

NBER WORKING PAPER SERIES

HOW DO WE CHOOSE OUR IDENTITY? A REVEALED PREFERENCE APPROACH  
USING FOOD CONSUMPTION

David Atkin  
Eve Colson-Sihra  
Moses Shayo

Working Paper 25693  
<http://www.nber.org/papers/w25693>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
March 2019

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, Jon Eguia, Armin Falk, David Genesove, Marco Gonzalez-Navarro, Matthew Gentzkow, Ori Heffetz, Supreet Kaur, David Laitin, Edward Lazear, ShabanaMitra, SharunMukand, Salvatore Nunnari, Ran Shorrer, Katia Zhuravskaya, the editors and four anonymous referees, and many seminar participants for valuable comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by David Atkin, Eve Colson-Sihra, and Moses Shayo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

How Do We Choose Our Identity? A Revealed Preference Approach Using Food Consumption  
David Atkin, Eve Colson-Sihra, and Moses Shayo  
NBER Working Paper No. 25693  
March 2019, Revised June 2020  
JEL No. D12,D74,D91,O1,Z1

### **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. Notably, our estimated identity changes correlate with changes in vote shares for ethnic and religious parties. While social-identity research has focused on status and salience, our results suggest that economic costs also play an important role.

David Atkin  
MIT Department of Economics  
The Morris and Sophie Chang Building, E52-550  
50 Memorial Drive  
Cambridge, MA 02142  
and NBER  
atkin@mit.edu

Eve Colson-Sihra  
Department of Economics  
The Hebrew University of Jerusalem  
Jerusalem 91905, Israel  
eve.cs@mail.huji.ac.il

Moses Shayo  
Department of Economics  
The Hebrew University of Jerusalem  
Jerusalem 91905, Israel  
mshayo@huji.ac.il

A data appendix is available at <http://www.nber.org/data-appendix/w25693>

# 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). Such identity shifts have also been linked to changes in trade policies and opposition to globalization (Grossman and Helpman Forthcoming), with the political repercussions seen in Brexit, the resurgence of the far right in Europe, and Donald Trump’s election victory. Similar forces appear to be playing 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. Identity is usually conceptualized in terms of preferences, i.e., it is part of the “self” whose self-interest we seek to maximize. Thus, to identify with different groups means to care about different things (for reviews see Haslam and Ellemers 2005; Charness and Chen 2020; Shayo 2020). However, to measure identity outside the lab, much of the existing literature relies on survey questions or on ethnographic and historical case studies. But while case studies can provide important insights, they are often very specific and evaluating causal arguments is hard. Surveys offer the possibility of broad representative samples, but it is at present less clear to what extent different self-reported identity measures reflect day-to-day economic behavior.<sup>1</sup> Finally, experiments 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), female labor supply (Bertrand et al. 2015), grading decisions (Feld et al. 2016; Lavy et al. 2018), and conflict (Depetris-Chauvin and Durante 2019). We build on these recent advances.

Our approach starts from the observation that consumption choices are both widely docu-

---

<sup>1</sup>With the increasing interest in identity and polarization, survey questions on identity are likely to become more common. Our approach suggests a method for validating such questions in the spirit of Falk et al. (2018)—by testing whether they capture non-hypothetical, costly decisions. For studies correlating survey and behavioral measures of identity see Ellemers et al. (1999); Klor and Shayo (2010); Bankert et al. (2017).

mented 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 India 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 they cannot (easily) choose to be Muslim or Tamil, they 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 that identifies with a group suffers a utility loss if her consumption bundle is far from the prescribed or prototypical bundle of that group. In other words, identification with a group means, among other things, seeking to consume the goods that are acceptable to 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. 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 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 into 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 implications for health and welfare. Furthermore, as a proof of concept, we compare the identity changes inferred from consumption data with an alternative proxy for identity choices—changes in votes for religious and ethnic parties.

We begin with outbreaks of Hindu-Muslim violence. A long line of scholarship in political science argues that inter-group conflict increases the salience of membership in—and identification with—the fighting groups (e.g. Horowitz 1985; Kaufmann 1996). Are consumption patterns consistent with this argument? We use a combination of cross-sectional and panel regressions exploiting variation in exposure to religious conflict for 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. Nonetheless, we cannot fully dismiss all concerns related to reverse causality (identity shifts leading to conflict).

A somewhat different interpretation of these findings (and potentially later ones related to

status and costs) is that rather than directly changing identities, conflict alters social connections within and across identity groups, and that these changed social circles affect consumption patterns. One possibility is that members of different social circles more strongly reward (or punish) a person's adherence to a given taboo. However, the fact that our consumption-based estimates correlate with voting behavior—which is less observable—suggests that the effects go beyond the mere observability of your consumption by others. A second possibility is that expanded interactions and greater behavioral congruence with a particular group enhances one's identification with that group. For example, interacting more with members of your religious group may make you more likely to identify with your religion. This would then be manifested even in strictly private acts. Our results are fully consistent with—and our conceptual framework isomorphic to—this possibility.

Turning to ethnic identity, we study 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. The reverse is also true: inter-ethnic differences in cereal consumption decline with religious conflict.

Next, we borrow from the sociology literature 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.

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 taboo goods. 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.

Our conceptual framework thus captures important and previously undocumented regularities in the consumption of prominent religious taboos and ethnic staples. While this has the advantage of using easily recognized identity goods, there is little reason to suppose that identity effects are limited only to these prominent examples. Our next step is therefore to estimate a full demand system where we do not impose which goods are ‘identity goods’, nor what the ‘appropriate’ level of consumption is. Generalizing the above analysis, households suffer a utility loss when their consumption *bundle* is far from the prototypical bundle of the group they identify with (which we take to be the observed mean bundle in the group). We thus include all 124 food items and jointly consider all three determinants of identity. We find that both religious conflict and a rising relative status of one’s religious group shift households towards the prototypical bundle of their religious group and away from that of their ethnic group. In contrast, household consumption diverges from their religion’s prototypical bundle and shifts toward their ethnic bundle when the cost of their religious bundle rises relative to that of their ethnic bundle (in all cases conditioning on standard price and income effects).

Our demand system estimates also allow us to quantify, at least in an approximate sense, the relative importance of these three determinants over India’s tumultuous reform period 1987-2000.<sup>2</sup> Our estimates suggest large shifts in identity choices. However, substantial heterogeneity across districts and religions led to only a small shift from ethnic to religious identities on net. While conflict-driven local 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 changes in the economic costs of identifying with a group have large impacts is surprising given that this channel has received little attention in the social-identity literature. However, such a finding can broaden our understanding of recent and much-discussed efforts by the Hindu-nationalist BJP party to raise the effective price of beef through bans and legislation.<sup>3</sup> Our results suggest that, intentionally or not, such a strategy may possibly contribute to the hardening of Hindu identities, much like inciting religious violence (a tactic documented by Wilkinson 2004). Indeed, we find that changes in identification patterns revealed by consump-

---

<sup>2</sup>This reform period has been extensively studied (e.g. see Aghion et al. (2008), Topalova (2010), or Martin et al. (2017) for the impacts of removing licensing, trade restrictions, and small-scale-industry reservations, respectively).

