we start by estimating the own- and cross-price elasticities (where all three cross-price elasticities are restricted to be identical). As expected, higher prices significantly increase abstinence. Goods are weak substitutes on average. Next (in columns 2, 5, and 8), we allow the own- and cross-price elasticities to vary with whether the own good is a taboo to test the first set of hypotheses. We find strong support: $\alpha_4 < 0$ and sign$\alpha_3 \neq$ signα_5 in all specifications. Demand for taboo goods is highly inelastic. Similarly, the cross-price terms imply goods are in general substitutes, but are highly cross-price inelastic when the own good is a taboo.

Finally, columns 3, 6, and 9 further allow the cross-price elasticities to vary depending on whether the *other* good is a taboo or not. Our main hypothesis, that taboo goods will tend to be complements with each other in a framework where identity choices depend on economic costs, is strongly supported. Across all three specifications, α_7 is significantly greater than zero. When both the own and other good are taboos, the cross-price terms are more positive and goods are more complementary.

As in any demand analysis, the endogeneity of prices is of concern and good instruments are elusive. We note, however, that while a standard endogeneity story would tend to attenuate the estimated magnitude of the own-price demand response, it is not clear why it should *differentially* affect cross-price elasticities between taboo and non-taboo goods, where which goods are taboo varies across households in the same location. We also note that the reverse-causality worry we discussed in previous sections would likely bias us against finding our key result, $\alpha_7 > 0$. Suppose the proportion of households that identify with their religion increases due to some other shock (e.g. status or salience changes). This would lead households to reduce their consumption of their religious taboo goods and prices should fall as we move down the supply curve.[32] Thus we would observe price reductions for one taboo good associated with

[32] As long, of course, as this reduction dominates any increased consumption from members of other religious groups moving away from a shared ethnic identity. As noted above, we indeed find that conflict-induced abstention leads to a fall in price (Appendix Table B.3.7).

Table 6: Costs, Price Elasticities and Identity

	LHS Variable: Abstain from Consuming Good i								
	Baseline			Cross-section			Panel		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
taboo$_i$=1	0.160***	0.166***	0.176***						
	(0.00219)	(0.0116)	(0.0104)						
ln p$_i$	0.0153***	0.0634***	0.0476***	0.00958***	0.0500***	0.0499***	0.00748***	0.0208***	0.0208***
	(0.00275)	(0.00374)	(0.00353)	(0.00256)	(0.00484)	(0.00483)	(0.00252)	(0.00397)	(0.00396)
sum ln p$_j$	-0.000974	-0.0161***	0.00686***	-0.000925	-0.00635***	-0.00255	-0.000848	-0.00459**	-0.00148
	(0.00170)	(0.00203)	(0.00194)	(0.00159)	(0.00245)	(0.00291)	(0.00159)	(0.00185)	(0.00210)
taboo$_i$=1 x ln p$_i$		-0.0612***	-0.0443***		-0.0507***	-0.0505***		-0.0164***	-0.0166***
		(0.00287)	(0.00260)		(0.00456)	(0.00454)		(0.00331)	(0.00331)
taboo$_i$=1 x sum ln p$_j$		0.0199***	-0.00683***		0.00712***	0.000706		0.00482***	0.00221
		(0.00137)	(0.00124)		(0.00226)	(0.00359)		(0.00113)	(0.00168)
sum (ln p$_j$ x taboo$_j$)			-0.0456***			-0.00776			-0.00681**
			(0.00114)			(0.00475)			(0.00284)
taboo$_i$=1 x sum (ln p$_j$ x taboo$_j$)			0.0486***			0.0106*			0.00627**
			(0.00117)			(0.00547)			(0.00291)
Observations	1,115,640	1,115,640	1,115,640	1,115,292	1,115,292	1,115,292	1,114,116	1,114,116	1,114,116
Adjusted R^2	0.539	0.540	0.550	0.576	0.576	0.576	0.594	0.594	0.594
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	No	No	Yes	Yes	Yes	No	No	No
religion*state*product*district*quarter	No	No	No	No	No	No	Yes	Yes	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. Columns 1, 4 and 7 include own and cross-price elasticities. Columns 2, 5 and 8 add the interaction between taboo and own and cross-price elasticities. Columns 3, 6 and 9 allow cross-price elasticities to differ depending on whether both own and other goods are taboos. Columns 1-3 include the baseline fixed effects, columns 4-6 add the fixed effects for cross-sectional identification and columns 7-9 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * p < 0.10, ** p < 0.05, *** p < 0.01.

higher abstention from other taboo goods, not lower as we find.

Following the logic of Hausman (1996), we also instrument prices with the price in a nearby village as in Atkin (2013). For the instrument to be valid, we require that supply shocks are correlated spatially within districts (driving the correlation between nearby village prices) but idiosyncratic village tastes are not. Atkin (2013) provides evidence that supports these assumptions for food products in India. Appendix Table E.1 reports these IV results. The first stages are strong, the IV estimates are close to the OLS ones, and we cannot reject either set of hypotheses.

5 An Almost Ideal Demand System with Identity Choice

The results in the previous section suggest that the tendency of Indian households to follow the prescribed behavior of their religious or ethnic group depends on several features of their social and economic environment: salience of group membership, group status, and prices. However, our analysis has so far focused on a rather narrow set of goods, where religiously or ethnically prescribed behavior is clearly documented in scriptures and historical texts. But social identity may affect a household's consumption of many other goods, even if they are not strict taboos or the most well-known ethnic staples. In this section we therefore pursue an analysis of the entire food-consumption bundle. We do not impose any assumption on which goods are "identity goods" and which are not. Nor do we assume ex-ante what the 'appropriate' level of consumption of a particular good for a particular group is. Rather, we take the prototypical bundle to be the observed mean bundle in the group. This is consistent with a long line of research showing that individuals tend to mimic the observed prevalent behavior of other members of their group—what is known in psychology as the *descriptive norm* (Cialdini et al. 1990, see Goldstein et al. 2008; Allcott 2011 for evidence of causal effects of descriptive norms on behavior).

Furthermore, we now use a specification that combines all three factors that can shape identification choices—salience, status and costs—and explicitly takes into account the choice between religious and ethnic identity. This allows us to examine the marginal importance of each of the three factors, taking into account that they may be correlated (for example, Mitra and Ray (2014) show that changes in group status can affect inter-group violence). And, by formalizing the choice between a household's ethnic and religious identities, we generate additional testable implications of the model and can quantify changes in identity choices over our study period.

5.1 A Structural Model of Identity

Consider a household h belonging to two groups, $G_h = \{r, s\}$, where r denotes the household's religious group and s denotes the household's ethnic group. To simplify notation, we use generic

r and s with the understanding that these groups are h-specific. Thus, for a Muslim Gujarati household, $\overline{x}_r$ is the prototypical Muslim bundle, and $\overline{x}_s$ is the prototypical Gujarati bundle.

The indirect utility of household h that identifies with group $J \in G_h$ is:

$$V_{hJ}(p, m_h, y_J, \kappa_{hJ}; \overline{x}_J) = \delta_1 v(p, m_h; \overline{x}_J) + \delta_2 y_J + \delta_3 \kappa_{hJ} + \xi_{hJ}, \tag{7}$$

where $v(\cdot)$ is an AIDS indirect utility function (made explicit below). Consistent with Definition 1, identifying with group J means utility is also affected by the status of group J, y_J, and by the group-specific salience shifter κ_{hJ}. Finally, ξ_{hJ} is an idiosyncratic utility shifter.

We adapt the AIDS indirect utility, defined over the vector of prices p and income m_h, to take into account the prototypical consumption bundle of group J, given by the vector $\overline{x}_J$:

$$v(p, m_h; \overline{x}_J) = \frac{\ln m_h - \ln a_J(p)}{\Pi_i p_i^{\beta_i}}, \tag{8}$$

where $\ln a_J(p) = a_0 + \sum_i \overline{x}_{iJ} \ln p_i + \frac{1}{2} \sum_i \sum_k \gamma_{ik} \ln p_i \ln p_k$ and $\sum_i \overline{x}_{iJ} = 1$. Note that, other things equal, a household that identifies with group J is worse off when $\overline{x}_J$ is more expensive. Solving for the budget share x_{hiJ} of good i for household h that identifies with group J we obtain:

$$x_{hiJ} = \overline{x}_{iJ} + \sum_k \gamma_{ik} \ln p_k + \beta_i \ln(\frac{m_h}{a_J(p)}). \tag{9}$$

The *observed* budget share x_{hi} will depend on the household's chosen identity. From Assumption 1, the household chooses its religious identity if $V_{hr} > V_{hs}$ and its ethnic identity if $V_{hs} \geq V_{hr}$.[33] Therefore, defining $\widetilde{x_{iJ}} \equiv \overline{x}_{iJ} - \beta_i \sum_i \overline{x}_{iJ} \ln p_i$, the observed budget share is:

$$x_{hi} = \widetilde{x_{is}} + (\widetilde{x_{ir}} - \widetilde{x_{is}}) \mathbf{1}[V_{hr} > V_{hs}] + \sum_k \gamma_{ik} \ln p_k + \beta_i (\ln m_h - a_0 - \frac{1}{2} \sum_i \sum_k \gamma_{ik} \ln p_i \ln p_k). \tag{10}$$

From equations (7) and (8), the difference in utilities is:

$$V_{hr} - V_{hs} = -\delta_1 (\frac{\sum_i (\overline{x}_{ir} - \overline{x}_{is}) \ln p_i}{\Pi_i p_i^{\beta_i}}) + \delta_2 (y_r - y_s) + \delta_3 (\kappa_{hr} - \kappa_{hs}) + (\xi_{hr} - \xi_{hs}). \tag{11}$$

Religious identity is more appealing when its prototypical bundle is relatively cheap (i.e. $\sum_i \overline{x}_{ir} \ln p_i$ is relatively small), or when either its status or salience is relatively high.

5.2 Estimating the Model

Given the discrete choice between identities, the standard way to proceed is to assume that the ξ_{hJ} terms are i.i.d. extreme value (type I). With this assumption, the probability of choosing

[33] To break ties we assume the household identifies ethnically in case of indifference.

religious identity r is:

$$Prob(V_{hr} > V_{hs}) = F\left(-\delta_1 \frac{\sum_i (\overline{x}_{ir} - \overline{x}_{is}) \ln p_i}{\Pi_i p_i^{\beta_i}} + \delta_2(y_r - y_s) + \delta_3(\kappa_{hr} - \kappa_{hs})\right), \tag{12}$$

where $F(x) = \frac{1}{1+e^{-x}}$ is the standard logistic function. When averaging over many households of the same type and location, this probability becomes a proportion.[34] We obtain a non-linear demand system by substituting this probability for the function $\mathbf{1}[V_{hr} > V_{hs}]$ in equation (10).

While elegant, the highly non-linear nature of such a demand system and the large number of unknowns make it infeasible to estimate. Instead, we turn to a local linear approximation by making the assumption that $\xi_{hr} - \xi_{hs}$ is distributed uniformly with mean zero and a range equal to 1. Then the change in proportion of households of religion r simply equals the change in $-\delta_1 \frac{\sum_i (\overline{x}_{ir} - \overline{x}_{is}) \ln p_i}{\Pi_i p_i^{\beta_i}} + \delta_2(y_r - y_s) + \delta_3(\kappa_{hr} - \kappa_{hs})$. This motivates running the linear specification:

$$\begin{aligned} x_{hi} = (\overline{x}_{ir} - \overline{x}_{is})(-\eta_1 \sum_i (\overline{x}_{ir} - \overline{x}_{is}) \ln p_i + \eta_2(y_r - y_s) + \eta_3(\kappa_{hr} - \kappa_{hs})) \\ + \overline{x}_{is} + \sum_k \gamma_{ik} \ln p_k + \beta_i(\ln m_h - \sum_i \overline{x}_{is} \ln p_i) + FEs + \varepsilon_{hi}, \end{aligned} \tag{13}$$

where in addition to linearizing $Prob(V_{hr} > V_{hs})$, we follow Deaton and Muellbauer (1980) by replacing the non-linear price index $\ln a_s(p)$ with a Stone price index, here using state-level average budget shares and village-level prices.[35] As discussed earlier in this section, we proxy the $\overline{x}_{iJ}$ by the prototypical bundles we observe in these populations in the data (excluding own village).[36] Intuitively, reductions in the cost of my religious bundle relative to my ethnic bundle, or increases in the relative status or salience of my religious identity, push me closer to consuming the typical bundle of my religious group and further away from that of my ethnic group.

Estimating equation (13) is useful for two reasons. First, it can be interpreted as a linear approximation of equation (10) under some simplifying assumptions. Thus, the estimates will allow us to consider counterfactual scenarios as well as to quantify the relative importance of

[34] As noted in Section 2, we will not be able to empirically distinguish between a model where households make a binary choice between identities and one where households choose the relative weights they place on each.

[35] The one term absent from equation (13) is the budget share changes that result from different income effect magnitudes under the two identities, $-\beta_i \sum_i (\overline{x}_{ir} - \overline{x}_{is}) \ln p_i Prob(V_{hr} > V_{hs})$. We assume that these differences are negligible, or at least that $cov[\beta_i \sum_i (\overline{x}_{ir} - \overline{x}_{is}), (\overline{x}_{ir} - \overline{x}_{is})] \approx 0$ so that our coefficients of interest are unbiased. Similarly, as we do not include the denominator $\Pi_i p_i^{\beta_i}$ in the cost effect, η_1 should be interpreted as approximately equal to $\delta_1 E[1/\Pi_i p_i^{\beta_i}]$. Of course, if preferences are homothetic, neither of these issues arises. We apply the approximation $\ln a_s(p) \equiv a_0 + \sum_i \overline{x}_{is} \ln p_i + \frac{1}{2} \sum_i \sum_k \gamma_{ik} \ln p_i \ln p_k \approx \sum_i \overline{x}_{is} \ln p_i$, where following Moschini (1995) we use average budget shares in the Stone price index, not household-specific ones, here taking averages over members of group s which are simply the prototypical bundles $\overline{x}_{is}$ as we note below.

[36] Specifically, we assume that $\overline{x}_J$ is the mean vector of budget shares in group J. To compute this reference consumption for each household, we keep all observations except those in the village of the household. The prototypical consumption of the religious group is computed at the national level, while the prototypical bundle of the ethnicity is simply the average bundle in the state.

the three factors in shifting identity patterns. Second, equation (13) also provides a direct assessment of how the factors we explored in Section 4 alter the proximity between a household's consumption of good i and the prototypical consumption of that good in group J. For example, $\eta_2 > 0$ indicates that as the status of one's religious group increases relative to that of one's ethnic group, consumption moves closer to one's religious bundle than to one's ethnic bundle. But in contrast to Section 4, we now evaluate the full consumption basket, recovering prototypical bundles from the data rather than the scriptures, while jointly considering all three forces.

We run this specification using the same sample and combinations of fixed effects as in Section 4, but now we include all 124 food items. For the cost terms $\sum_i(\overline{x}_{ir}-\overline{x}_{is})\ln p_i$, we use the cost of the two identity bundles leaving out the cost of good i itself, $\sum_{j\neq i}(\frac{\overline{x}_{jr}}{\sum_{j\neq i}\overline{x}_{jr}}\ln p_j - \frac{\overline{x}_{js}}{\sum_{j\neq i}\overline{x}_{js}}\ln p_j)$, to minimize the worry that own-price effects drive these results. Once again, we proxy status shocks $y_r - y_s$ with the Bartik version of the occupational-returns-based instruments we introduced in Section 4.3, $\sum_o \log w_{od^-t}\theta_{odrt_0} - \sum_o \log w_{od^-t}\theta_{odst_0}$. Finally, for salience shocks we rely on the data concerning Hindu-Muslim violence introduced in Section 4.1.

Table 7 presents the results.[37] As before, all columns control for local supply and demand conditions via δ_{idt} fixed effects. Column 2 includes cross-sectional fixed effects δ_{irst} that deal with temporal shocks to adherence at the good-religion-ethnicity level. Column 3 includes panel fixed effects δ_{irsdq} that accommodate persistent deviations in good-religion-ethnicity-district-season consumption patterns. We refer the reader to discussions in Sections 4.1-4.4 regarding threats to identification for each of the three shocks. We cluster standard errors at the rdt level.

In both the cross-sectional fixed effect specification in column 2 of Table 7, and in our preferred panel specification in column 3, we find that relative cost, relative status and religious conflict shocks all have the expected signs and are significant at the 1 percent level. These results generalize our previous findings concerning taboo goods. They support the notion that identity is fungible, and that choices react systematically to these three forces. We discuss magnitudes in Section 6 where, among other things, we quantify the net changes in identity driven by each of the three forces over our sample period.

The Appendix contains several additional results. Appendix Table H.1 relaxes the symmetry implicit in equation (13)—that a shock moves the household as far away from one identity's

[37] Table 7 does not include the cross-price terms, $\sum_{k\neq i}\gamma_{ik}\ln p_k$, as they are infeasible to estimate for 124 foods in a sample with 34 millions observations. To alleviate the concern that cross-price terms may affect our results, we first impose that cross-price effects for good i are common within each of the thirteen food categories such as cereals or fruits in the NSS surveys (see Appendix Table A.1) and aggregate category prices using a Stone price index. We then draw a random sample of half the population within religion-district-time cells and estimate equation (13) on the restricted sample both with and without cross-price terms. Appendix Table H.3 shows that the results with cross-price terms are qualitatively and quantitatively similar, both to the half-sample results without cross-price terms and to our full-sample baseline in Table 7.

Table 7: Linear Approximation of Identity Choice with Cost, Status and Conflict

	LHS Variable: Share Spent on Good i		
	(1) Baseline	(2) Cross-section	(3) Panel
$(\overline{x_{ir}} - \overline{x_{is}}) \times (cost_r - cost_s)$	-0.0228	-0.660***	-0.692***
	(0.0545)	(0.0975)	(0.102)
$(\overline{x_{ir}} - \overline{x_{is}}) \times (status_r - status_s)$	0.481***	0.237***	0.222***
	(0.0274)	(0.0273)	(0.0639)
$(\overline{x_{ir}} - \overline{x_{is}}) \times conflict_r \ +/- \ 6\, months$	0.577***	0.0981***	0.273***
	(0.0474)	(0.0374)	(0.106)
Observations	32,523,464	32,515,776	32,435,920
Adjusted R^2	0.766	0.772	0.780
log price and total expenditure controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No
religion*state*product*district*quarter	No	No	Yes

Notes: Dependent variable is the share spent on good i in total food expenditure. $\overline{x}_{ir} - \overline{x}_{is}$ is the difference between prototypical religious and ethnic budget shares spent on good i. $cost_r - cost_s$ is the difference in religious and ethnic Stone price indexes leaving out the cost of good i. $status_r - status_s$ is the difference between religious and ethnic status measured by national returns to the initial local occupational mix of religion and ethnicity. $conflict_r + / - 6\, months$ is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * p < 0.10, ** p < 0.05, *** p < 0.01.

prescribed behavior as it does towards the other identity's prescribed behavior. More precisely, we evaluate our assumption that households make a choice between their religious and ethnic identity by separately interacting $\overline{x}_{ir}$ and $\overline{x}_{is}$ with each of the three shocks and examining whether the coefficients on the $\overline{x}_{ir}$ and $\overline{x}_{is}$ interactions are of similar magnitudes but opposite signs. We find reasonable support for the symmetry imposed by our model with the coefficients on all three $\overline{x}_{ir}$ interactions of opposite signs to the coefficients on the $\overline{x}_{is}$ interactions and magnitudes similar for costs and conflict. Appendix Table H.2 allows the impacts of each of the three shocks to vary by the religion of the household. The heterogeneity is somewhat limited, but upper-caste Hindus appear to care significantly more about status than other religious groups.

6 Counterfactuals and Implications

In this section, we explore the implications of our demand-system estimates for changes in identity, health and welfare over our sample period. This exercise serves two purposes. First, we quantify the absolute and relative importance of the three drivers of identity choice: economic cost, status, and salience. Second, the period 1987-2000 that our data span was a time of great change in India. The economic reforms that began in 1991 moved India towards a market economy, liberalized trade and dismantled the "license Raj". These reforms led to significant changes in prices as well as in occupational returns. The relationships between these once-

in-a-generation events and identity choices are of obvious interest and may have effects across multiple domains, including health and welfare.

6.1 Changes in Identity 1987-2000

We first explore the changes in identity induced by changes in cost, status and salience. Drawing on the local linear approximation in equation (13), the change in the proportion of households in a religion-ethnicity-district cell that identifies with their religion, dP_r, approximately equals

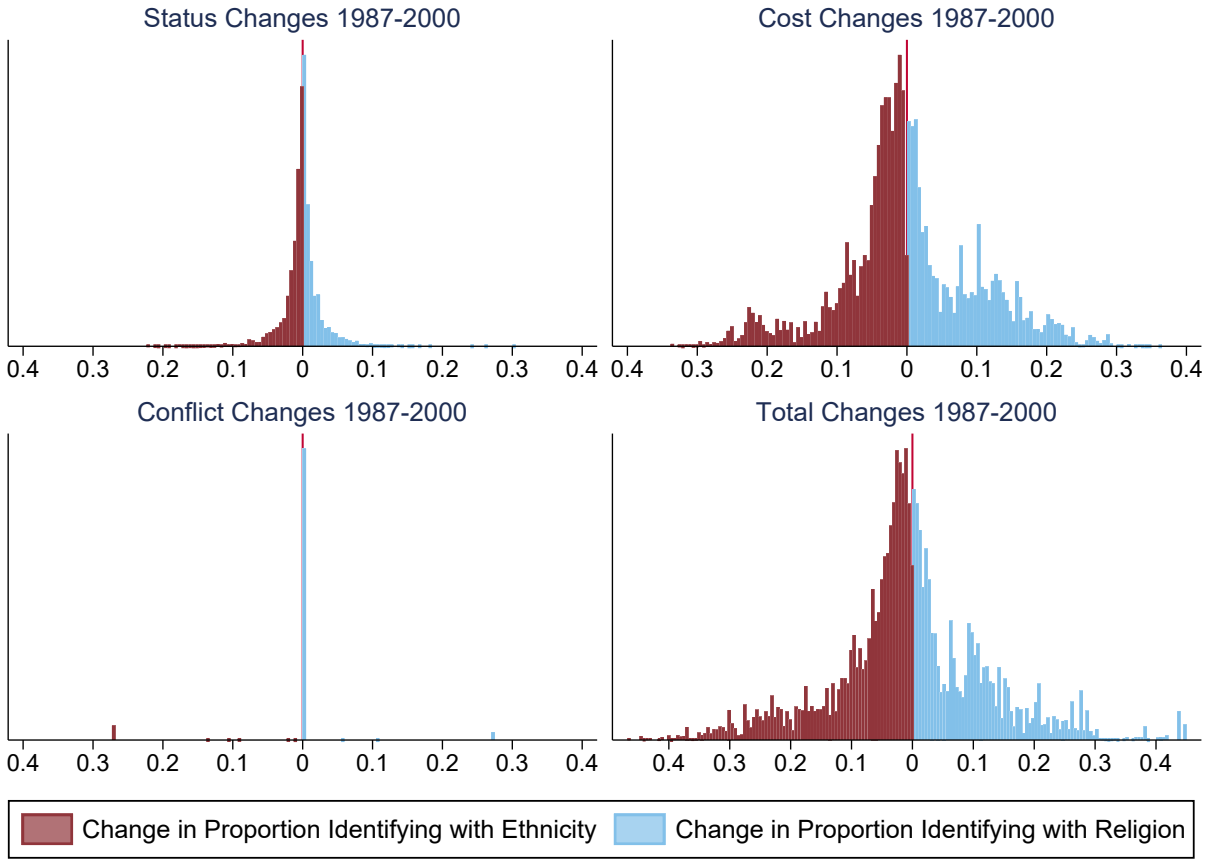
$$dP_r \approx -\eta_1 d \sum_i (\overline{x}_{ir} - \overline{x}_{is}) \ln p_i + \eta_2 d(y_r - y_s) + \eta_3 d(\kappa_{hr} - \kappa_{hs}).$$

Thus we can use the η estimates from Table 7 and combine them with the long changes in cost, status and salience (i.e. conflict) over the 1987-2000 period to calculate the (net) change in the proportion of households that identify with their religion. Figure 3 presents the distributions of these changes in proportions for each district-religion cell, both for the total change and for each of the three components. Appendix Figure I.1 breaks out these changes by religious group, aggregating over all districts using population weights.

Three features stand out. First, identity changes over this period were substantial. This is shown by the significant mass away from zero in the lower right panel of Figure 3. District-religion cells contributing to the mass to the left of zero are those where there was an increase in the proportion of households identifying with their ethnicity (with the proportion changing identity denoted on the x-axis). Cells contributing mass to the right saw increases in the proportion of households identifying with their religion. For the most affected district-religion cells (those in the 5 percent tail at either end of the distribution), almost one-quarter of households switched identity. Second, over this period there were more households shifting from their religious to ethnic identity than vice versa, but substantial heterogeneity across district-religion cells led to only a small net change of 0.65 percent of the population.

But perhaps the most unexpected result concerns the relative importance of the three forces. While we have shown that salience shocks due to conflict have a significant effect on identification decisions (consistent with previous literature and common narratives), quantitatively, prices and status have much larger impacts on identity choice. The muted effects of conflict are less surprising when realizing that our conflict shocks are both relatively uncommon and temporary, with the effects on identity fading out approximately 9 months after the shock. In contrast, changes in prices and occupational returns are ubiquitous and are much more persistent. Thus, when looking at changes in identity over a decade or more, conflict only plays a substantial role if there were shocks in that location at the start or end of the period.

Figure 3: Proportion of Population Changing Identity across District-Religion Cells, 1987-2000



The fact that the economic costs due to price changes have the largest impacts is particularly striking given that this channel has been largely overlooked in both public discussion and scholarly work. In fact, it can rationalize recent and much-discussed efforts by the Hindu-nationalist BJP party to raise the effective price of beef through bans and legislation.[38] Our results suggest that such a strategy may be more effective at hardening Hindu identities than is inciting religious violence, a tactic documented by Wilkinson (2004).

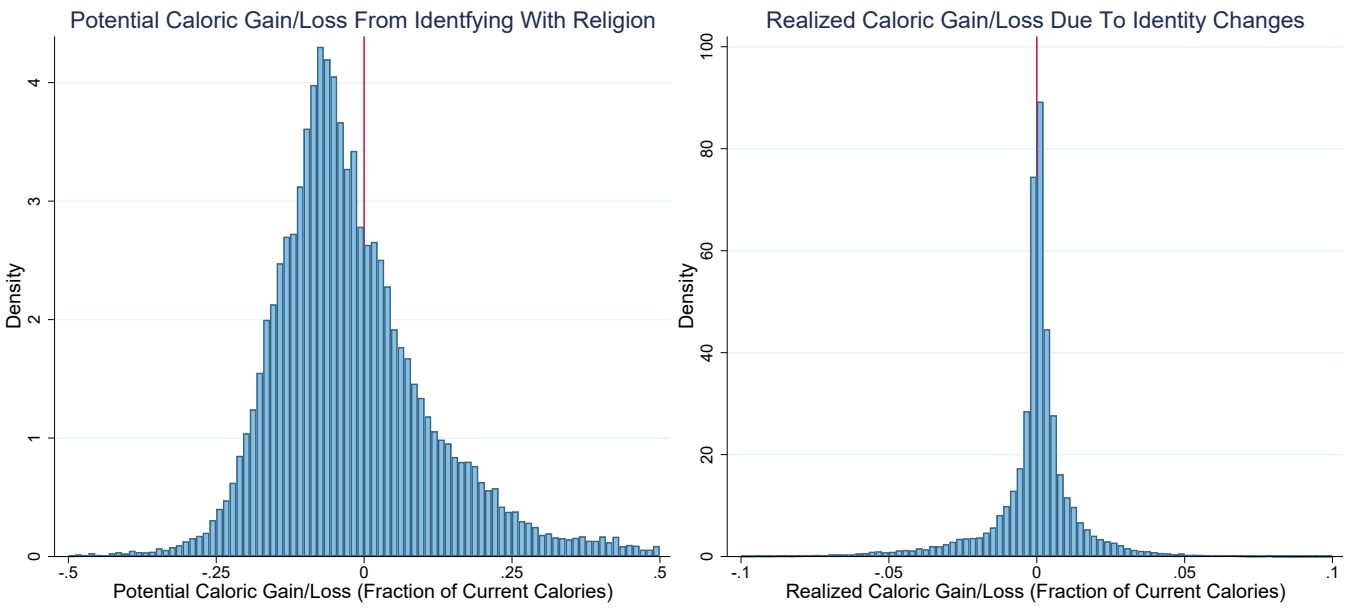
6.2 Effects of Identity Changes on Health 1987-2000

We next turn to documenting the impacts of identity changes on an easily measurable proxy for health: caloric intake. During this period, around half of Indian children were malnourished, and so changes in caloric intake due to changing identity choices have clear health implications.[39] The change in caloric intake per capita can be calculated for each religion-district cell

[38] See, for example, https://www.nytimes.com/2017/07/11/world/asia/india-cows-slaughter-beef-leather-hindu-supreme-court-ban.html for coverage of the government's attempted ban on cow slaughter.

[39] The 1992/93 NFHS finds that 52.0 percent of children aged 0-4 were stunted (more than 2 s.d. below the median WHO height-for-age) and 53.4 percent were underweight (more than 2 s.d. below the median weight-for-age).

Figure 4: Caloric Gains from Identity Changes, 1987-2000



by taking the change in identity choices and multiplying by the difference in the caloric intake of households under their two possible identities:

$$dCalories_{ih} \approx calories_per_kg_i \times \frac{foodexp_h}{p_{ih}}(\overline{x}_{ir} - \overline{x}_{is})dP_r.$$

Before reporting results, the left panel of Figure 4 plots the distribution of the difference in caloric intake assuming everyone starts from their ethnic identity and switches to their religious one. The potential for caloric gains from identity changes is substantial, with possible gains of 20 percent or more at the tails of the distribution. By and large, identifying with one's religion rather than with one's ethnicity tends to reduce caloric intake (the distribution is left-skewed). This is because in general the ethnic bundle is relatively less expensive at local prices than the religious one (see Atkin 2013 for a model of habit formation that generates this pattern due to developing a taste for comparative-advantage foods that are locally inexpensive).

The right panel of Figure 4 plots the distribution of realized caloric gains and losses due to identity changes over this period. As with the choices themselves, the distribution is reasonably symmetric. Thus, on average, caloric consumption declined by only 0.77 percent due to identity shifts. At the tails, however, caloric changes associated with shifting identity are large, with the 5th percentile of religion-district cells losing 3.2 percent on average and the 95th percentile gaining 2.2 percent. Note that if this average comes from a binary choice of identity, only a fraction of households are switching identity but those few that are actually switching are losing or

gaining many more calories—the magnitudes in the left panel. Either way, identity choices have real ramifications in terms of calories and hence health. To put these magnitudes in context, between 1987 and 2000 caloric intake per capita actually fell by 1.4 percent in our Indian sample. This decline in the face of rising real incomes has spawned much discussion (e.g. Deaton and Drèze (2009)). Our findings suggest identity changes may have contributed to this decline.