<sup>3</sup>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.

tion choices are associated with changes in the vote shares of religious and ethnic parties in state elections. This insight might potentially 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.<sup>4</sup>

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 of how norms influence economic behavior (e.g. Young 2015), 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). 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

---

<sup>4</sup>Item level inflation between 2003 and February 2019 from the Turkish Statistics Institute (TUIK), <https://biruni.tuik.gov.tr/medas/?kn=84&locale=en>.



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. 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, compares our consumption results to changes in voting patterns, and evaluates health and welfare implications. Section 7 concludes.

## 2 Conceptual Framework

We present a brief exposition of the social identity framework that generates the three key comparative statics we bring to the data. See Shayo (2020) for a more detailed exposition as well as a review of evidence supporting the basic assumptions of the model.

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  $\bar{x}_J$  the *prescribed* consumption bundle of social group  $J$ . For example, if group  $J$  has a taboo on the consumption of good  $i$ , then  $\bar{x}_{iJ} = 0$ . Identifying with group  $J$  then implies seeking to reduce the distance  $d(x_h, \bar{x}_J)$  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  $\kappa_{h,J}$  to denote shocks to the salience of  $h$ 's membership in group  $J$ .<sup>5</sup> The second feature of identification with a group is caring about its status (Tajfel and Turner 1979). We

---

<sup>5</sup>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.

denote group  $J$ 's status by  $y_J$ . Defining identity in terms of preferences—rather than in terms of beliefs—leads us to the following simple definition.

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

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

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

Let  $x_{hJ}^*(p, m_h, y_J, \kappa_{hJ}; \bar{x}_J)$  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*

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

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

$$V_{hJ}(p, m_h, y_J, \kappa_{hJ}; \bar{x}_J) = U(x_{hJ}^*, y_J, \kappa_{hJ}; \bar{x}_J). \quad (2)$$

Thus, the choice of identity itself responds systematically to the social and economic environment.<sup>6</sup> 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  $\bar{x}_J$  (e.g., respect group  $J$ 's taboos) when:

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

---

<sup>6</sup>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  $\frac{\partial}{\partial p_1}(V_{hA} - V_{hB}) < 0$ .<sup>7</sup> 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.

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 and 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). As noted in the introduction, this modeling choice is also consistent with shocks to salience, status and costs altering the type and nature of social interactions within and across identity groups, with these changed interactions leading to identity changes as defined above. However, there may also be non-psychological benefits associated with a low distance  $d(x_h, \bar{x}_J)$  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 clearly distinguish between there being psychological or material benefits from respecting the group's prototypical behavior (although, to the extent that voting is not as easy for others to observe, our voting results tend to support the former interpretation). This distinction only matters for our welfare calculations in Section 6.4.

A final remark. While in a more general framework, group status  $y_J$ , salience  $\kappa_{hJ}$ , 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.

<sup>7</sup>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  $\beta^J$ '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:<sup>8</sup> 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.<sup>9</sup> 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 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.<sup>10</sup>

**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.

---

<sup>8</sup>The Mitra and Ray (2014) conflict data end in 2000. The resulting sample restriction has the additional benefit of spanning the major (and externally-enforced) economic reforms that India embarked on in 1991, reshaping its economy. These policy changes plausibly drive much of the price and income variation we use to analyze the status and economic cost channels. That said, as we show in Appendix Table I.5, the status and price results are similar when including the 61st (2004-2005) and 66th (2009-2019) rounds.

<sup>9</sup>As only region identifiers are available for urban households in round 50, we exclude these households from any district-level analysis. Results are robust to excluding the whole of round 50, as shown in Appendix G.

<sup>10</sup>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.

## 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” 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, and 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. Defining and measuring ethnicity in India is, however, conceptually challenging (see e.g. Fearon 2003; Somanathan 2018 for discussion), and the NSS does not provide linguistic or genetic data.

We propose two approaches to address this problem. Our primary approach leverages the fact that in response to the demands of ethnic groups following Independence, India chose to draw its state borders along ethno-linguistic lines.<sup>11</sup> We can thus use the state of residence of the household in the NSS surveys as a proxy for ethnicity, sidestepping the conceptual issues by employing an unambiguous geographic notion of identity. One may also interpret the results using this measure as indicating the strength of identification with one’s state.<sup>12</sup>

Our second measure of ethnicity uses the mother tongue most commonly spoken in the district (e.g. Hindi, Marathi, Telugu, Bengali, or Gujarati), as recorded in the 2001 Census of India. An ethnic group is then proxied by households living in the set of districts where the most-commonly-spoken mother tongue is the same. This yields 44 ethnic groups across India.<sup>13</sup>

---

<sup>11</sup>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.

<sup>12</sup>See Atkin (2016) for evidence that inter-state migrants in India often carry their origin-state food preferences with them. Our results are robust to excluding inter-state migrants, see Appendix Table I.1.

<sup>13</sup>Appendix Figure A.2 shows linguistic and religious fractionalization (one minus the Herfindahl of group shares)

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

### 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.<sup>14</sup> Pork consumption is equally if not more taboo for Muslims.<sup>15</sup> 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.<sup>16</sup> Lower-caste Hindus do not widely share this vegetarian norm, and in fact have developed a non-vegetarian cuisine particular to their identity.<sup>17</sup> Our primary definition of vegetarianism includes avoidance of eggs, poultry and fish. Appendix H 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.<sup>18</sup> Table 1 provides a schematic summary of these taboos.

Direct evidence for these 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

by state. Religious fractionalization is broadly similar in most states of India with an average of 0.58 across states, close to the overall level of religious fractionalization in India (0.61). In contrast, linguistic fractionalization is much higher when looking at India as a whole (0.8) than within states (average 0.31).

<sup>14</sup>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).