6.3 Effects of Identity Changes on Welfare 1987-2000

Finally, if we take our conceptual framework literally, we can say something about the additional welfare gains made possible by the fungibility of identity. In our framework, households are more flexible than consumer theory typically assumes, as they have an additional means to respond to shocks to their environment. By switching identity, households are essentially able to change the set of parameters that enter their (indirect) utility function $V_{hJ}(p, m_h, y_J, \kappa_{hJ}; \overline{x}_J)$, which can mitigate adverse price or status shocks. For example, if there is a shock that substantially raises the price of an important ethnic good, shifting into your religious identity ameliorates the resulting welfare loss.

We can calculate the change in compensating variation (CV) due to ability to shift identity. To derive an expression for this change, we implicitly define $z_{JJ'}$ as the proportional increase in income required in post-shock period 1 under identity J' to maintain the utility level of period 0 under identity J:

$$V_{hJ'}(p_1, m_1 z_{JJ'}, y_{J'1}, \kappa_{hJ'1}; \overline{x}_{J'}) = V_{hJ}(p_0, m_0, y_{J0}, \kappa_{hJ0}; \overline{x}_J).$$

Our interest is to compare $z_{JJ'}$ to the (standard) compensating variation z_{JJ} which obtains when identities are rigid. Substituting the indirect utility function from equation (7) and solving for $\ln z_{JJ'}$, it is straightforward to show that the log difference in the CV from maintaining the initial identity compared to switching identity is:

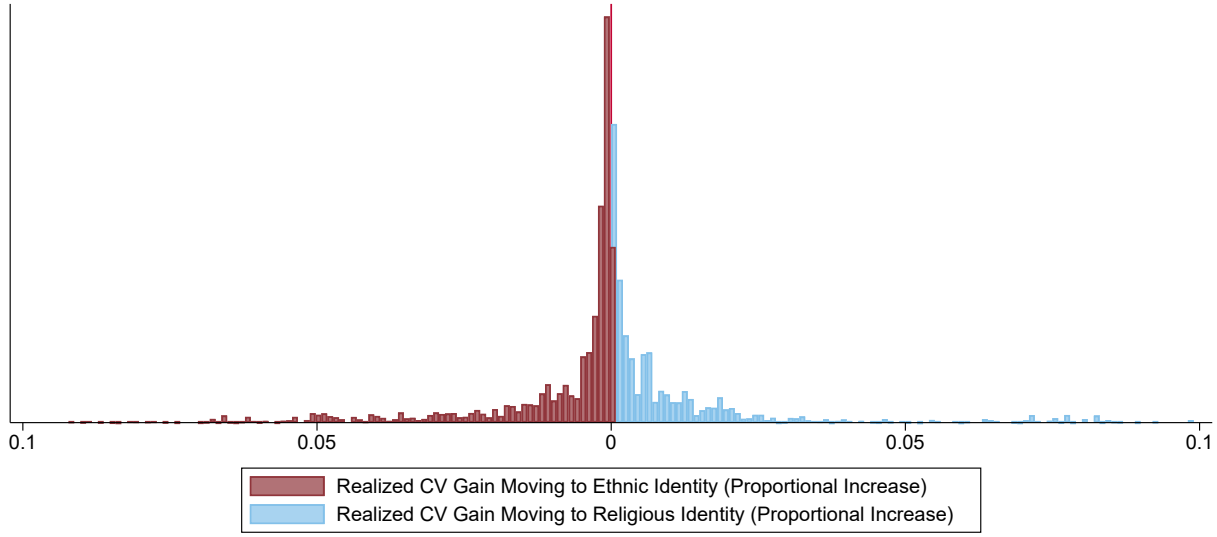
$$\ln z_{JJ} - \ln z_{JJ'} = \sum_i (\overline{x}_{iJ} - \overline{x}_{iJ'}) \ln p_{i1} - \frac{\Pi_i p_{i1}^{\beta_i}}{\delta_1}(\delta_2(y_{J1} - y_{J'1}) + \delta_3(\kappa_{J1} - \kappa_{J'1})). \tag{14}$$

To see the intuition, suppose status and salience are equal across identity groups. In this case, the potential CV gain from being able to switch identities is essentially the difference in the (post-shock) cost of the two prototypical bundles.

We evaluate equation (14) using the estimated ηs from the linear approximation above.[40] We then multiply the estimated $(\ln z_{ss} - \ln z_{sr})$ by positive values of dP_r, and $(\ln z_{rr} - \ln z_{rs})$ by negative values of dP_r to obtain the distribution of these additional welfare gains made possible

[40] Specifically, $\ln z_{JJ} - \ln z_{JJ'} \approx \sum_i (\overline{x}_{iJ} - \overline{x}_{iJ'}) \ln p_{i1} - \Pi_i p_{i1}^{\beta_i} E[\frac{1}{\Pi_i p_i^{\beta_i}}](\frac{\eta_2}{\eta_1}(y_{J1} - y_{J'1}) + \frac{\eta_3}{\eta_1}(\kappa_{J1} - \kappa_{J'1}))$.

Figure 5: Realized Compensating Variation Gains from Identity Changes, 1987-2000



due to the fungibility of identity. Figure 5 plots this distribution. For households in religion-district cells that are on average moving towards their ethnic identity, there is a long tail with 5 percent of cells enjoying a proportional increase in their compensating variation of more than 0.08 and an increase of 0.05 for cells moving towards their religious identity. As above, these cell-level averages are small as only a fraction of households switch. Appendix Figure I.3 shows numbers about six times larger if all households in the cell change identity. Such welfare gains are substantial. By changing the norms they follow in the face of adverse shocks, households can better cope with their new economic environment.

7 Conclusions

Recent political developments in both the developed and developing worlds have made more urgent the need to understand whether social identities are fungible, how they are chosen, and what the implications of those choices are. However, understanding the nature and implications of identity is difficult since identity choices are not directly observable. But consumption choices are observable and affected by norms and taboos of groups people identify with. This paper draws on this insight to explore how Indian households choose between their religious and ethnic identities.

We find that the consumption of prominent identity goods responds systematically to several forces, which we capture through a simple theoretical framework. Two of these forces feature prominently in prior social-identity research—group status and group salience, with the

latter proxied here by inter-group tensions. Consistent with economic theory, revealed identity choices also respond to a third, less studied, force: the cost of identifying with a group. To understand the relative magnitudes of these forces, we propose and estimate an Almost Ideal Demand System that incorporates endogenous identity choice. The estimates suggest that economic forces loom large, with changes in economic costs leading to the largest identity shifts over the period spanning India's 1991 economic reforms.

Such a finding may go some way to rationalize recent attempts by India's Hindu nationalist BJP party to raise beef prices through a nationwide ban on the sale of cattle for slaughter. In future work we plan to explore the electoral implications of these revealed identity choices directly. More broadly, our revealed preference approach draws on widely-available consumption survey data—including historical data—and can be fruitfully used in other contexts to provide a deeper understanding of identity and its ramifications.

References

Abdelal, R., Y.M. Herrera, A.I. Johnston, and R. McDermott, *Measuring Identity: A Guide for Social Scientists*, Cambridge, New York: Cambridge University Press, 2009.

Akerlof, G. and R.E. Kranton, "Economics and Identity," *Q.J.E.*, 2000, *115* (3), 715–753.

Alesina, A. and E. La Ferrara, "Ethnic Diversity and Economic Performance," *J. Econ. Literature*, 2005, *43* (3), 762–800.

Allcott, H., "Social Norms and Energy Conservation," *J. Pub. Econ.*, 2011, *95* (9-10), 1082–1095.

Appadurai, A., "How to Make a National Cuisine: Cookbooks in Contemporary India," *Comparative Studies in Society and History*, 1988, *30* (1), 3–24.

Atkin, D., "Trade, Tastes, and Nutrition in India," *A.E.R.*, 2013, *103* (5), 1629–63.

_ , "The Caloric Costs of Culture: Evidence from Indian Migrants," *A.E.R.*, 2016, *106* (4), 1144–81.

Basu, A., N. Sarkar-Roy, and P.P. Majumder, "Genomic Reconstruction of the History of Extant Populations of India Reveals Five Distinct Ancestral Components and a Complex Structure," *Proceedings of the National Academy of Sciences*, 2016, *113* (6), 1594–1599.

Bellet, Clément and Eve Colson-Sihra, "The Conspicuous Consumption of the Poor: Forgoing Calories for Aspirational Goods," *Working Paper, Hebrew Univ. of Jerusalem*, 2018.

Benjamin, D.J., J.J. Choi, and A.J. Strickland, "Social Identity and Preferences," *A.E.R.*, 2010, *100* (4), 1913–1928.

Bertrand, M. and E. Kamenica, "Coming apart? Cultural Distances in the United States over Time," *Working Paper no. 24771, NBER, Cambridge, MA*, 2018.

Bettencourt, B., K. Charlton, N. Dorr, and D.L. Hume, "Status Differences and In-group Bias: A Meta-analytic Examination of the Effects of Status Stability, Status Legitimacy, and Group Permeability," *Psychological Bulletin*, 2001, *127* (4), 520.

Burnkrant, R.E. and A. Cousineau, "Informational and Normative Social Influence in Buyer Behavior," *J. Consumer Res.*, 1975, *2* (3), 206–215.

Bursztyn, L., B. Ferman, S. Fiorin, M. Kanz, and G. Rao, "Status Goods: Experimental Evidence from Platinum Credit Cards," *Q.J.E.*, 2018, *133* (3), 1561–1595.

Chakravarti, A.K., "Regional Preference for Food: Some Aspects of Food Habit Patterns in India," *Canadian Geographer/Le Géographe Canadien*, 1974, *18* (4), 395–410.

Chandra, K., *Constructivist Theories of Ethnic Politics*, New York: Oxford University Press, 2012.

Charles, K.K., E. Hurst, and N. Roussanov, "Conspicuous Consumption and Race," *Q.J.E.*, 2009, *124* (2), 425–467.

Cialdini, R.B., R.R. Reno, and C.A. Kallgren, "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places," *J. Personality and Soc. Psych.*, 1990, *58* (6), 1015.

Cohn, A., E. Fehr, and M.A. Maréchal, "Business Culture and Dishonesty in the Banking Industry," *Nature*, 2014, *516* (7529), 86.

_ , M.A. Maréchal, and T. Noll, "Bad Boys: How Criminal Identity Salience Affects Rule Violation," *Rev. Econ. Studies*, 2015, *82* (4), 1289–1308.

Connelly, R., V. Gayle, and P.S. Lambert, "A Review of Occupation-based Social Classifications for Social Survey Research," *Methodological Innovations*, 2016, *9*, 2059799116638003.

Deaton, A. and J. Drèze, "Food and Nutrition in India: Facts and Interpretations," *Econ. and Polit. Weekly*, 2009, pp. 42–65.

_ and J. Muellbauer, "An Almost Ideal Demand System," *A.E.R.*, 1980, *70* (3), 312–326.

Depetris-Chauvin, E. and R. Durante, "One Team, One Nation: Football, Ethnic Identity, and Conflict in Africa," *Discussion Paper no. 12233, CEPR, Washington, DC*, 2017.

Eifert, B., E. Miguel, and D.N. Posner, "Political Competition and Ethnic Identification in Africa," *American J. Poli. Sci.*, 2010, *54* (2), 494–510.

Ellemers, N., P. Kortekaas, and J.W. Ouwerkerk, "Self-categorisation, Commitment to the Group and Group Self-esteem as Related but Distinct Aspects of Social Identity," *European J. Soc. Psych.*, 1999, *29* (23), 371–389.

Fearon, J.D., "Ethnic and Cultural Diversity by Country," *J. Econ. Growth*, 2003, *8* (2), 195–222.

Feld, J., N. Salamanca, and D.S. Hamermesh, "Endophilia or Exophobia: Beyond Discrimination," *Econ. J.*, 2016, *126* (594), 1503–1527.

Gneezy, U., A. Imas, and K. Madarász, "Conscience Accounting: Emotion Dynamics and Social Behavior," *Management Sci.*, 2014, *60* (11), 2645–2658.

Goldstein, N.J., R.B. Cialdini, and V. Griskevicius, "A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels," *J. Consumer Res.*, 2008, *35* (3), 472–482.

Grossman, G. and E. Helpman, "Identity Politics and Trade Policy," *Working Paper no. 25348, NBER, Cambridge, MA*, 2018.

Gruber, J. and D.M. Hungerman, "The Church versus the Mall: What Happens when Religion Faces Increased Secular Competition?," *Q.J.E.*, 2008, *123* (2), 831–862.

Hausman, J.A., "Valuation of New Goods under Perfect and Imperfect Competition," in Timothy F. Bresnahan and Robert J. Gordon, eds., *The Economics of New Goods*, Chicago: University of Chicago Press, 1996, pp. 207–248.

Heffetz, Ori, "A test of conspicuous consumption: Visibility and income elasticities," *Review of Economics and Statistics*, 2011, *93* (4), 1101–1117.

Hjort, J., "Ethnic Divisions and Production in Firms," *Q.J.E.*, 2014, *129* (4), 1899–1946.

Kalyvas, S.N., "Ethnic Defection in Civil War," *Compar. Polit. Stud.*, 2008, *41* (8), 1043–1068.

Kaufmann, C., "Possible and Impossible Solutions to Ethnic Civil Wars," *International Security*, 1996, *20* (4), 136–175.

Lavy, V., E. Sand, and M. Shayo, "Charity Begins at Home (and at School): Effects of Religion-Based Discrimination in Education," *Working Paper no. 24922, NBER, Cambridge, MA*, 2018.

Manski, C.F. and J. Mayshar, "Private Incentives and Social Interactions: Fertility Puzzles in Israel," *J. European Econ. Assoc.*, 2003, *1* (1), 181–211.

Mitra, A. and D. Ray, "Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India," *J. P.E.*, 2014, *122* (4), 719–765.

Moschini, Giancarlo, "Units of Measurement and the Stone Index in Demand System Estimation," *American J. Agricultural Econ.*, 1995, *77* (1), 63–68.

Nosofsky, R.M., "Similarity Scaling and Cognitive Process Models," *Annual Rev. Psych.*, 1992, *43* (1), 25–53.

Parkin, F., *Class Inequality and Political Order*, London: MacGibbon & Kee, 1971.

Reich, D., K. Thangaraj, N. Patterson, A.L. Price, and L. Singh, "Reconstructing Indian Population History," *Nature 461*, 2009, *461* (7263), 489.

Sambanis, N. and M. Shayo, "Social Identification and Ethnic Conflict," *American Polit. Sci. Rev.*, 2013, *107* (2), 294–325.

Shayo, M., "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution," *American Polit. Sci. Rev.*, 2009, *103* (02), 147–174.

_ and A. Zussman, "Judicial Ingroup Bias in the Shadow of Terrorism," *Q.J.E.*, 2011, *126* (3), 1447–1484.

_ and _ , "Conflict and the Persistence of Ethnic Bias," *American Econ. J.: Appl. Econ.*, 2017, *9* (4), 137–65.

Smith, A., *An Inquiry into the Nature and Causes of the Wealth of Nations*, London: Cadell and Strahan, 1776.

Srinivas, M.N., "A Note on Sanskritization and Westernization," *Far Eastern Quart.*, 1956, *15* (4), 481–496.

Subramanian, S. and A. Deaton, "The Demand for Food and Calories," *J.P.E.*, 1996, *104* (1), 133–162.

Tajfel, H. and J.C. Turner, "An Integrative Theory of Intergroup Conflict," in W.G. Austin and S. Worchel, eds., *The Social Psychology of Intergroup Relations*, Monterey, CA: Brooks-Cole, 1979, chapter 3.

Turner, J.C., M.A. Hogg, P.J. Oakes, S.D. Reicher, and M.S. Wetherell, *Rediscovering the Social Group: A Self-Categorization Theory*, New York: Basil Blackwell, 1987.

Weiss, Y. and C. Fershtman, "Social Status and Economic Performance: A Survey," *European Econ. Rev.*, 1998, *42* (3-5), 801–820.

Wilkinson, S.I., *Votes and Violence: Electoral Competition and Ethnic Riots in India*, Cambridge, UK: Cambridge University Press, 2004.

How Do We Choose Our Identity?
A Revealed Preference Approach Using Food Consumption

David Atkin, Eve Colson-Sihra and Moses Shayo

Online Appendices

Contents

A Data **2**

B Religious Conflict **6**
B.1 Conditional Event Study 6
B.2 Event Study: Other Tests 8
B.3 Religious Conflict and Taboo Adherence Regressions 9

C State Splits **16**

D Status Shocks **19**

E Cost of Identity **22**

F Regressions with Household Controls **23**

G Baseline Taboo Regressions with Ovo-Pesco Vegetarianism **27**

H Linear Approximation of Identity Choice **30**

I Counterfactuals **33**

A Data

Figure A.1: Fraction Population by Religious Groups in each District, all NSS Rounds

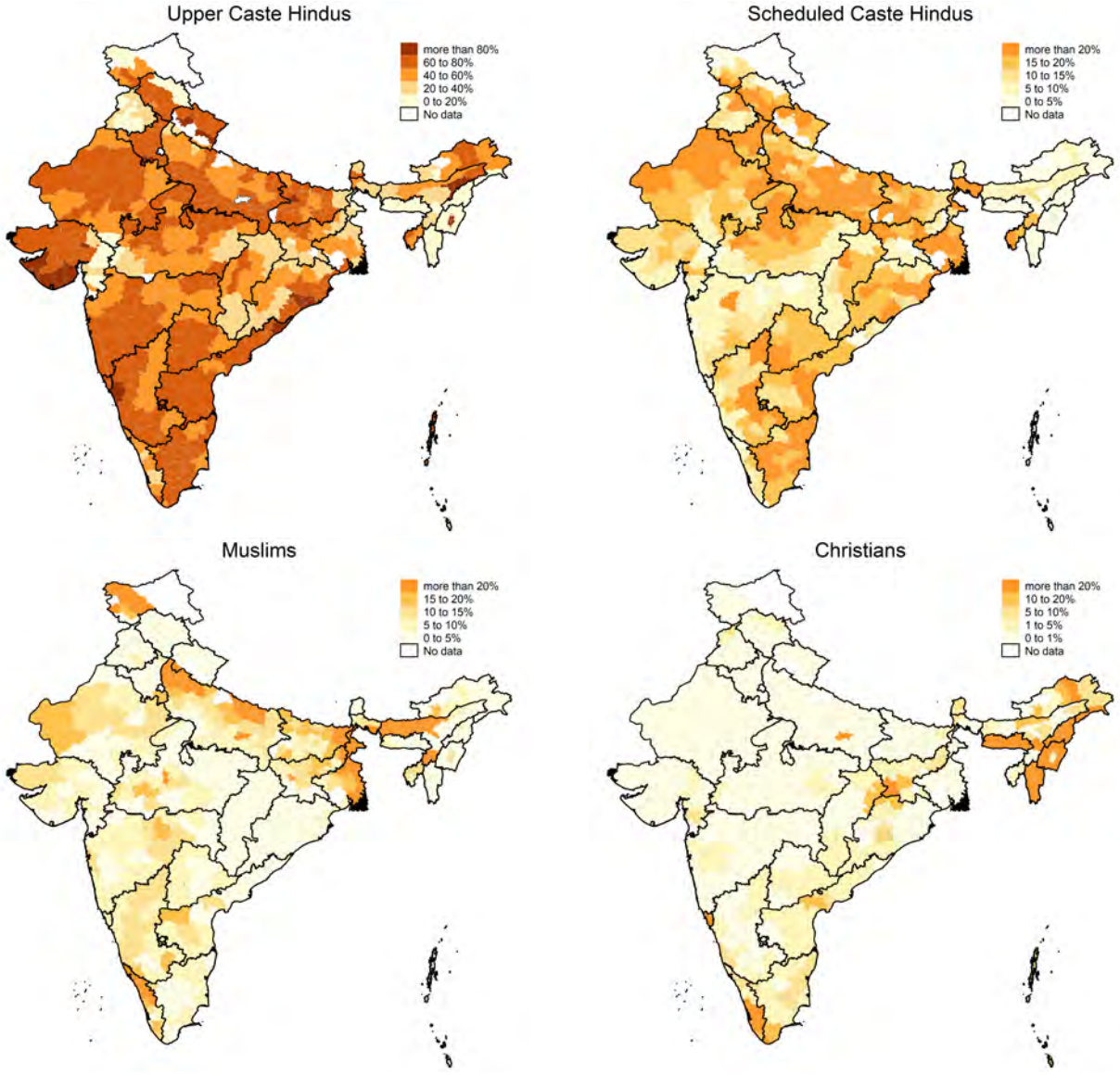


Figure A.2: Share of Rice and Wheat in Total Cereal Expenditures by District, all NSS Rounds

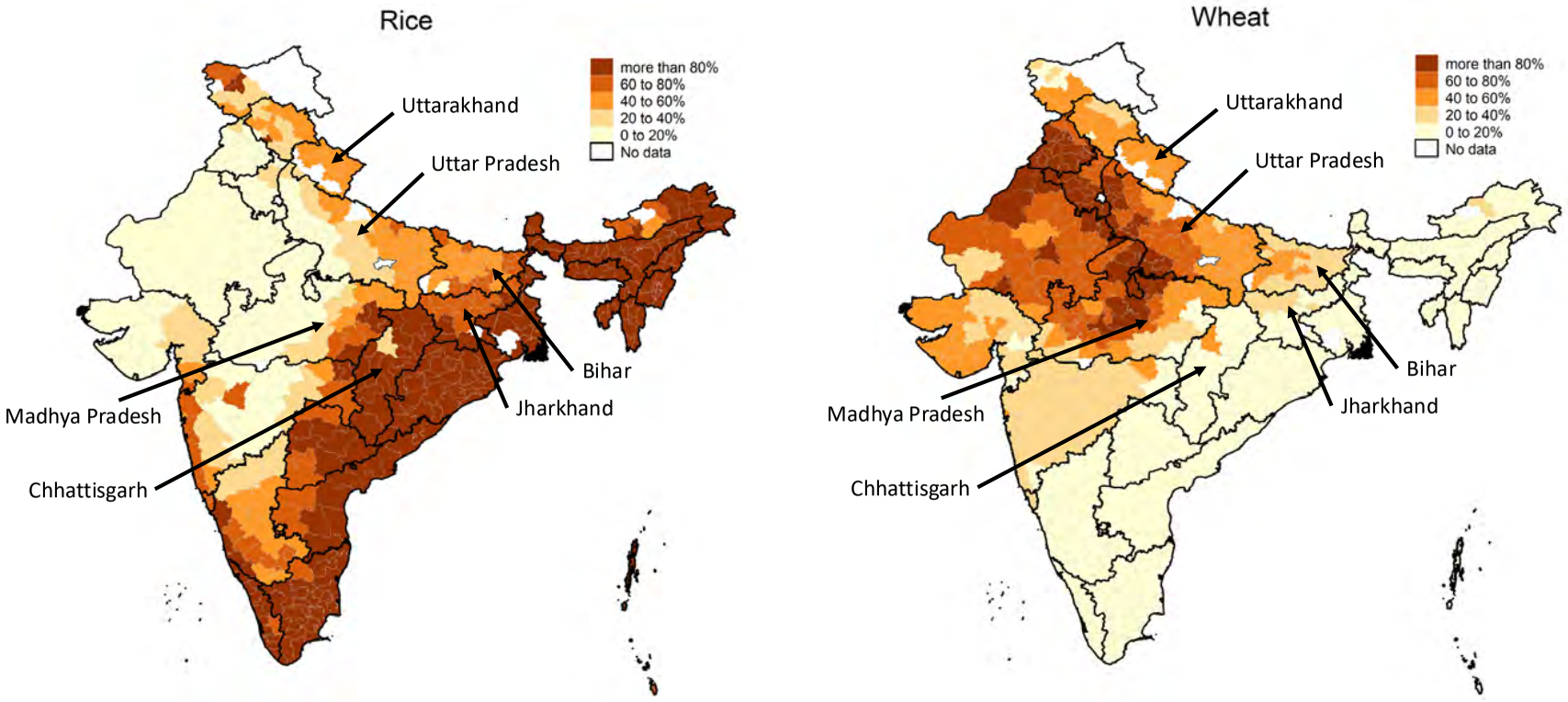


Table A.1: List of Food Items by NSS Categories

Category	**Items**
Cereals	bajra, barley, jowar, maize, millet, ragi, rice, wheat, other cereals
Pulses	gram, arhar, moong, masur, urd, khesari, peas, soya, other pulses
Dairy products	butter, curd, ghee, milk, baby food, condensed milk, ice cream, other milk products
Oils	vanaspati oil, mustard oil, groundnut oil, coconut oil, other oils
Meat	beef, chicken, eggs, fish, mutton, pork, other meats
Sugar	sugar, gur, misri, honey
Vegetables	onion, potato, radish, carrot, turnip, beet, sweet potato, arum, pumpkin, gourd, bitter gourd, cucumber, parwal, jhinga, snake gourd, cauliflower, cabbage, brinjal, bhindi, other leaf vegetables, french beans, tomato, green peas, chilli, capsicum, plantain, jackfruit, lemon, other vegetables
Fruits	banana, watermelon, pineapple, coconut, guava, singara, orange, mango, kharbooza, pear, berries, leechi, apple, grape, other fruits
Dry fruits	copra, groundnut, date, cashewnut, walnut, other nuts, kishmish, other dry fruits
Spices	garlic, turmeric, black pepper, dry chilli, tamarind, ginger, curry, other spices
Drinks	tea leaves, coffee beans, tea cup, coffee cup, cold drink, fruit juice, coconut juice, other drinks
Processed products	biscuits, salted refreshments, sweets, cooked meal, cake, pickle, sauce, jam, other processed food
Alcohol	beer, country liquor, foreign liquor, toddy
Intoxicant	pan

Table A.2: Hindu-Muslim Conflict by State and NSS Round

	1987-88		1993-94		1999-2000	
State	Incidence	No. Killed	Incidence	No. Killed	Incidence	No. Killed
Andhra Pradesh	0	0	1	0	1	0
Arunachal Pradesh	0	0	0	0	0	0
Assam	0	0	0	0	0	0
Bihar	2	17	0	0	2	5
Goa	0	0	0	0	0	0
Gujarat	24	49	8	54	8	11
Haryana	0	0	1	4	0	0
Himachal Pradesh	0	0	0	0	0	0
Jammu and Kashmir	7	7	0	0	3	9
Karnataka	3	1	9	49	1	0
Kerala	0	0	0	0	2	7
Madhya Pradesh	3	1	0	0	0	0
Maharashtra	14	37	5	564	11	2
Manipur	0	0	1	94	0	0
Meghalaya	0	0	0	0	0	0
Mizoram	0	0	0	0	0	0
Nagaland	0	0	0	0	0	0
Orissa	0	0	0	0	1	0
Punjab	0	0	0	0	0	0
Rajasthan	3	0	1	0	1	0
Sikkim	0	0	0	0	0	0
Tamil Nadu	1	1	1	1	0	0
Tripura	0	0	0	0	0	0
Uttar Pradesh	15	181	4	3	8	13
West Bengal	4	15	1	1	1	1

Notes: Table reports incidents of Hindu-Muslim conflict and numbers of people killed in each State for the period six months before, during and six months after each round based on the Varshney-Wilkinson Dataset.

B Religious Conflict

B.1 Conditional Event Study

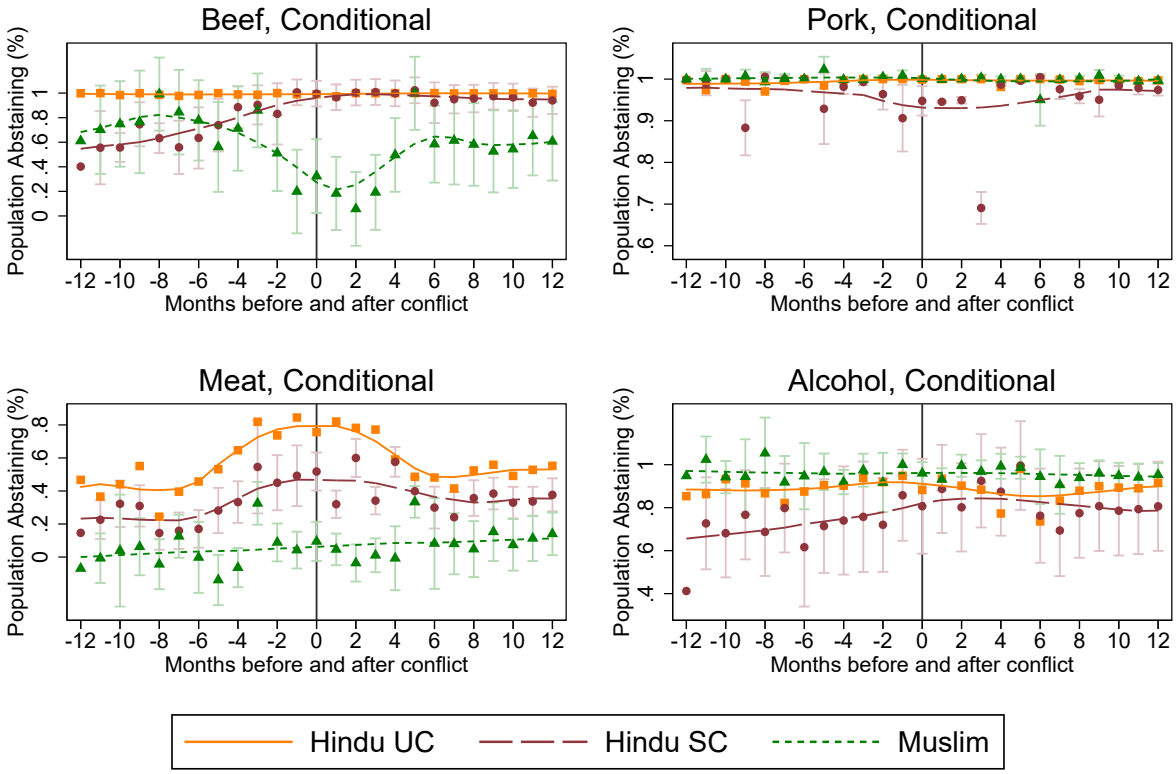
The non-parametric plots that show taboo abstention in the period building up to and after local conflict (Figure 2) do not account for potential confounds coming from price and income changes or other factors. For example, conflicts may be more likely in certain regions (those with different endowments or histories) or at certain moments of the year (religious festivals). We can potentially account for these factors by explicitly controlling for prices, total food expenditures and good-region-month fixed effects:

$$\text{Abstain}_{ihgm} = \sum_{m=-12}^{12} \theta_{im}^{SC} \text{SC}_h \times \text{Conflict}_{gm} + \sum_{m=-12}^{12} \theta_{im}^{M} \text{Muslim}_h \times \text{Conflict}_{gm} + \text{SC}_h + \text{Muslim}_h + \sum_j \gamma_{1ij} \ln price_{jh} + \gamma_{2i} \ln realfoodexp_h + \delta_{igm} + \epsilon_{ihgm}, \quad (15)$$

where Abstain_{ihgm} is an indicator variable that takes the value 1 if a householed does not consume good i; SC_h and Muslim_h are indicators that take the value 1 if a household h is scheduled-caste Hindu or Muslim (upper-caste Hindu is the reference group); Conflict_{gm} is an indicator for being surveyed m months before or after the first Hindu/Muslim conflict in region g; $\ln price_{jh}$ is the village median price of good j that controls for own- and cross-price effects; $\ln realfoodexp_h$ is the log of per capita food expenditure deflated by a Stone price index that controls for income effects; and δ_{igm} are good-region-month fixed effects that control for any local supply and demand conditions that are potentially correlated with conflict and are not adequately captured by prices. Standard errors are clustered at the gm level.