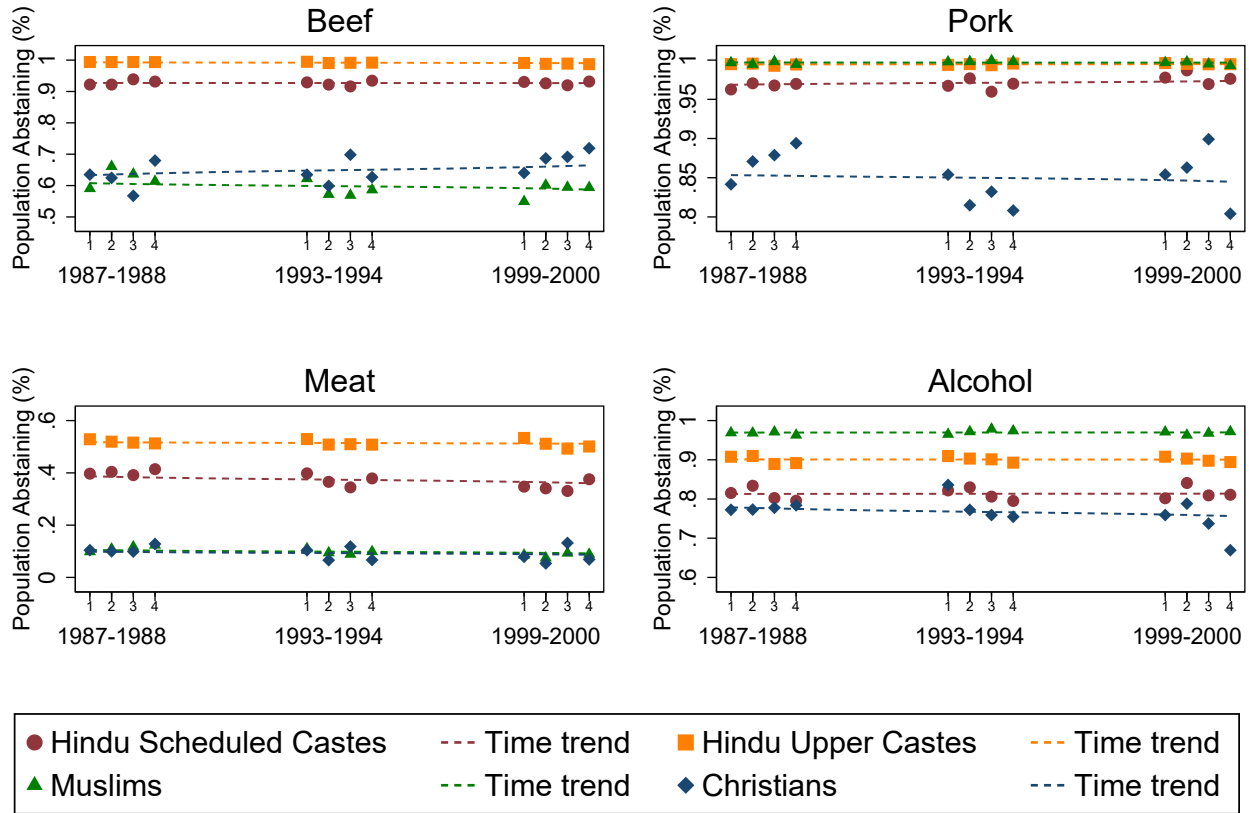
<sup>15</sup>The Qur’an, Surah Al-Baqarah 2:173 says: “He has forbidden you only carrion, blood, the flesh of the swine, and that which has been offered to other than Allah.”

<sup>16</sup>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, outside of an offering to gods or ancestors, wants to make his own flesh thrive at the expense of someone else’s.”

<sup>17</sup>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.

<sup>18</sup>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 of his own caste.”

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



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).

**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.<sup>19</sup> The most prominent example is the strong regional preferences for rice and wheat (Chakravarti 1974). Despite the fact that these two staple cereals provide similar nutrition per rupee, there is dramatic regional variation in their consumption and cooking practices. For example, despite similar relative prices, households in the state of Kerala consume thirteen times

<sup>19</sup>For example, Appadurai (1988) mentions the “longstanding and distinct regional cuisines” in India and the diversity of ethno-regional cookbooks. Sen (2014) lists regional food specificities in her history of food in India. Atkin (2013; 2016) also provide evidence of a regional component of food cultures in India.



more rice than wheat, while those in Punjab 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.3).

### 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.<sup>20</sup> The dataset records 505 Hindu-Muslim riots between 1987 and 2000, with about 4100 fatalities. Appendix Table A.3 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.

## 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 this analysis 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, focusing on two shocks to salience: conflict between religious groups and changes in political autonomy for ethnic groups. We then turn to group status and to the costs of identity goods.

---

<sup>20</sup>The dataset also records the duration and the reported proximate cause of the riot, but no information on which side initiated the violence. Appendix Table A.2 reports the causes of Hindu-Muslim conflicts during our period.



## 4.1 Religious Conflict and Identity

Inter-group conflict has long been associated with hardening group identities.<sup>21</sup> Exploring how conflict affects consumption therefore serves as a useful first test of our approach. This necessitates introducing our common empirical approach and discussing the key identification concerns in this section. That said, this is not a paper about conflict and, as we previewed in the introduction, our results suggest that costs and status are more important determinants of identity changes between 1987 and 2000.

Through the lens of our conceptual framework, conflict along religious lines raises the salience of membership in one's religious group ( $\kappa_{hr}$ ), thereby stimulating identification with that group. Indeed, in the Indian context, it has been argued that politicians use religious violence precisely to that end (Wilkinson 2004). This section therefore examines whether religious conflict is also associated with changes in consumption, consistent with a movement towards the prescribed behaviors of one's religion. Note that, from a standard economic perspective, while conflict can affect prices and incomes, conditional on those it should not affect consumption choices.<sup>22</sup>

### 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, 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.<sup>23</sup> The patterns are

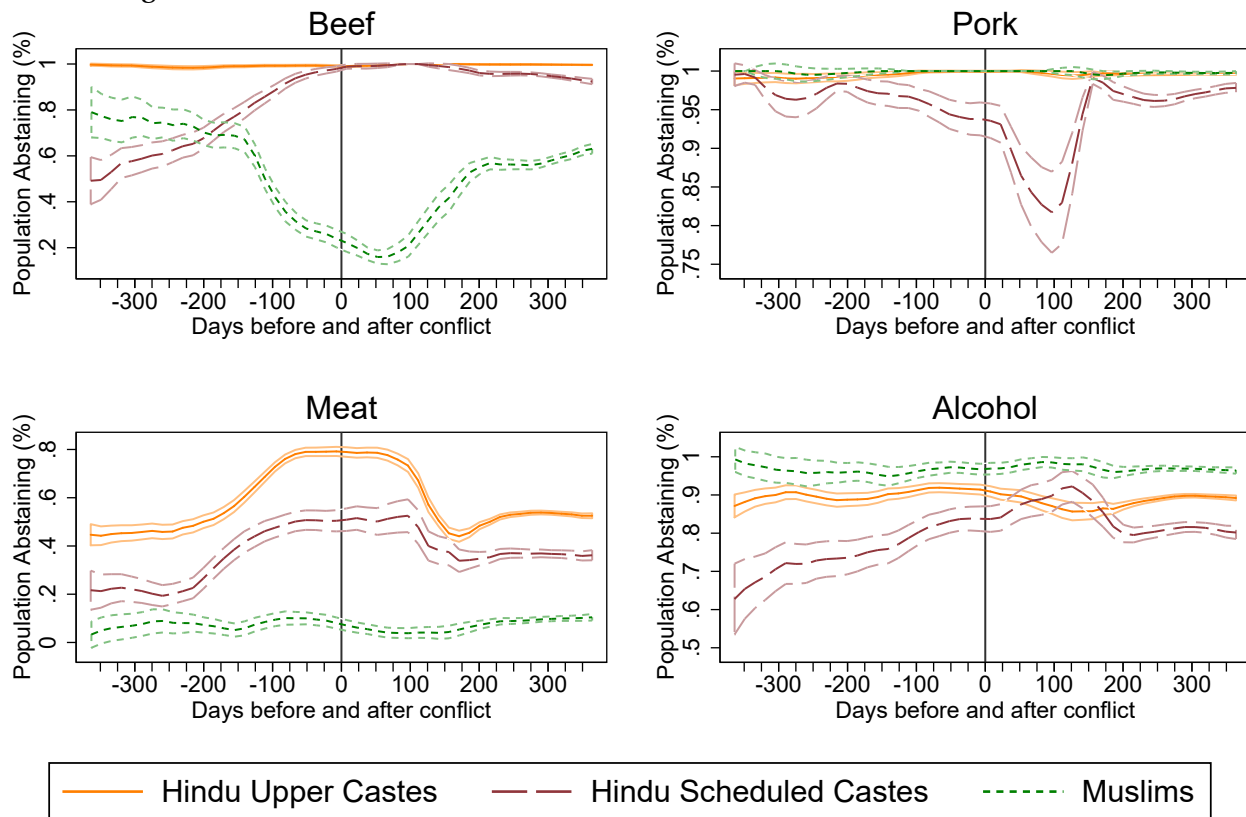
---

<sup>21</sup>See Shayo and Zussman (2017) for a recent discussion. *Security Studies* (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). Most relevant to our study, Iyer and Shrivastava (2018) show that Hindu-Muslim riots in the year preceding an election significantly increase the vote share of the Hindu nationalist BJP.

<sup>22</sup>Conflict could also lead individuals to want to hide or downplay their religious identity to avoid harm (see e.g. Kalyvas 2008).

<sup>23</sup>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.

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.

qualitatively similar when conditioning on a battery of controls and using a balanced sample.<sup>24</sup>

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.

This figure suggests that consumption patterns may potentially serve as leading indicators for the eruption of conflicts. Recorded conflict—as manifested by reports that reach the national press in India—is plausibly preceded by mounting inter-religious tensions that make religious identity salient even before these tensions boil over. However, as we discuss below, once we include multiple survey rounds and a battery of fixed effects, the association of taboo adherence with future (as opposed to present and past) reported conflict dissipates.

<sup>24</sup>Appendix Figure B.1.1 conditions on household per capita expenditure, local prices, and region-month fixed effects. Appendix Figure B.2.1 restricts attention to locations surveyed both in the 6 months before and after a conflict.

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. If religious tensions lead Muslims to identify more with their religious group and less with their ethnic group, then beef consumption should rise as they move between the two groups’ prescribed behaviors. Such behavior is consistent with 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.<sup>25</sup>

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. 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.2.2, 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.<sup>26</sup>

---

<sup>25</sup>An alternative explanation is that Muslims’ reduced pork consumption during times of Hindu-Muslim conflict leads to substitution into non-pork meats. While possible, such a large shift into beef requires large unreported changes in Muslim pork consumption. Additionally, Section 4.2 shows that within cereals—where a simple substitution story is less plausible—the share of ethnic staples declines with religious conflict.

<sup>26</sup>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.3.1). 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.2.3 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.

## 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  $i$ :

$$\begin{aligned} \text{Abstain}_{ihrdst} = & \alpha_1 \text{Taboo}_{ir} + \alpha_2 \text{Conflict}_{rdt} + \alpha_3 \text{Taboo}_{ir} \times \text{Conflict}_{rdt} \\ & + \sum_j \gamma_{1ij} \ln \text{price}_{jht} + \gamma_{2i} \ln \text{realfoode}xp_{ht} + \delta_{idt} + \delta_{rdt} + \text{Additional\_FE} + \epsilon_{iht} \quad (3) \end{aligned}$$

where  $\text{Abstain}_{ihrdst}$  is an indicator variable that takes the value 1 if the household does not consume good  $i$ ;  $\text{Taboo}_{ir}$  is an indicator for good  $i$  being a taboo for religion  $r$  (from Table 1); and  $\text{Conflict}_{rdt}$  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  $\text{Conflict}_{rdt} = 0$  for Christian households.

The key coefficient of interest is  $\alpha_3$ , which captures how consumption differences across taboo and non-taboo goods vary with conflict. To take a concrete example, we expect Muslims are more likely to abstain from pork than beef ( $\alpha_1 > 0$ ). Conflict may also affect the consumption of goods that are taboo for *other* religions (captured by  $\alpha_2$ ). However, a positive  $\alpha_3$  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:  $\ln \text{price}_{jht}$  is the village median price of good  $j$  that controls for own- and cross-price effects; and  $\ln \text{realfoode}xp_{ht}$  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 product-district-time fixed effects  $\delta_{idt}$ . These fixed effects also absorb other factors that affect consumption of good  $i$  in district  $d$  at time  $t$  and may correlate with conflict.<sup>27</sup> Thus, we explore how taboo/non-taboo consumption differences between religions *within* a location vary with conflict. Finally, we include religion-district-time fixed effects,  $\delta_{rdt}$ , to account for possible forces affecting total consumption of our four prominent religious goods (relative to other goods) for a particular religion-location pair.

<sup>27</sup>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.

We cluster standard errors at the  $rdt$  level.<sup>28</sup>

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).

**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 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). In columns 1-3 we code  $Conflict_{rdt} = 1$  if any conflict incident occurred in the quarter in which the survey was administered or in the 6 months preceding that quarter. Appendix Table B.3.4 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 the current and preceding two quarters.

Our main interest is the interaction term in the second row. (Note that the main effect of conflict is swept out by the  $\delta_{rdt}$  fixed effects, and the main effect of taboo by either the  $\delta_{irst}$  or  $\delta_{irsdq}$  additional fixed effects). 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

---

<sup>28</sup>Results are robust to higher spatial clustering at the religion-region-time level (Appendix Table B.3.2), or clustering at the religion-district level to account for serial correlation (Appendix Table B.3.3).

Table 2: Religious Conflict and Taboo Adherence

	LHS Variable: Abstain from Consuming Good $i$						
	Baseline	Cross-section	Panel	Panel			
	(1)	(2)	(3)	Geographic Area		Cause of Conflict	
				Urban	Rural	Narrow	W/o Relig
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
taboo=1	0.181*** (0.00364)						
taboo $\times$ conflict	0.106*** (0.0162)	0.0396*** (0.0107)	0.0599*** (0.00981)			0.0660*** (0.0101)	0.0536*** (0.0175)
taboo $\times$ conflict past 2 quarters				0.0431** (0.0194)	0.0550*** (0.0134)		
taboo $\times$ conflict present quarter				0.0587** (0.0293)	0.0293 (0.0187)		
Observations	1,115,640	1,115,292	1,114,116	347,556	764,344	1,114,116	1,114,116
Adjusted $R^2$	0.560	0.585	0.596	0.612	0.603	0.596	0.596
log prices and total expenditure	Yes	Yes	Yes	Yes	Yes	Yes	Yes
prod*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*state*prod*round*quarter	No	Yes	No	No	No	No	No
relig*state*prod*district*quarter	No	No	Yes	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 quarter in which the household is surveyed or in the preceding two quarters. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and columns 3-7 for panel identification. Columns 4-5 differentiate the effect of a conflict occurring in the quarter of the survey from conflict in the preceding six months. Column 4 restricts the analysis to the urban population, and column 5 to the rural population. Column 6 excludes conflicts due to animal slaughter, economic, or political factors. Column 7 additionally excludes conflicts with religious reported causes. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

by 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 somewhat smaller in size but similar in significance, despite drawing on very different variation.<sup>29</sup> 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 18 percentage points higher if the good is a taboo for the household's religion. This difference increases by more than a half during times of conflict.