The θ_{im}^{r} coefficients capture consumption deviations relative to omitted group, upper-caste Hindus. Figure B.1.1 displays the predicted values from estimating Equation (15) for upper-caste Hindus, and adding the estimated θ_{im}^{r} coefficients for scheduled-caste Hindus and Muslims to this baseline consumption. The resulting patterns are very similar to those obtained using non-parametric regressions in Figure 2.

Figure B.1.1: Conflict and Taboo Avoidance, Conditional on Price, Income, Religion and Good-Region-Month FE, NSS 50th Round (1993-1994)



B.2 Event Study: Other Tests

Figure B.2.1: Conflict and Taboo Avoidance, 6 Months Before/After Conflict, NSS 50 (1993-1994)

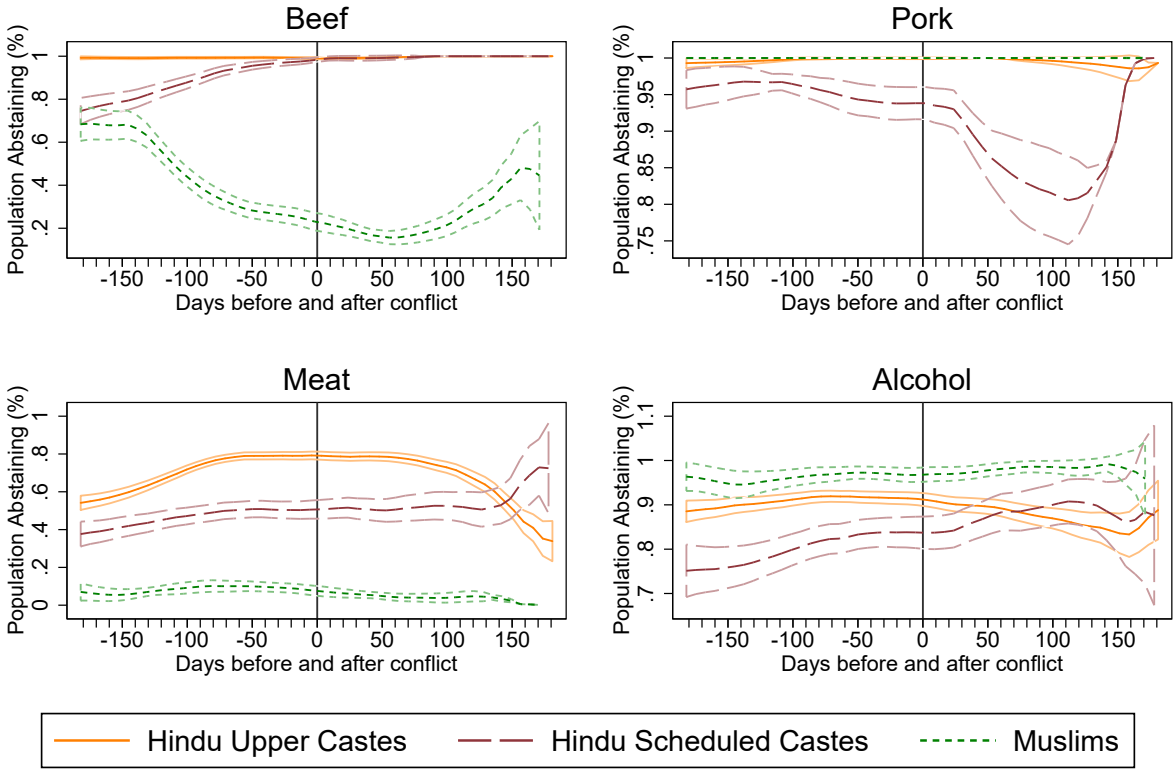


Figure B.2.2: Conflict and Beef/Pork Avoidance, Round 1993-1994, High vs. Low Local Religious Fractionalization

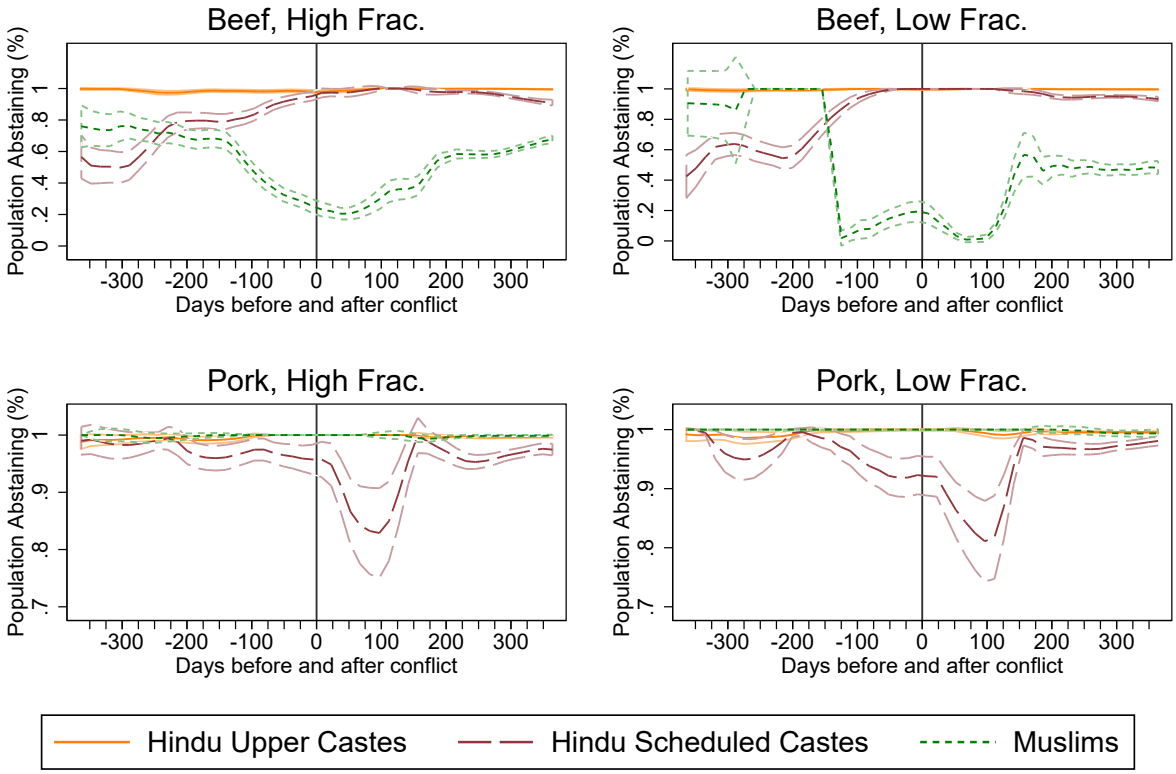
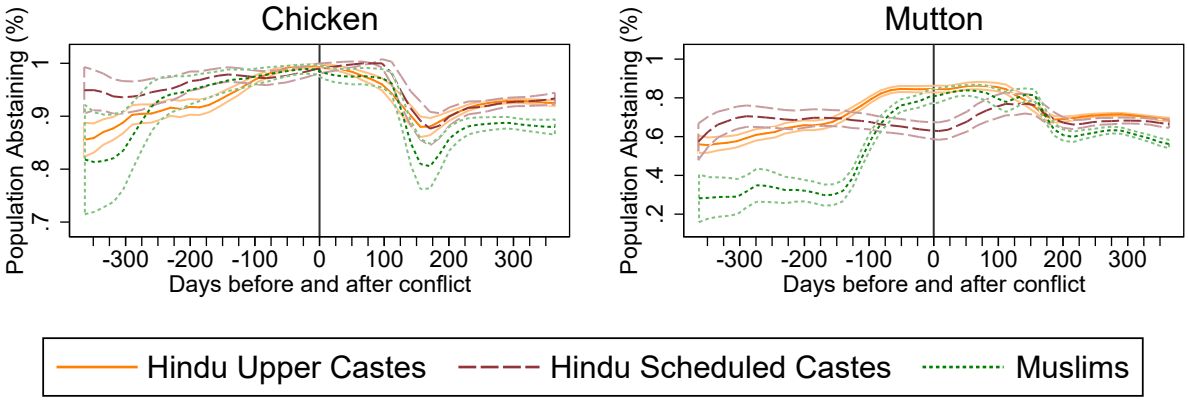


Figure B.2.3: Conflict and Chicken/Mutton Avoidance, NSS 50th Round (1993-1994)



B.3 Religious Conflict and Taboo Adherence Regressions

Table B.3.1: Religious Conflict and Taboo Adherence, Clustering at Higher Geographic Level

	LHS Variable: Abstain from Consuming Good i		
	Baseline	Cross-section	Panel
	(1)	(2)	(3)
	All	All	All
taboo=1	0.155***		
	(0.00430)		
conflict +/- 6 months	-0.102***	-0.0215	-0.0445*
	(0.0224)	(0.0240)	(0.0268)
taboo=1 $\times$ conflict +/- 6 months	0.0920***	0.0275***	0.0358***
	(0.0130)	(0.00736)	(0.00707)
Observations	1,115,640	1,115,292	1,114,116
Adjusted R^2	0.540	0.576	0.594
log prices and total expenditure controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No
religion*state*product*district*quarter	No	No	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. conflict +/- 6 months is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Robust standard errors clustered at religion-region-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3.2: Number of Religious Conflict Fatalities and Taboo Adherence

	LHS Variable: Abstain from Consuming Good i		
	Baseline	Cross-section	Panel
	(1)	(2)	(3)
	All	All	All
taboo=1	0.158***		
	(0.00219)		
log fatalities +/- 6 months	-0.0322**	0.0102	-0.00404
	(0.0147)	(0.0129)	(0.0140)
taboo=1 $\times$ log fatalities +/- 6 months	0.0385***	0.00767*	0.0100**
	(0.00677)	(0.00413)	(0.00397)
Observations	1,115,640	1,115,292	1,114,116
Adjusted R^2	0.539	0.576	0.594
log prices and total expenditure controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No
religion*state*product*district*quarter	No	No	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. Log fatalities is the log of the number of people killed in Hindu-Muslim conflicts in the district in the six months before or after the household is surveyed. It is computed using the inverse hyperbolic sine transformation to account for the zero observations. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3.3: Religious Conflict and Taboo Adherence, Lags and Leads

	LHS Variable: Abstain from Consuming Good i		
	Panel		
	(1)	(2)	(3)
taboo=1 × conflict t-0 quarters	0.0283**	0.0281**	0.0276**
	(0.0129)	(0.0130)	(0.0130)
taboo=1 × conflict t-1 quarters		0.0235**	0.0220**
		(0.0107)	(0.0109)
taboo=1 × conflict t-2 quarters		0.0334***	0.0359***
		(0.0105)	(0.0104)
taboo=1 × conflict t-3 quarters		-0.00287	-0.00224
		(0.00946)	(0.00934)
taboo=1 × conflict t-4 quarters		-0.00746	-0.00786
		(0.00980)	(0.00987)
taboo=1 × conflict t+1 quarters			0.0278**
			(0.0132)
taboo=1 × conflict t+2 quarters			0.00677
			(0.0136)
taboo=1 × conflict t+3 quarters			-0.00256
			(0.0128)
taboo=1 × conflict t+4 quarters			-0.0247
			(0.0216)
Observations	1,114,116	1,114,116	1,114,116
Adjusted R^2	0.594	0.594	0.594
log prices and total expenditure controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*state*product*round*quarter	No	No	No
religion*state*product*district*quarter	Yes	Yes	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. Conflict is an indicator for at least one occurrence of Hindu-Muslim conflict in the district. Column 1 shows the effect of conflict in the quarter in which the household is surveyed (t-0). Column 2 additionally includes lags of conflict from quarters t-1 to t-4. Column 3 further includes leads of conflict from quarters t+1 to t+4. All regressions include the main effects of taboo and conflict, including lags and leads of conflict in columns 2 and 3 (not shown). All regressions include the baseline fixed effects and the fixed effecs for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3.4: Religious Conflict and Beef Abstention in NSS 50 (1993-1994)

	LHS Variable: Abstain from Beef	
	Baseline	Cross-section
	(1)	(2)
taboo=1	0.323***	
	(0.0124)	
conflict +/- 6 months	-0.377*	-0.359***
	(0.221)	(0.0845)
taboo=1 $\times$ conflict +/- 6 months	0.385***	0.311***
	(0.0421)	(0.0453)
Observations	59,279	59,248
Adjusted R^2	0.379	0.480
log prices and total expenditure controls	Yes	Yes
district*quarter	Yes	Yes
religion*state*quarter	No	Yes

Notes: Dependent variable is an indicator for abstaining from beef. Taboo is an indicator equal to 1 if beef is a taboo for the religion of the household. conflict +/- 6 months is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed. Column 1 includes the baseline fixed effects and column 2 adds the fixed effects for cross-sectional identification. The regression is run using the NSS 50 round (1993-1994). Robust standard errors clustered at religion-district-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3.5: Religious Conflict in Nearby Districts and Taboo Adherence

	LHS Variable: Abstain from Consuming Good i		
	Baseline	Cross-section	Panel
	(1)	(2)	(3)
	All	All	All
taboo=1	0.138***	-41.76	1.204
	(0.00216)	(1155515.9)	(204460.1)
conflict +/- 6 months	-0.0900***	-0.0253	-0.0429
	(0.0228)	(0.0223)	(0.0271)
conflict, other districts in region	-0.0372**	-0.0199	0.0114
	(0.0147)	(0.0125)	(0.0108)
taboo=1 × conflict +/- 6 months	0.0843***	0.0340***	0.0359***
	(0.0128)	(0.00729)	(0.00714)
taboo=1 × conflict, other districts in region	0.0755***	0.0344***	0.00209
	(0.00638)	(0.00499)	(0.00392)
Observations	1,115,640	1,115,292	1,114,116
Adjusted R^2	0.541	0.576	0.594
log prices and total expenditure controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No
religion*state*product*district*quarter	No	No	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. conflict +/- 6 months is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed. conflict, other districts in region is an indicator for a conflict occurrence in other districts in the same region. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * p $<$ 0.10, ** p $<$ 0.05, *** p $<$ 0.01.