Appendix Table B.3.6 repeats the analysis using a non-binary measure of conflict, log fatalities, with similar results. There is also some evidence that the effect of conflict spills over to nearby districts, at least when exploiting cross-sectional variation (see Appendix Table B.3.7). Finally, as suggested by the earlier plots of beef consumption for Upper Caste Hindus and pork consumption for Muslims, effects may be more muted for products with very high rates of (local) abstention. Appendix Table B.3.8 provides some supportive evidence for such ceiling effects,

<sup>29</sup>Figure 2 showed larger unconditional effects for beef abstention using the date of survey that is only available in NSS 50. Appendix Table B.3.5 finds similarly large effects using the regression analysis restricted to beef in NSS 50.

although differences are generally insignificant for our preferred panel specification.

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.3.9, a 1-percentage-point increase in conflict-induced abstention rates leads to a 0.8 percent drop in prices.<sup>30</sup>

### **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 consumption and generate conflict. 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 Hindu-Muslim conflict that would explicitly rule out reverse causality, we can dig deeper into the time structure and the reported causes of conflict to provide some reassurance.

Our first approach exploits the fact, noted by both Varshney (2001) and Mitra and Ray (2014), that Hindu-Muslim riots are primarily an urban phenomenon.<sup>31</sup> Thus, if the conflict itself raises religious salience—thus causing changes in identification and consumption—the timing of the response would likely differ between rural and urban areas. Specifically, these urban conflicts would immediately increase religious salience in urban areas and hence affect consumption. In contrast, it is reasonable to expect the salience of religious cleavages to take time to spill over to adjacent rural areas, and so identity changes and the observed consumption response would likely be delayed.

To investigate this hypothesis, columns 4-5 of Table 2 break the conflict variable into two: conflicts that occurred in the six months preceding the quarter of the survey, and conflicts that occurred during the quarter of the survey itself. Consistent with the above scenario, among

---

<sup>30</sup>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.3.9 for details.

<sup>31</sup>Varshney (2001) states “Villages constitute a remarkably small portion of communal rioting. Between 1950-95, rural India, where a majority of Indians still live, accounted for a mere 3.6 per cent of the deaths in communal violence.”



urban households (column 4), consumption changes are significantly associated with religious conflict in both the current and the preceding quarters. By contrast, among rural households (column 5), the association is only with reports of religious conflict in the preceding quarters. This timing of consumption changes is consistent with the causal chain above whereby urban conflict raises religious salience but only with a lag in rural areas. For the simple reverse-causation worry outlined at the top of this subsection, forces that change identity other than salience would also have to be urban phenomena that spill over to rural areas with lags (a possibility of course, but not one previously documented by scholars).

A second and more direct approach to causality is to narrow the set of conflicts we consider, using the reported cause of conflict in the Varshney-Wilkinson dataset (see list of causes in Appendix Table A.2). For the association between taboo adherence and conflict to provide causal evidence for our hypothesis—that changes in religious salience causally affect identity choice—we require that either: 1a) exogenous shocks to inter-religious violence cause religious salience and hence taboo adherence to rise, or 1b) exogenous shocks to religious salience cause both inter-religious violence and taboo adherence to rise. However, our associations would be spurious if identity changes are driven by a shock to some force other than religious salience; and either 2a) those changes induce both greater taboo adherence and inter-religious violence, or 2b) the inter-religious violence is itself caused by the changes in taboo adherence.

To help dismiss these two spurious interpretations, we exclude conflicts that are most subject to the concern of being driven by consumption changes, or most at risk of being driven by other factors that affect taboo consumption differentially across religions (beyond raising the salience of religion that our *conflict* variable is designed to capture). Column 6 of Table 2 reports our preferred panel specification after removing conflicts that are due either to animal slaughter (where food consumption might potentially cause conflict as in case 2b), or to economic and political factors (which may be driven by changes in group status, identity costs, or other endogenous forces as in case 2a). This removes 11 percent of the conflicts with a reported cause, and leaves us only with more plausibly exogenous shocks to religious salience—primarily public rituals, the desecration of religious sites, previous violence, as well as private quarrels, homicides and sexual harassment. Reassuringly, the coefficients on the Taboo  $\times$  Conflict interaction change little (if anything, the coefficient is larger).<sup>32</sup>

A direct test of case 2a—non-salience induced identity changes causing conflict—is reported in column 7. If some external force changed identities which in turn spurred inter-religious violence, we would expect these conflicts to be most commonly attributed to religious causes. Thus,

---

<sup>32</sup>Appendix Table B.3.10 reports similar results with cross-section fixed effects and sequentially removing causes.



in column 7 we further remove all 40 percent of conflicts that have religious reported causes.<sup>33</sup> Again, the coefficient on the taboo interaction changes little.

Finally, Appendix Table B.3.11 shows that results are remarkably stable across multiple categories of reported cause (with the exception of the more endogenous causes noted above—animal slaughter and economic/political factors). As omitted variable concerns differ markedly across causes (e.g. previous communal violence versus public rituals), the similarity of the estimates suggests that any endogeneity bias is likely small.

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.<sup>34</sup> 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.<sup>35</sup> 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.<sup>36</sup> 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 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.<sup>37</sup> “Religious Fractionalization” (third row in Table 3) is calculated using all seven religions

---

<sup>33</sup>This removal also excludes case 1b above, and so leans only on the shocks to conflict in case 1a.

<sup>34</sup>A remaining concern is that butchers increase non-market transactions to specific religions rather than lowering prices. Note, however, that gifting of taboo goods is very uncommon (e.g. 0.26 percent of Muslims received beef gifts and 0.01 percent of Hindus received pork).

<sup>35</sup>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.3.1 shows butcher numbers and shares by religion.

<sup>36</sup>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.

<sup>37</sup>Fractionalization =  $1 - \sum \pi_r^2$  where  $\pi_r$  is the proportion of people in religion  $r$  in the neighborhood.

Table 3: Conflict and Religious Composition

	LHS Variable: Abstain from Consuming Good $i$					
	Cross-section			Panel		
	(1)	(2)	(3)	(4)	(5)	(6)
taboo $\times$ conflict	0.0430*** (0.0112)	0.0467*** (0.0120)	0.0490*** (0.0114)	0.0599*** (0.0101)	0.0635*** (0.0128)	0.0677*** (0.0122)
taboo $\times$ conflict $\times$ non hindu/muslim butcher share	-0.173 (0.222)			0.00218 (0.0613)		
taboo $\times$ conflict $\times$ religious fractionalization		-0.0324 (0.0368)			-0.0163 (0.0375)	
taboo $\times$ conflict $\times$ hindu/muslim fractionalization			-0.0491 (0.0410)			-0.0410 (0.0417)
Observations	1,107,484	1,115,292	1,115,292	1,106,292	1,114,116	1,114,116
Adjusted $R^2$	0.585	0.585	0.585	0.596	0.596	0.596
main effects and double interactions	Yes	Yes	Yes	Yes	Yes	Yes
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*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 quarter in which the household is surveyed or in the preceding two quarters. 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 1-3 include fixed effects for cross-sectional identification. Columns 4-6 include 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

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 are likely to 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 social 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,<sup>38</sup> but we have no information about

<sup>38</sup>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

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). The voting results in Section 6.2 are also helpful here since voting is less visible than consumption; hence social desirability bias is a less plausible explanation. A final point is that while there may be reasons other than identity for Muslims to misreport their pork consumption in times of religious tensions, it is harder to see why they would misreport their pork consumption when the price of alcohol changes (Section 4.4 below).

We view these robustness exercises as supportive of our preferred interpretation—that increases in the salience of intergroup differences that accompany intergroup conflict cause changes in identification. This interpretation is consistent with studies documenting the effect of conflict on identification patterns in other contexts, or using other outcome variables. Nonetheless, we acknowledge that threats to a causal interpretation may still remain, and one should not completely dismiss closely-related competing explanations. For example, differences in urban and rural life may mean that taboo adherence responds more quickly in urban areas. Additionally, conflict may alter social connections between and within identity groups in complicated ways, and these can also affect consumption patterns—an interesting finding in its own right.<sup>39</sup> As mentioned above, the fact that our estimates correlate with voting behavior suggests that the effects go beyond the mere observability of one’s consumption that is required for many social mechanisms to operate. It is, however, consistent with our preferred interpretation, or isomorphic variants where altered social interactions affect preferences themselves.

## 4.2 Ethnic Autonomy 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

---

headquarters and one by the state NSSO office, and then checks for discrepancies.

<sup>39</sup>For example, an alternative explanation for our results is that conflict reduces social interactions between members of different religious groups. Thus, Muslims no longer need to observe Hindu taboos at social gatherings, and Hindus no longer need to observe Muslim ones. However, under this explanation the effects should be attenuated when excluding the third of households who served meals to guests during the period covered by the expenditure survey (the number of meals served to guests, including at ceremonies, is recorded in NSS rounds 43 and 50). As shown in Appendix Table B.3.12, the effect size actually increases rather than falls with this restriction.

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 separated from Bihar; and the former kingdoms of Garwhal and Kumaon split from Uttar Pradesh to become the new state of Uttaranchal (now called Uttarakhand). These splits were preceded by a decade of political activism and agitation, and were the outcome of drawn-out political processes.<sup>40</sup> 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.3, 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. (And, as shown in Appendix Figures C.1 and C.2, the splits also largely followed linguistic lines.) 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  $i$ :

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

where  $\text{CerealShare}_{ihdst}$  is  $h$ 's share of expenditure spent on  $i \in \{\text{rice, wheat, other cereals}\}$ ,  $\text{EthnicCereal}_{is}$  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  $\text{Round}_{xx-yy}$  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  $\delta_{its_o}$  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  $\delta_{idq}$  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  $\delta_{idq}$  fixed effects absorb an  $\text{EthnicCereal}_{is} \times \text{Round}_{87-88}$  dummy. Hence,  $\theta_1$  reveals the increase in consumption of the ethnic cereal in the future state between the 43rd

---

<sup>40</sup>The 1990s saw region-wide strikes and rallies in Chhattisgarh led by the Chhattisgarh Rajya Nirman Manch political forum, with a resolution demanding an independent Chhattisgarh passing the Madhya Pradesh legislative assembly in 1994 followed by the Separate Chhattisgarh Bill in 1998. In Jharkhand, regional parties including the Jharkhand Mukti Mocha, All India Jharkhand Party and the All Jharkhand Students Union grew in importance before uniting with the local BJP and Congress parties to pass the Jharkhand Act in 1998.

and 50th round, and  $\theta_2$  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  $\theta_2$  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 districts so the  $\text{EthnicCereal}_{i,s} \times \text{Round}_{93-94}$  coefficient from a district-level analysis would conflate 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. Second, rice/wheat consumption was more mixed in these locations (see Figure A.3), 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 and 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

Table 4: Ethnic Goods and State Splits Along Ethnic Lines

	LHS Variable: Share Spent on Cereal $i$							
	Regions				Districts			
	(1) All	(2) Border	(3) All	(4) Border+Neighbor	(5) Border	(6) All	(7) Border+Neighbor	(8) Border
Ethnic Cereal $\times$ 1987-1988	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Ethnic Cereal $\times$ 1993-1994	0.0297** (0.0130)	0.0322** (0.0130)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Ethnic Cereal $\times$ 1999-2000	0.0689*** (0.0124)	0.0697*** (0.0122)	0.0554*** (0.00857)	0.0625*** (0.0129)	0.0928*** (0.0182)	0.0546*** (0.00858)	0.0608*** (0.0130)	0.0928*** (0.0182)
Ethnic Cereal $\times$ Hindu-Muslim Conflict						-0.0919** (0.0358)	-0.127*** (0.0273)	0 (.)
Observations	128,029	102,070	93,114	39,710	23,730	93,114	39,710	23,730
Adjusted $R^2$	0.732	0.755	0.793	0.830	0.836	0.793	0.830	0.836
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
original state*product*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
region*product*quarter	Yes	Yes	No	No	No	No	No	No
district*product*quarter	No	No	Yes	Yes	Yes	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 3-8 are district-level regressions: column 3 includes all districts, column 4 restricts to border and border-adjacent districts, column 5 to border districts, and columns 6-8 reproduce columns 3-8 adding an interaction between ethnic cereal and Hindu-Muslim conflict. All regressions include the baseline fixed effects controlling for local supply and demand conditions (original state-product-time) and the fixed effects for panel identification (region- or district-product-quarter). Robust standard errors clustered at region-round-quarter (columns 1-2) or district-round-quarter (columns 3-8) in parentheses. Regressions weighted by survey population weights. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

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.<sup>41</sup>

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.3, 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 religious conflict on the consumption of ethnic staples. Columns 6-8 of Table 4 add an interaction between the ethnic cereal dummy and Hindu-Muslim conflict to the previous specification. Consistent with our model of dueling identities—where religious salience leads households to move towards their religious identity and away from their ethnic identity—episodes of Hindu-Muslim conflict reduce inter-ethnic differences in staple food shares. There is also evidence suggesting that the converse relationship holds, i.e., that rising ethnic salience increases the consumption of *religious* taboo goods. Specifically, Appendix Table C.3 adapts equation (4) to the demand for taboo goods and shows that households adhere less strongly to their religious taboos as a state split approaches, especially in border districts (significantly so only 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.

### 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

<sup>41</sup>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):  $\ln\text{Price}_{idt} = \theta_1\text{EthnicCereal}_{is} \times \text{Round}_{93-94} + \theta_2\text{EthnicCereal}_{is} \times \text{Round}_{99-00} + \delta_{its_o} + \delta_{idq} + \epsilon_{idt}$ .

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  $Status_{r dt}$  which measures the status of religious group  $r$  in district  $d$  at time  $t$ :

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

As in the conflict regressions, we control for price and income effects, as well as for local supply and demand conditions via  $\delta_{idt}$  and  $\delta_{r dt}$  fixed effects. 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  $\delta_{irst}$  that deal with temporal shocks to adherence at the religion-ethnicity level, and the panel fixed effects  $\delta_{irsdq}$  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.<sup>42</sup> 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.