Table B.3.6: Butcher Shares by Religion, All Survey Rounds

	Butchers		Households	
	Count	Weighted Share	Count	Weighted Share
Hindus	703	0.514	284,905	0.827
Muslims	561	0.451	42,145	0.119
Christians	55	0.022	19,549	0.023
Sikhs	12	0.006	8,561	0.019
Jains	0	0.000	1,478	0.003
Budhists	4	0.005	3,175	0.006
Zoroastrians	1	0.000	126	0.000
Other Religions	6	0.004	3,593	0.004
Total	1,342	1	363,532	1

Table B.3.7: Demand-Side Effects of Conflict on Prices

	LHS Variable: log price by good-district-time							
	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) RF	(6) RF	(7) RF	(8) RF
fraction abstaining$_{idt}$	-2.081***	-0.317***	-1.000***	-0.836**	-2.780***	-0.763***	-1.329***	-1.244**
	(0.500)	(0.121)	(0.371)	(0.351)	(0.664)	(0.289)	(0.493)	(0.521)
conflict +/- 6 months				0.0627				0.0245
				(0.0471)				(0.0505)
Observations	12,369	13,187	12,369	12,369	12,369	13,187	12,369	12,369
Adjusted R^2	0.257	0.223	0.528	0.528	0.253	0.220	0.521	0.521
district*product*quarter	Yes	No	Yes	Yes	Yes	No	Yes	Yes
product*round*quarter	No	Yes	Yes	Yes	No	Yes	Yes	Yes
First-stage F-statistic (CDF)	621.8	1665.5	612.3	691.9				
First-stage F-statistic (RKF)	295.6	564.1	292.3	317.5				

Notes: Dependent variable is the log price at the good-district-quarter-round level. Fraction abstaining$_{idt}$ is the fraction of population abstaining, instrumented by the predicted rate of abstention $\widehat{ShareAbstain}_{idt}$ in the district, based on the estimated parameters from equation (3). Specifically for each household h we compute the predicted likelihood of abstaining $\widehat{Abstain}_{ihdt} = \widehat{\alpha}_1 \text{Taboo}_{ir} + \widehat{\alpha}_2 \text{Conflict}_{rdt} + \widehat{\alpha}_3 \text{Taboo}_{ir} \times \text{Conflict}_{rdt}$ using the estimated $\widehat{\alpha}$'s from the baseline regression, and then compute $\widehat{ShareAbstain}_{idt}$ as the weighted mean of $\widehat{Abstain}_{ihdt}$ by product-district-quarter-round. Columns 4 and 8 also include conflict as an independent variable (conflict +/- 6 months, a dummy for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed). Columns 1-4 are estimated using 2SLS, while columns 5-8 are the reduced-form results. Columns 1 and 5 add district-product-quarter fixed effects (panel identification), columns 2 and 6 add product-round-quarter fixed effects (cross-sectional identification), and columns 3-4 and 7-8 add both sets of fixed effects. Robust standard errors clustered at district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C State Splits

Figure C.1: Cross-District Migration and State Splits

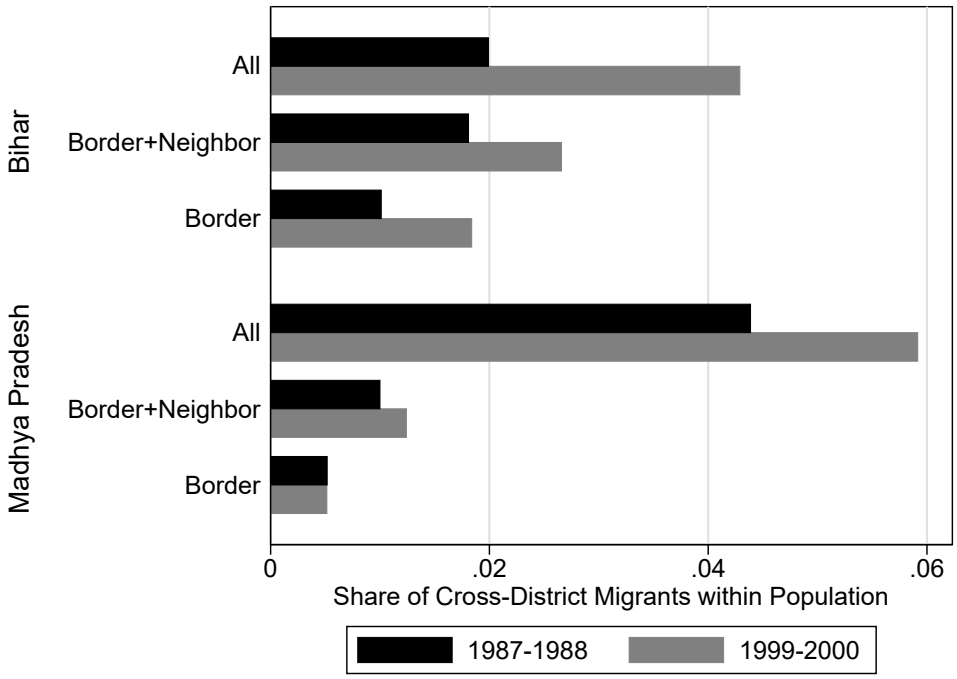


Table C.1: Ethnic Goods and State Splits

	LHS Variable: Share Spent on Cereal i				
	(1) All Regions	(2) Border Regions	(3) All Districts	(4) Border+Neighbor Districts	(5) Border Districts
Wheat-loving × Ethnic Cereal × 1993-1994	0.0373**	0.0379***			
	(0.0183)	(0.0138)			
Rice-loving × Ethnic Cereal × 1993-1994	0.0227**	0.0369***			
	(0.0103)	(0.0138)			
Wheat-loving × Ethnic Cereal × 1999-2000	0.0953***	0.0724***	0.0845***	0.0590***	0.0787***
	(0.0174)	(0.0134)	(0.0100)	(0.0129)	(0.0186)
Rice-loving × Ethnic Cereal × 1999-2000	0.0428***	0.0797***	0.0260***	0.0656***	0.107***
	(0.00966)	(0.0116)	(0.00866)	(0.0139)	(0.0185)
Observations	128,023	70,379	93,114	39,710	23,730
Adjusted R^2	0.732	0.772	0.793	0.830	0.836
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes
oldstate*round*quarter*product	Yes	Yes	Yes	Yes	Yes
district*quarter*product	No	No	Yes	Yes	Yes
region*quarter*product	Yes	Yes	No	No	No

Notes: Dependent variable is the share of cereal i (rice, wheat or other cereals) in total cereal expenditure. Ethnic Cereal is an indicator variable that takes the value 1 if cereal i is the ethnic cereal in future State. 1987-1988, 1993-1994 and 1999-2000 are round dummies with the initial round 1987-1988 as the reference group. In this table we break out the round effects separately for wheat- and rice-loving ethnicities (northwest and southeast of the fault line, respectively). Columns 1-2 are region-level regressions: column 1 includes all regions and column 2 restricts to border regions. Columns 3-5 are district-level regressions: column 3 includes all districts, column 4 restricts to border and border-adjacent districts, and column 5 to border districts. All regressions include the baseline fixed effects controlling for local supply and demand conditions (original state-time-product) and the fixed effects for panel identification (region-quarter-product for columns 1-2, district-quarter-product for columns 3-5). Robust standard errors clustered at region-round-quarter (columns 1-2) or district-round-quarter (columns 3-5) in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.2: Demand-Side Effects of Anticipated State Split

	LHS Variable: Log Price by Cereal-District-Time				
	(1) All Regions	(2) Border Regions	(3) All Districts	(4) Border+Neighbor Districts	(5) Border Districts
Ethnic Cereal × 1987-1988	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
Ethnic Cereal × 1993-1994	-0.00271	-0.00230			
	(0.00663)	(0.00726)			
Ethnic Cereal × 1999-2000	0.0167**	0.00862	0.0183*	0.0236	0.0291
	(0.00654)	(0.00592)	(0.0106)	(0.0151)	(0.0214)
Observations	2,840	1,452	1,880	696	456
Adjusted R^2	0.676	0.702	0.765	0.774	0.768
oldstate*round*quarter*product	Yes	Yes	Yes	Yes	Yes
district*quarter*product	No	No	Yes	Yes	Yes
region*quarter*product	Yes	Yes	No	No	No

Notes: Dependent variable is the log price of cereal i at the district-quarter-round level. Ethnic Cereal is an indicator variable that takes the value 1 if cereal i is the ethnic cereal in future state. 1987-1988, 1993-1994 and 1999-2000 are round dummies with the initial round 1987-1988 as reference group. Columns 1-2 are region-level regressions: column 1 includes all regions and column 2 restricts to border regions. Columns 3-5 are district-level regressions: column 3 includes all districts, column 4 restricts to border and border-adjacent districts, and column 5 to border districts. All regressions include the baseline fixed effects controlling for local supply and demand conditions (original state-time-product) and the fixed effects for panel identification (region-quarter-product for columns 1-2, district-quarter-product for columns 3-5). Robust standard errors clustered at region-round-quarter (columns 1-2) or district-round-quarter (columns 3-5) in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

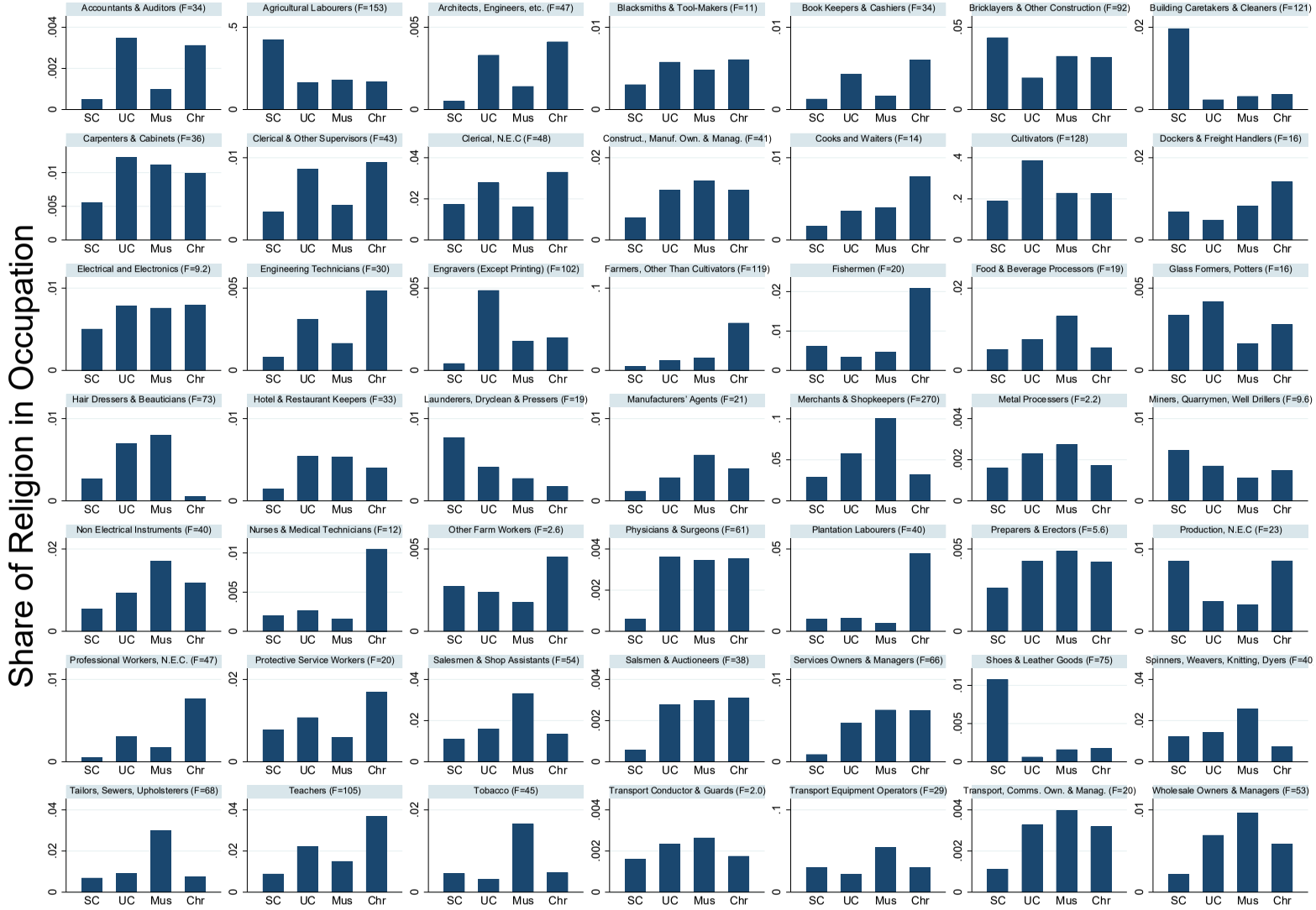
Table C.3: State Splits and Taboo Abstention

	LHS Variable: Abstain from Consuming Good i				
	(1) All Regions	(2) Border Regions	(3) All Districts	(4) Border+Neighbor Districts Districts	(5) Border Districts
Taboo Good x 1987–1988	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
Taboo Good x 1993–1994	-0.0107	-0.00137			
	(0.0121)	(0.0135)			
Taboo Good x 1999–2000	-0.0130	-0.0119	-0.0147	-0.0164	-0.0479**
	(0.0105)	(0.0119)	(0.00948)	(0.0162)	(0.0226)
Observations	171,780	94,600	124,708	53,280	31,796
Adjusted R^2	0.405	0.370	0.472	0.437	0.438
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes
oldstate*round*quarter*product	Yes	Yes	Yes	Yes	Yes
district*quarter*product*religion	No	No	Yes	Yes	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. 1987-1988, 1993-1994 and 1999-2000 are round dummies with the initial round 1987-1988 as reference group. Columns 1-2 are region-level regressions: column 1 includes all regions and column 2 restricts to border regions. Columns 3-5 are district-level regressions: column 3 includes all districts, column 4 restricts to border and border-adjacent districts, and column 5 to border districts. All regressions include the baseline fixed effects controlling for local supply and demand conditions (original state-time-product) and the fixed effects for panel identification (region-quarter-product-religion for columns 1-2, district-quarter-product-religion for columns 3-5). Robust standard errors clustered at region-round-quarter-religion (columns 1-2) or district-round-quarter-religion (columns 3-5) in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Status Shocks

Figure D.1: Heterogeneity in Occupational Shares by Religion, All Rounds (49 Most Common Occupations)



Joint F-test of equality across religions for every occupation: F(294,277877)=53.43, Prob > F =0.0000

Figure D.2: Heterogeneity in the Growth of Returns by Occupation, 1987-2000 (49 Most Common Occupations)