**Status<sub>r dt</sub><sup>national-occ(r)</sup>**: 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 vari-

---

<sup>42</sup>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”.



ation in local occupational returns:

$$\text{Status}_{rdt}^{\text{national-occ}(r)} = \sum_o \log w_{odt} \theta_{od^{-}rt},$$

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.<sup>43</sup>  $\theta_{od^{-}rt}$  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  $\delta_{irst}$ , 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** $_{rdt}^{\text{national-w}(o)}$ : 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:

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

where  $w_{od^{-}t}$  are national occupational returns leaving out own district, and  $\theta_{odrt_o}$  are occupation shares by religion-district-quarter in the initial survey round  $t_0$ . Coupled with the panel fixed effects  $\delta_{irsdq}$ , 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  $\text{Status}_{rdt}^{\text{national-occ}(r)}$  measure for the three fixed-effect specifications (with cross-sectional fixed effects in column 2 most appropriate), and columns 4-6 use  $\text{Status}_{rdt}^{\text{national-w}(o)}$  (with panel fixed effects in column 6 most appropriate).<sup>44</sup>

Consistent with our conceptual framework, for all six columns,  $\alpha_3$  is positive and significant:

<sup>43</sup>Absent reliable wage data, we use the total per capita expenditure of households with primary occupation  $o$  deflated by the all-India CPI.

<sup>44</sup>To ease interpretation of the taboo main effect in columns 1 and 4, we demean both status variables.

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	0.196*** (0.00379)			0.213*** (0.00383)		
taboo=1 $\times$ status <sub>rdt</sub> <sup>national_occ(r)</sup>	0.120*** (0.0128)	0.0715*** (0.0120)	0.0439*** (0.0102)			
taboo=1 $\times$ status <sub>rdt</sub> <sup>national_w(o)</sup>				0.490*** (0.0204)	0.0669*** (0.0197)	0.0556*** (0.0213)
Observations	1,111,072	1,110,724	1,109,544	1,089,132	1,088,876	1,088,280
Adjusted $R^2$	0.560	0.584	0.595	0.561	0.583	0.593
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

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. In terms of magnitudes, we focus on the more 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 1.1 percentage points relative to non taboo goods.

A closely related approach measures Status<sub>rdt</sub> in equation (5) by the average per capita expenditure of your religious group in your district, and then instruments status with one of the two shift-share measures described above. As average group expenditures are strongly correlated with both instruments—i.e. we have a strong first stage—we draw very similar conclusions from this alternative approach (see Appendix Tables D.1 and D.2).

#### 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, 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). A second implication of our framework is also worth noting. If a group has a strong taboo against consuming a particular good, then identifying with the group tends to make demand for that good less sensitive to prices, as the consumer could be at a corner solution.

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:

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

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  $\text{sign}(\alpha_3) \neq \text{sign}(\alpha_5)$ .

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, as outlined above, 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). 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,

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 $_i=1$	0.187*** (0.00359)	0.170*** (0.0185)	0.228*** (0.0158)						
ln $p_i$	0.0127*** (0.00270)	0.0427*** (0.00363)	0.0323*** (0.00488)	0.00884*** (0.00257)	0.0453*** (0.00511)	0.0424*** (0.00564)	0.00721*** (0.00255)	0.0207*** (0.00413)	0.0211*** (0.00527)
sum ln $p_j$	-0.00294* (0.00167)	-0.0140*** (0.00221)	0.00659* (0.00367)	-0.00166 (0.00165)	-0.00883*** (0.00286)	-0.00503 (0.00396)	-0.00112 (0.00165)	-0.00644*** (0.00199)	-0.00351 (0.00377)
taboo $_i=1$ x ln $p_i$		-0.0382*** (0.00278)	-0.0287*** (0.00454)		-0.0457*** (0.00488)	-0.0424*** (0.00556)		-0.0168*** (0.00351)	-0.0173*** (0.00509)
taboo $_i=1$ x sum ln $p_j$		0.0146*** (0.00171)	-0.0293*** (0.00159)		0.00946*** (0.00282)	-0.0117*** (0.00416)		0.00681*** (0.00131)	0.00209 (0.00267)
sum (ln $p_j$ x taboo $_j$ )			-0.0618*** (0.00422)			-0.0142** (0.00626)			-0.00791 (0.00574)
taboo $_i=1$ x sum (ln $p_j$ x taboo $_j$ )			0.0905*** (0.00171)			0.0340*** (0.00664)			0.0101* (0.00575)
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 $R^2$	0.560	0.560	0.568	0.585	0.585	0.585	0.595	0.595	0.595
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
region*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
region*state*product*round*quarter	No	No	No	Yes	Yes	Yes	No	No	No
region*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 region 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

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 indeed find that  $\alpha_4 < 0$  and  $\text{sign}\alpha_3 \neq \text{sign}\alpha_5$  in these 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.

Most importantly, 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,  $\alpha_7$  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. In terms of magnitudes, abstention from taboo good  $i$  rises by between 0.1 (column 9) and 0.9 percentage points (column 3) with a ten percent increase in the price of other taboo goods (relative to the effect on non-taboo goods).

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 any reverse-causality would likely bias us against finding our key result,  $\alpha_7 > 0$ . Suppose the proportion of households that identify with their religion falls due to another shock (e.g. status or salience changes). This would lead households to abstain less from their religious taboo goods and prices should rise as we move up the supply curve.<sup>45</sup> Thus we would observe price increases for one taboo good associated with *lower* abstention from other taboo goods, not higher 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.

---

<sup>45</sup>As long, of course, as these consumption changes dominate any movements in the opposite direction by members of other religions. Indeed, as noted above, conflict-induced abstention does lower prices (Appendix Table B.3.9).

## 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 or 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. We must, however, be careful in interpreting the relative importance of the three factors since the total effect of changing a particular factor may come in part through changing another one. For example, Mitra and Ray (2014) argue that changes in group status can affect inter-group violence. Similarly, political moves that increase the price of beef may also increase religious salience. Finally, 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,  $\bar{x}_r$  is the prototypical Muslim bundle, and  $\bar{x}_s$  is the prototypical Gujarati bundle.

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

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

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  $\kappa_{hJ}$ . Finally,  $\xi_{hJ}$  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  $\bar{x}_J$ :

$$v(p, m_h; \bar{x}_J) = \frac{\ln m_h - \ln a_J(p)}{\prod_i p_i^{\beta_i}}, \quad (8)$$

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

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

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  $V_{hs} \geq V_{hr}$ .<sup>46</sup> Therefore, defining  $\widetilde{x}_{iJ} \equiv \bar{x}_{iJ} - \beta_i \sum_i \bar{x}_{iJ} \ln p_i$ , the observed budget share is:

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

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

$$V_{hr} - V_{hs} = -\delta_1 \left( \frac{\sum_i (\bar{x}_{ir} - \bar{x}_{is}) \ln p_i}{\prod_i p_i^{\beta_i}} \right) + \delta_2 (y_r - y_s) + \delta_3 (\kappa_{hr} - \kappa_{hs}) + (\xi_{hr} - \xi_{hs}). \quad (11)$$

Religious identity is more appealing when its prototypical bundle is relatively cheap (i.e.  $\sum_i \bar{x}_{ir} \ln p_i$  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  $\xi_{hJ}$  terms are i.i.d. extreme value (type I). The probability of choosing religious identity  $r$  is then:

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

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.<sup>47</sup> We obtain a non-linear

<sup>46</sup>To break ties we assume the household identifies ethnically in case of indifference.

<sup>47</sup>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.