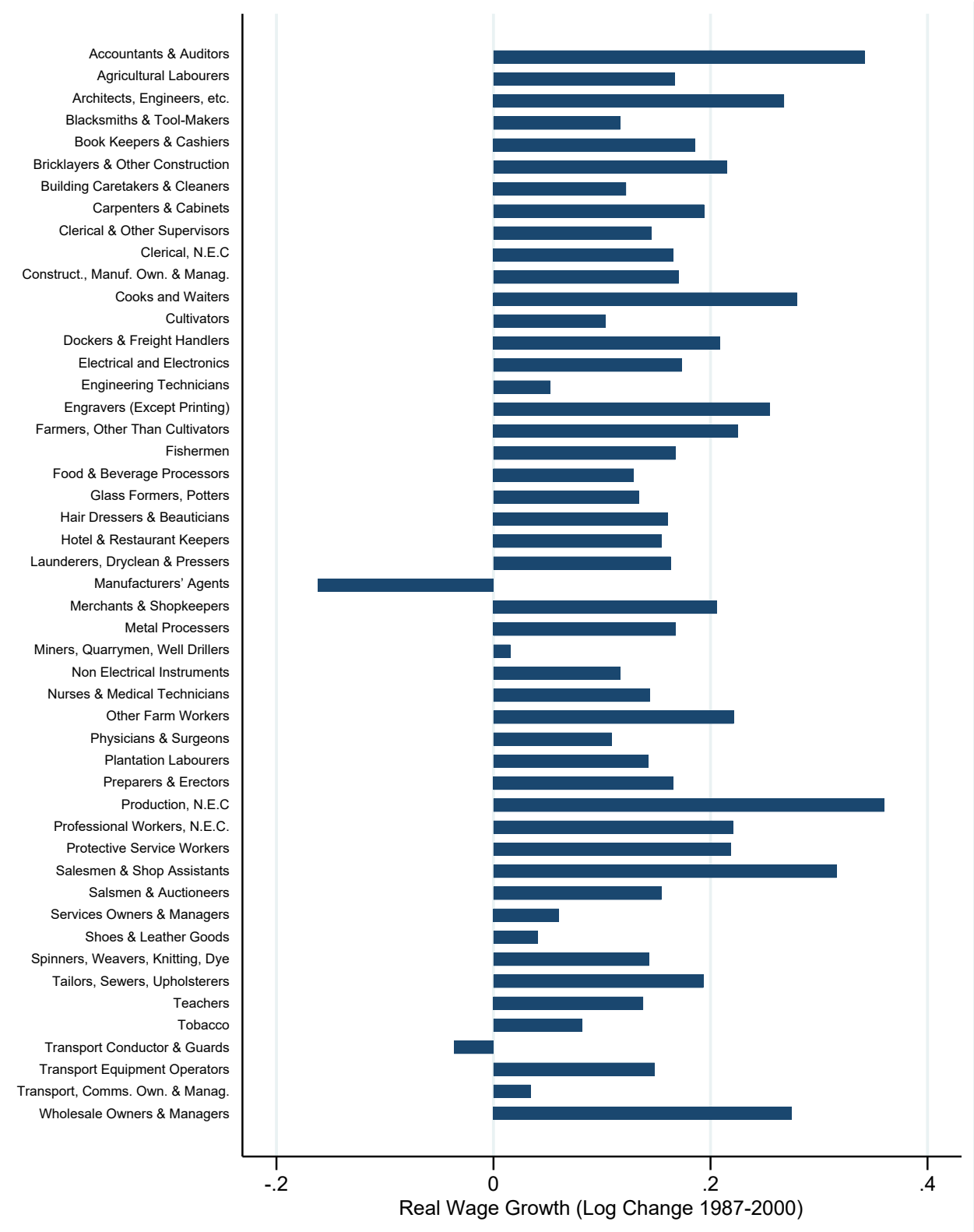


Table E.1: Costs, Price Elasticities and Identity with Instrumented Prices

	LHS Variable: Abstain from Consuming Good i								
	Baseline			Cross-section			Panel		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$taboo_i$=1	0.161***	0.167***	0.175***						
	(0.00223)	(0.0119)	(0.0107)						
ln p_i	-0.0261	0.0309	0.00660	-0.0286	0.0274	0.0268	-0.0152	-0.00821	-0.00847
	(0.0404)	(0.0410)	(0.0401)	(0.0379)	(0.0422)	(0.0421)	(0.0396)	(0.0430)	(0.0431)
sum ln p_j	0.000652	-0.0150***	0.00853***	0.000561	-0.00547*	-0.00171	0.0000184	-0.00192	0.00124
	(0.00234)	(0.00266)	(0.00262)	(0.00220)	(0.00298)	(0.00335)	(0.00221)	(0.00300)	(0.00324)
$taboo_i$=1 x ln p_i		-0.0620***	-0.0439***		-0.0576***	-0.0575***		-0.00728	-0.00741
		(0.00337)	(0.00312)		(0.00843)	(0.00840)		(0.00613)	(0.00614)
$taboo_i$=1 x sum ln p_j		0.0202***	-0.00692***		0.00741***	0.00108		0.00244	-0.000391
		(0.00140)	(0.00128)		(0.00228)	(0.00361)		(0.00176)	(0.00227)
sum (ln p_j x $taboo_j$)			-0.0457***			-0.00765			-0.00691**
			(0.00115)			(0.00474)			(0.00285)
$taboo_i$=1 x sum (ln p_j x $taboo_j$)			0.0488***			0.0105*			0.00657**
			(0.00118)			(0.00547)			(0.00293)
Observations	1,115,640	1,115,640	1,115,640	1,115,292	1,115,292	1,115,292	1,114,116	1,114,116	1,114,116
Adjusted R^2	0.539	0.540	0.549	0.576	0.576	0.576	0.593	0.593	0.593
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	No	No	Yes	Yes	Yes	No	No	No
religion*state*product*district*quarter	No	No	No	No	No	No	Yes	Yes	Yes
First-stage F-statistic (CDF)	5928.1	2954.7	2951.3	5707.9	2788.2	2788.7	4993.9	2454.2	2451.7
First-stage F-statistic (RKF)	40.17	20.10	20.09	39.10	19.61	19.62	33.57	16.82	16.81

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. In all regressions, the price of good i is instrumented by its price in a nearby village. Columns 1, 4 and 7 include own and cross-price elasticities. Columns 2, 5 and 8 add the interaction between taboo and own and cross-price elasticities. Columns 3, 6 and 9 allow cross-price elasticities to differ depending on whether both goods are taboos. Columns 1-3 include the baseline fixed effects, columns 4-6 add the fixed effects for cross-sectional identification and columns 7-9 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

F Regressions with Household Controls

Table F.1: Religious Conflict and Taboo Adherence, with Household Controls

	LHS Variable: Abstain from Consuming Good i					
	Baseline	Cross-section	Panel	Panel		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	Urban	Rural
taboo=1	0.153***					
	(0.00218)					
conflict +/- 6 months	-0.0942***	-0.0217	-0.0403			
	(0.0207)	(0.0214)	(0.0255)			
taboo=1 × conflict +/- 6 months	0.0899***	0.0276***	0.0357***			
	(0.0125)	(0.00714)	(0.00706)			
conflict past (6 months)				-0.0248	0.0244	-0.133***
				(0.0207)	(0.0259)	(0.0301)
conflict present/future (6 months)				-0.0339	-0.0202	-0.0795
				(0.0366)	(0.0255)	(0.0673)
taboo=1 × conflict past (6 months)				0.0387***	0.0338**	0.0331***
				(0.00843)	(0.0167)	(0.0102)
taboo=1 × conflict present/future (6 months)				0.0235**	0.0453**	0.00817
				(0.00980)	(0.0181)	(0.0126)
Observations	1,112,876	1,112,536	1,111,356	1,111,356	344,880	764,264
Adjusted R^2	0.544	0.580	0.597	0.597	0.618	0.605
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No	No	No	No
religion*state*product*district*quarter	No	No	Yes	Yes	Yes	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. Conflict is an indicator for at least one occurrence of Hindu-Muslim conflict in the district. Columns 1-3 consider a conflict occurrence in the six months before or after the household is surveyed. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Columns 4-6 differentiate the effect of a conflict occurrence in the six months preceding the quarter of the survey, and the six months covering the quarter of the survey and the subsequent quarter. Column 5 restricts the analysis to the urban population, and column 6 to the rural population. All regressions include the household controls used in Subramanian and Deaton (1996): log of household size, household demographic shares by age and gender, and indicators for being self-employed and working in the agricultural sector. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F.2: Status and Choice of Identity with Household Controls

	LHS Variable: Abstain from Consuming Good i					
	Baseline	Cross-section	Panel	Baseline	Cross-section	Panel
	(1)	(2)	(3)	(4)	(5)	(6)
taboo=1	-0.0308			-0.668***		
	(0.0227)			(0.0485)		
$\text{status}_{rdt}^{national_o cc(r)}$	-0.290***	-0.0104	-0.0516***			
	(0.0165)	(0.0172)	(0.0155)			
taboo=1 $\times$ $\text{status}_{rdt}^{national_o cc(r)}$	0.0677***	0.0469***	0.0333***			
	(0.00738)	(0.00769)	(0.00711)			
$\text{status}_{rdt}^{national_w(o)}$				-0.225***	-0.00397	-0.0221
				(0.0136)	(0.0121)	(0.0191)
taboo=1 $\times$ $\text{status}_{rdt}^{national_w(o)}$				0.265***	0.0839***	0.0261*
				(0.0156)	(0.0147)	(0.0151)
Observations	1,108,308	1,107,968	1,106,784	1,086,368	1,086,120	1,085,524
Adjusted R^2	0.544	0.579	0.596	0.545	0.579	0.595
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No	No	Yes	No
religion*state*product*district*quarter	No	No	Yes	No	No	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. In columns 1-3, status is measured by local returns to the national occupational mix of each religion. In columns 4-6, status is measured by national returns to the initial local occupational mix of each religion. All regressions include the household controls used in Subramanian and Deaton (1996): log of household size, household demographic shares by age and gender, and indicators for being self-employed and working in the agricultural sector. Columns 1 and 4 include the baseline fixed effects, columns 2 and 5 add the fixed effects for cross-sectional identification and columns 3 and 6 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F.3: Costs, Price Elasticities and Identity, with Household Controls

	LHS Variable: Abstain from Consuming Good i								
	Baseline			Cross-section			Panel		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$taboo_i$=1	0.160***	0.166***	0.176***						
	(0.00219)	(0.0116)	(0.0104)						
ln p_i	0.0153***	0.0634***	0.0476***	0.00958***	0.0500***	0.0499***	0.00748***	0.0208***	0.0208***
	(0.00275)	(0.00374)	(0.00353)	(0.00256)	(0.00484)	(0.00483)	(0.00252)	(0.00397)	(0.00396)
sum ln p_j	-0.000974	-0.0161***	0.00686***	-0.000925	-0.00635***	-0.00255	-0.000848	-0.00459**	-0.00148
	(0.00170)	(0.00203)	(0.00194)	(0.00159)	(0.00245)	(0.00291)	(0.00159)	(0.00185)	(0.00210)
$taboo_i$=1 x ln p_i		-0.0612***	-0.0443***		-0.0507***	-0.0505***		-0.0164***	-0.0166***
		(0.00287)	(0.00260)		(0.00456)	(0.00454)		(0.00331)	(0.00331)
$taboo_i$=1 x sum ln p_j		0.0199***	-0.00683***		0.00712***	0.000706		0.00482***	0.00221
		(0.00137)	(0.00124)		(0.00226)	(0.00359)		(0.00113)	(0.00168)
sum (ln p_j x $taboo_j$)			-0.0456***			-0.00776			-0.00681**
			(0.00114)			(0.00475)			(0.00284)
$taboo_i$=1 x sum (ln p_j x $taboo_j$)			0.0486***			0.0106*			0.00627**
			(0.00117)			(0.00547)			(0.00291)
Observations	1,115,640	1,115,640	1,115,640	1,115,292	1,115,292	1,115,292	1,114,116	1,114,116	1,114,116
Adjusted R^2	0.539	0.540	0.550	0.576	0.576	0.576	0.594	0.594	0.594
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	No	No	Yes	Yes	Yes	No	No	No
religion*state*product*district*quarter	No	No	No	No	No	No	Yes	Yes	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. Columns 1, 4 and 7 include own and cross-price elasticities. Columns 2, 5 and 8 add the interaction between taboo and own and cross-price elasticities. Columns 3, 6 and 9 allow cross-price elasticities to differ depending on whether both goods are taboos. All regressions include the household controls used in Subramanian and Deaton (1996): log of household size, household demographic shares by age and gender, and indicators for being self-employed and working in the agricultural sector. Columns 1-3 include the baseline fixed effects, columns 4-6 add the fixed effects for cross-sectional identification and columns 7-9 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F.4: Linear Approximation of Identity Choice with Cost, Status and Conflict, Household Controls

	LHS Variable: Share Spent on Good i		
	(1)	(2)	(3)
	Baseline	Cross-section	Panel
$(\overline{x_{ir}} - \overline{x_{is}}) \times (cost_r - cost_s)$	-0.0234	-0.660***	-0.692***
	(0.0545)	(0.0975)	(0.102)
$(\overline{x_{ir}} - \overline{x_{is}}) \times (status_r - status_s)$	0.481***	0.237***	0.222***
	(0.0274)	(0.0273)	(0.0639)
$(\overline{x_{ir}} - \overline{x_{is}}) \times conflict_r \ +/- \ 6\, months$	0.577***	0.0982***	0.273***
	(0.0474)	(0.0374)	(0.106)
Observations	32,437,780	32,430,340	32,350,360
Adjusted R^2	0.766	0.772	0.780
log price and total expenditure controls	Yes	Yes	Yes
household controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No
religion*state*product*district*quarter	No	No	Yes

Notes: Dependent variable is the share spent on good i in total food expenditure. $\overline{x_{ir}} - \overline{x_{is}}$ is the difference between prototypical religious and ethnic budget share spent on good i. $cost_r - cost_s$ is the difference in religious and ethnic Stone price indexes leaving out the cost of good i. $status_r - status_s$ is the difference between religious and ethnic status measured by national returns to the initial local occupational mix of religion and ethnicity. $conflict_r + / - 6\, months$ is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed. Columns 1-3 include the household controls used in Subramanian and Deaton (1996): log of household size, household demographic shares by age and gender, and indicators for being self-employed and working in the agricultural sector. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * p < 0.10, ** p < 0.05, *** p < 0.01.

G Baseline Taboo Regressions with Ovo-Pesco Vegetarianism

Table G.1: Religious Conflict and Taboo Adherence, Ovo-Pesco Vegetarianism

	LHS Variable: Abstain from Consuming Good i					
	Baseline	Cross-section	Panel	Panel		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	Urban	Rural
taboo=1	0.160***					
	(0.00242)					
conflict +/- 6 months	-0.102***	-0.0204	-0.0802**			
	(0.0246)	(0.0250)	(0.0328)			
taboo=1 × conflict +/- 6 months	0.106***	0.0377***	0.0447***			
	(0.0138)	(0.00747)	(0.00773)			
conflict past (6 months)				-0.0756**	0.0302	-0.303***
				(0.0345)	(0.0289)	(0.0707)
conflict present/future (6 months)				-0.0346	-0.0148	-0.0849
				(0.0387)	(0.0286)	(0.0703)
taboo=1 × conflict past (6 months)				0.0481***	0.0383**	0.0474***
				(0.00905)	(0.0174)	(0.0117)
taboo=1 × conflict present/future (6 months)				0.0314***	0.0640***	0.00868
				(0.0108)	(0.0208)	(0.0135)
Observations	1,115,640	1,115,292	1,114,116	1,114,116	347,556	764,344
Adjusted R^2	0.391	0.441	0.463	0.463	0.531	0.462
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No	No	No	No
religion*state*product*district*quarter	No	No	Yes	Yes	Yes	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. The vegetarian taboo is restricted to abstention of red meat and chicken (excluding fish and eggs). Conflict is an indicator for at least one occurrence of Hindu-Muslim conflict in the district. Columns 1-3 consider a conflict occurrence in the six months before or after the household is surveyed. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Columns 4-6 differentiate the effect of a conflict occurrence in the previous 6 months (past) and in the current or next 6 months (present/future) after the household is surveyed. Column 5 restricts the analysis to the urban population, and column 6 to the rural population. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table G.2: Status and Choice of Identity, Ovo-Pesco Vegetarianism

	LHS Variable: Abstain from Consuming Good i					
	Baseline	Cross-section	Panel	Baseline	Cross-section	Panel
	(1)	(2)	(3)	(4)	(5)	(6)
taboo=1	-0.129***			-0.858***		
	(0.0247)			(0.0535)		
status$_{rdt}^{national_o cc(r)}$	-0.397***	0.00116	-0.0256			
	(0.0188)	(0.0185)	(0.0167)			
taboo=1 $\times$ status$_{rdt}^{national_o cc(r)}$	0.104***	0.0492***	0.0229***			
	(0.00805)	(0.00817)	(0.00772)			
status$_{rdt}^{national_w(o)}$				-0.304***	-0.0245*	0.00282
				(0.0155)	(0.0132)	(0.0211)
taboo=1 $\times$ status$_{rdt}^{national_w(o)}$				0.330***	0.104***	-0.0155
				(0.0173)	(0.0158)	(0.0170)
Observations	1,111,072	1,110,724	1,109,544	1,089,132	1,088,876	1,088,280
Adjusted R^2	0.393	0.441	0.463	0.393	0.440	0.460
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No	No	Yes	No
religion*state*product*district*quarter	No	No	Yes	No	No	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. The vegetarian taboo is restricted to abstention of red meat and chicken (excluding fish and eggs). In columns 1-3, status is measured by local returns to the national occupational mix of each religion. In columns 4-6, status is measured by national returns to the initial local occupational mix of each religion. Columns 1 and 4 include the baseline fixed effects, columns 2 and 5 add the fixed effects for cross-sectional identification and columns 3 and 6 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table G.3: Costs, Price Elasticities and Identity, Ovo-Pesco Vegetarianism

	LHS Variable: Abstain from Consuming Good i								
	Baseline			Cross-section			Panel		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
taboo$_i$=1	0.166***	0.185***	0.202***						
	(0.00244)	(0.0133)	(0.0118)						
ln p$_i$	0.0139***	0.0297***	0.0260***	0.00608**	0.0492***	0.0489***	0.00382	0.0189***	0.0188***
	(0.00287)	(0.00369)	(0.00355)	(0.00263)	(0.00513)	(0.00509)	(0.00261)	(0.00433)	(0.00432)
sum ln p$_j$	-0.00112	-0.00477**	0.0174***	-0.00155	-0.00776***	0.000441	-0.00183	-0.00435**	0.000562
	(0.00189)	(0.00220)	(0.00214)	(0.00174)	(0.00280)	(0.00341)	(0.00172)	(0.00204)	(0.00239)
taboo$_i$=1 x ln p$_i$		-0.0200***	-0.0194***		-0.0536***	-0.0533***		-0.0185***	-0.0184***
		(0.00268)	(0.00255)		(0.00477)	(0.00472)		(0.00367)	(0.00365)
taboo$_i$=1 x sum ln p$_j$		0.00487***	-0.0194***		0.00804***	-0.00389		0.00325***	-0.000696
		(0.00134)	(0.00129)		(0.00252)	(0.00386)		(0.00123)	(0.00186)
sum (ln p$_j$ x taboo$_j$)			-0.0542***			-0.0166***			-0.0109***
			(0.00134)			(0.00571)			(0.00351)
taboo$_i$=1 x sum (ln p$_j$ x taboo$_j$)			0.0573***			0.0207***			0.00993***
			(0.00136)			(0.00637)			(0.00355)
Observations	1,115,640	1,115,640	1,115,640	1,115,292	1,115,292	1,115,292	1,114,116	1,114,116	1,114,116
Adjusted R^2	0.390	0.390	0.406	0.441	0.441	0.441	0.463	0.463	0.463
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	No	No	Yes	Yes	Yes	No	No	No
religion*state*product*district*quarter	No	No	No	No	No	No	Yes	Yes	Yes

Notes: Dependent variable is an indicator for abstaining from good i. Taboo is an indicator equal to 1 if the good is considered a taboo for the religion of the household. The vegetarian taboo is restricted to abstention of red meat and chicken (excluding fish and eggs). Columns 1, 4 and 7 include own and cross-price elasticities. Columns 2, 5 and 8 add the interaction between taboo and own and cross-price elasticities. Columns 3, 6 and 9 allow cross-price elasticities to differ depending on whether both goods are taboos. Columns 1-3 include the baseline fixed effects, columns 4-6 add the fixed effects for cross-sectional identification and columns 7-9 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

H Linear Approximation of Identity Choice

Table H.1: Linear Approximation of Identity Choice with Cost, Status and Conflict, Not Restricting Symmetry of Religious and Ethnic Identities

	LHS Variable: Share Spent on Good i		
	(1) Baseline	(2) Cross-section	(3) Panel
$\overline{x_{ir}} \times (cost_r - cost_s)$	0.0836*	-0.340***	-0.388***
	(0.0479)	(0.0947)	(0.0980)
$\overline{x_{is}} \times (cost_r - cost_s)$	0.0410	0.586***	0.625***
	(0.0586)	(0.0935)	(0.0981)
$\overline{x_{ir}} \times (status_r - status_s)$	0.311***	0.152***	0.0486
	(0.0235)	(0.0239)	(0.0594)
$\overline{x_{is}} \times (status_r - status_s)$	-0.492***	-0.249***	-0.249***
	(0.0268)	(0.0279)	(0.0653)
$\overline{x_{ir}} \times conflict_r \ +/- \ 6\,months$	0.586***	0.0869**	0.219**
	(0.0429)	(0.0362)	(0.0971)
$\overline{x_{is}} \times conflict_r \ +/- \ 6\,months$	-0.408***	-0.147**	-0.592**
	(0.0650)	(0.0647)	(0.275)
Observations	32,523,464	32,515,776	32,435,920
Adjusted R^2	0.766	0.772	0.780
log price and total expenditure controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No
religion*state*product*district*quarter	No	No	Yes

Notes: Dependent variable is the share spent on good i in total food expenditure. $\overline{x_{ir}}$ and $\overline{x_{is}}$ are, respectively, the prototypical religious and ethnic budget share spent on good i. $cost_r - cost_s$ is the difference in religious and ethnic Stone price indexes leaving out the cost of good i. $status_r - status_s$ is the difference between religious and ethnic status measured by national returns to the initial local occupational mix of religion and ethnicity. $conflict_r + / - 6\,months$ is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table H.2: Linear Approximation of Identity Choice with Cost, Status and Conflict, by Religion

	LHS Variable: Share Spent on Good i		
	(1) Baseline	(2) Cross-section	(3) Panel
Hindu SC $\times(\overline{x_{ir}} - \overline{x_{is}}) \times (cost_r - cost_s)$	0.201***	-0.631***	-0.656***
	(0.0557)	(0.134)	(0.125)
Hindu UC $\times(\overline{x_{ir}} - \overline{x_{is}}) \times (cost_r - cost_s)$	-0.0905	-0.687***	-0.692***
	(0.0582)	(0.121)	(0.119)
Muslim $\times(\overline{x_{ir}} - \overline{x_{is}}) \times (cost_r - cost_s)$	0.304***	-0.526***	-0.792***
	(0.0733)	(0.186)	(0.167)
Christian $\times(\overline{x_{ir}} - \overline{x_{is}}) \times (cost_r - cost_s)$	-0.121	-0.696	-0.359
	(0.213)	(0.452)	(0.369)
Hindu SC $\times(\overline{x_{ir}} - \overline{x_{is}}) \times (status_r - status_s)$	-0.0445	0.0388	-0.147
	(0.0436)	(0.0436)	(0.103)
Hindu UC $\times(\overline{x_{ir}} - \overline{x_{is}}) \times (status_r - status_s)$	1.576***	0.839***	1.261***
	(0.0763)	(0.0891)	(0.268)
Muslim $\times(\overline{x_{ir}} - \overline{x_{is}}) \times (status_r - status_s)$	0.356***	0.142***	0.342**
	(0.0576)	(0.0475)	(0.138)
Christian $\times(\overline{x_{ir}} - \overline{x_{is}}) \times (status_r - status_s)$	0.202*	0.208**	0.673***
	(0.104)	(0.0846)	(0.258)
Hindu SC $\times(\overline{x_{ir}} - \overline{x_{is}}) conflict_r \ +/- \ 6\, months$	0.489***	0.0960***	0.236**
	(0.0510)	(0.0364)	(0.109)
Hindu UC $\times(\overline{x_{ir}} - \overline{x_{is}}) conflict_r \ +/- \ 6\, months$	0.533***	0.112**	0.280**
	(0.0579)	(0.0441)	(0.132)
Muslim $\times(\overline{x_{ir}} - \overline{x_{is}}) conflict_r \ +/- \ 6\, months$	0.591***	0.120***	0.377***
	(0.0520)	(0.0423)	(0.108)
Christian $\times(\overline{x_{ir}} - \overline{x_{is}}) conflict_r \ +/- \ 6\, months$	0	0	0
	(.)	(.)	(.)
Observations	32,523,464	32,515,776	32,435,920
Adjusted R^2	0.766	0.772	0.780
log prices and total expenditure controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No
religion*state*product*district*quarter	No	No	Yes

Notes: Dependent variable is the share spent on good i in total food expenditure. $\overline{x_{ir}} - \overline{x_{is}}$ is the difference between prototypical religious and ethnic budget share spent on good i. $cost_r - cost_s$ is the difference in religious and ethnic Stone price indexes leaving out the cost of good i. $status_r - status_s$ is the difference between religious and ethnic status measured by national returns to the initial local occupational mix of religion and ethnicity. $conflict_r + / - 6\, months$ is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed. All differences are interacted with the religion of the household: Hindu scheduled caste (SC), Hindu upper caste (UC), Muslim or Christian. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and column 3 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table H.3: Linear Approximation of Identity Choice with Cost, Status and Conflict, Including Cross-Price Effects

	(1) Baseline	(2) Cross-section	(3) Panel	(4) Baseline	(5) Cross-section	(6) Panel
$(\overline{x_{ir}} - \overline{x_{is}}) \times (cost_r - cost_s)$	-0.00542	-0.698***	-0.769***	0.0342	-0.640***	-0.707***
	(0.0610)	(0.108)	(0.115)	(0.0600)	(0.107)	(0.114)
$(\overline{x_{ir}} - \overline{x_{is}}) \times (status_r - status_s)$	0.478***	0.237***	0.233***	0.475***	0.237***	0.230***
	(0.0302)	(0.0321)	(0.0743)	(0.0301)	(0.0321)	(0.0741)
$(\overline{x_{ir}} - \overline{x_{is}}) \times conflict_r \ +/- \ 6\,months$	0.566***	0.0932**	0.306**	0.563***	0.0931**	0.297**
	(0.0522)	(0.0453)	(0.122)	(0.0522)	(0.0452)	(0.121)
Observations	16,258,355	16,249,006	16,137,953	16,126,647	16,117,367	16,007,214
Adjusted R^2	0.768	0.775	0.784	0.770	0.776	0.785
log price and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
cross-price effects	No	No	No	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No	No	Yes	No
religion*state*product*district*quarter	No	No	Yes	No	No	Yes

Notes: Dependent variable is the share spent on good i in total food expenditure. $\overline{x_{ir}} - \overline{x_{is}}$ is the difference between prototypical religious and ethnic budget share spent on good i. $cost_r - cost_s$ is the difference in religious and ethnic Stone price indexes leaving out the cost of good i. $status_r - status_s$ is the difference between religious and ethnic status measured by national returns to the initial local occupational mix of religion and ethnicity. $conflict_r +/- 6\,months$ is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the six months before or after the household is surveyed. For computational feasibility, the table is based on a random 50 percent subsample at the religion-district-time level. Columns 4-6 include cross-price terms with respect to a Stone price aggregator of thirteen food product groups designated in the NSS product classification (e.g. cereals, fruits etc.). Columns 1 and 4 include the baseline fixed effects, columns 2 and 5 add the fixed effects for cross-sectional identification and columns 3 and 6 for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I Counterfactuals

Figure I.1: Population Changing Identity by Religion, 1987-2000

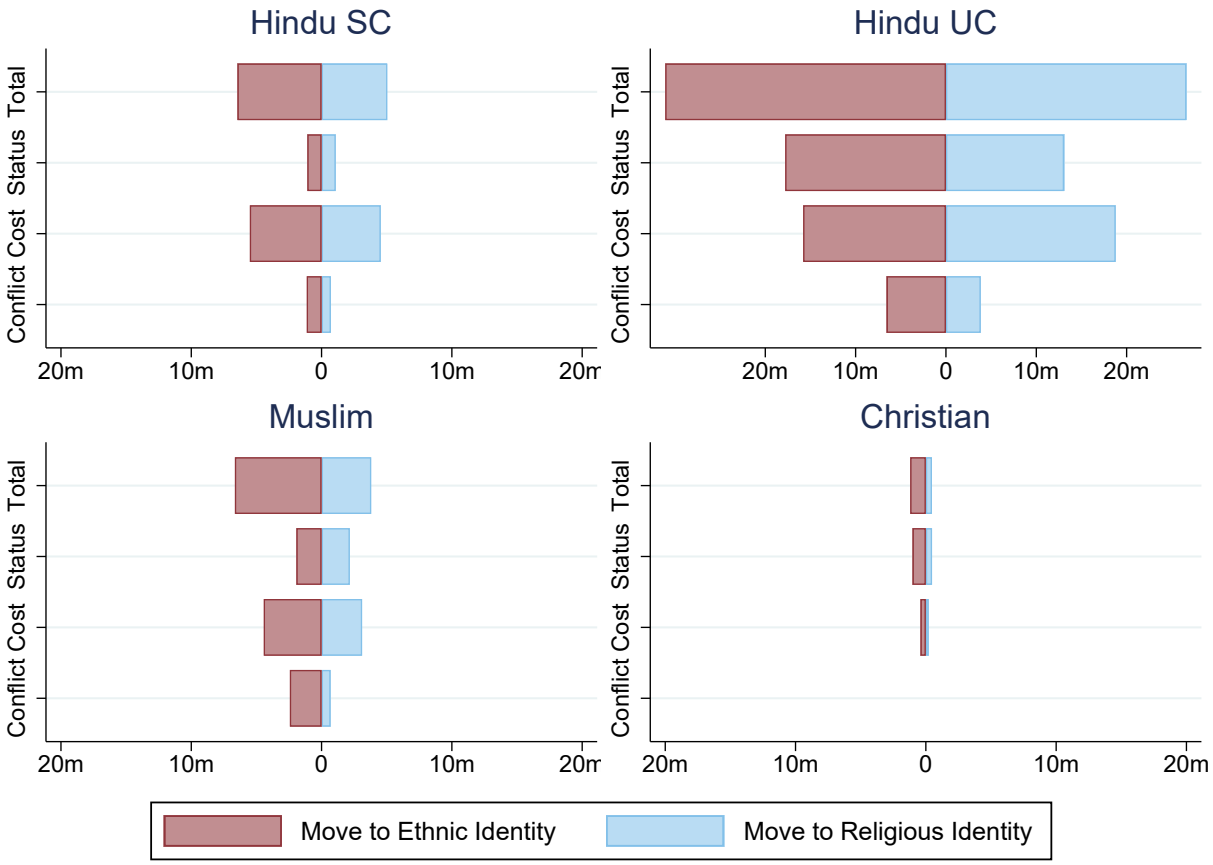


Figure I.2: Realized Compensating Variation Gains from Identity Changes, 1987-2000

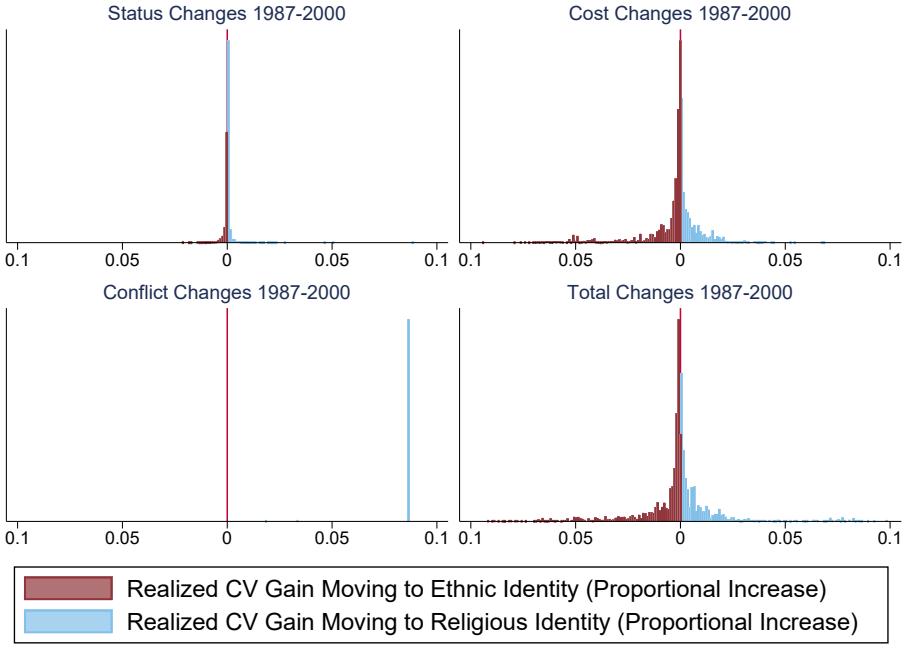


Figure I.3: Potential Compensating Variation Gains from Identity Changes, 1987-2000