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  $-\delta_1 \frac{\sum_i (\bar{x}_{ir} - \bar{x}_{is}) \ln p_i}{\prod_i p_i^{\beta_i}} + \delta_2 (y_r - y_s) + \delta_3 (\kappa_{hr} - \kappa_{hs})$ . This motivates running the linear specification:

$$x_{hi} = (\bar{x}_{ir} - \bar{x}_{is}) \left( -\eta_1 \sum_i (\bar{x}_{ir} - \bar{x}_{is}) \ln p_i + \eta_2 (y_r - y_s) + \eta_3 (\kappa_{hr} - \kappa_{hs}) \right) \quad (13)$$

$$+ \bar{x}_{is} + \sum_k \gamma_{ik} \ln p_k + \beta_i (\ln m_h - \sum_i \bar{x}_{is} \ln p_i) + FES + \varepsilon_{hi},$$

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.<sup>48</sup> As discussed earlier in this section, we proxy the  $\bar{x}_{iJ}$  by the prototypical bundles we observe in these populations in the data (excluding own village).<sup>49</sup> 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 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 eth-

<sup>48</sup>The one term absent from equation (13) is the budget share changes that result from different income effect magnitudes under the two identities,  $-\beta_i \sum_i (\bar{x}_{ir} - \bar{x}_{is}) \ln p_i Prob(V_{hr} > V_{hs})$ . We assume that these differences are negligible, or at least that  $cov[\beta_i \sum_i (\bar{x}_{ir} - \bar{x}_{is}), (\bar{x}_{ir} - \bar{x}_{is})] \approx 0$  so that our coefficients of interest are unbiased. Similarly, as we do not include the denominator  $\prod_i p_i^{\beta_i}$  in the cost effect,  $\eta_1$  should be interpreted as approximately equal to  $\delta_1 E[1/\prod_i p_i^{\beta_i}]$ . Of course, if preferences are homothetic, neither of these issues arises. We apply the approximation  $\ln a_s(p) \equiv a_0 + \sum_i \bar{x}_{is} \ln p_i + \frac{1}{2} \sum_i \sum_k \gamma_{ik} \ln p_i \ln p_k \approx \sum_i \bar{x}_{is} \ln p_i$ , 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  $\bar{x}_{is}$  as we note below.

<sup>49</sup>Specifically, we assume that  $\bar{x}_J$  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  $h$ 's village. The prototypical consumption of the religious group is computed at the national level, while the prototypical ethnic bundle is simply the average bundle among the relevant ethnicity (either the state or linguistic group). Although relatively limited in India (see Atkin (2016)), inter-state migration may bias our ethnic bundle estimates. To assess the likely size of the bias, Appendix Table I.1 reproduces our cross section analysis for the 43rd round—the only round containing state of birth information matched to consumption. Estimated coefficients are very similar when excluding interstate migrants.



nic 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  $\sum_i (\bar{x}_{ir} - \bar{x}_{is}) \ln p_i$ , we use the cost of the two identity bundles leaving out the cost of good  $i$  itself,  $\sum_{j \neq i} (\frac{\bar{x}_{jr}}{\sum_{j \neq i} \bar{x}_{jr}} \ln p_j - \frac{\bar{x}_{js}}{\sum_{j \neq i} \bar{x}_{js}} \ln p_j)$ , 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, columns 1-3, present the results using our primary ethnicity measure, the state of residence of the household.<sup>50</sup> As before, all columns control for local supply and demand conditions via  $\delta_{idt}$  and  $\delta_{rdt}$  fixed effects. Column 2 includes cross-sectional fixed effects  $\delta_{irst}$  that deal with temporal shocks to adherence at the good-religion-ethnicity level. Column 3 includes panel fixed effects  $\delta_{irsdq}$  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 5 percent level or lower.<sup>51</sup> 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 I.3 relaxes the symmetry implicit in equation (13)—that a shock moves the household as far away from one identity’s prescribed behavior as it does towards the other identity’s prescribed behavior. More precisely,

<sup>50</sup>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 I.2 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.

<sup>51</sup>The linear approximation appears to be reasonable despite 79 percent of budget shares being zero across our 124 goods. Three quarters of predicted values lie between 0 and 1, and the inadmissible values are all negative and typically tiny (a median of -0.0003).

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

	LHS Variable: Share Spent on Good $i$					
	Ethnicity as State			Ethnicity as Language		
	(1) Baseline	(2) Cross-section	(3) Panel	(4) Baseline	(5) Cross-section	(6) Panel
$(\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	-0.0346 (0.0535)	-0.696*** (0.0967)	-0.624*** (0.0868)	-0.249*** (0.0583)	-0.915*** (0.112)	-0.801*** (0.0989)
$(\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	0.479*** (0.0264)	0.246*** (0.0265)	0.234*** (0.0371)	0.480*** (0.0289)	0.226*** (0.0311)	0.178*** (0.0432)
$(\bar{x}_{ir} - \bar{x}_{is}) \times conflict_r$	0.549*** (0.0419)	0.0818** (0.0335)	0.110** (0.0435)	0.519*** (0.0384)	0.0439 (0.0305)	0.0700* (0.0393)
Observations	33,763,092	33,755,156	33,736,680	33,763,092	33,754,908	33,736,680
Adjusted $R^2$	0.767	0.773	0.778	0.767	0.773	0.778
log prices and total expenditure	Yes	Yes	Yes	Yes	Yes	Yes
prod*dist*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
relig*dist*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
relig*state*prod*round*quarter	No	Yes	No	No	No	No
relig*state*prod*dist*quarter	No	No	Yes	No	No	No
relig*lang*prod*round*quarter	No	No	No	No	Yes	No
relig*lang*prod*dist*quarter	No	No	No	No	No	Yes

Notes: Dependent variable is the share spent on good  $i$  in total food expenditure.  $\bar{x}_{ir} - \bar{x}_{is}$  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.  $conflict_r$  is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. Columns 4-6 define ethnicity as the set of districts where the most-commonly-spoken mother tongue is the same based on the 2001 Census of India. Prototypical ethnic bundles and the cost of these bundles are recalculated accordingly. 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

we evaluate our assumption that households make a choice between their religious and ethnic identity by separately interacting  $\bar{x}_{ir}$  and  $\bar{x}_{is}$  with each of the three shocks, and examining whether the coefficients on the  $\bar{x}_{ir}$  and  $\bar{x}_{is}$  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  $\bar{x}_{ir}$  interactions of opposite signs to the coefficients on the  $\bar{x}_{is}$  interactions and broadly similar in magnitude. Appendix Table I.4 allows the impacts of each of the three shocks to vary by the religion of the household. The heterogeneity is rather limited, but upper-caste Hindus appear to be somewhat more sensitive to status and Muslims are potentially more sensitive to conflict.

Columns 4-6 of Table 7 present the results using our linguistic definition of ethnicity—the set of districts in which the most-commonly-spoken mother tongue is the same. This implies new prototypical ethnic bundles, as well as new local bundle costs. The results are qualitatively similar although there is some change in relative magnitudes (the estimated cost effects are somewhat stronger using this alternative measure, and the conflict effects somewhat weaker).

In summary, households move closer to their religious bundle and further away from their

ethnic bundle: as their ethnic bundle becomes relatively more expensive; as the relative status of their religious group increases; and following religious conflict involving their religious group.

## 6 Counterfactuals and Implications

In this section, we explore the implications of our demand-system estimates for changes in identity, voting for religious and ethnic parties, 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 politics, health, and welfare.

### 6.1 Changes in Identity, 1987-2000

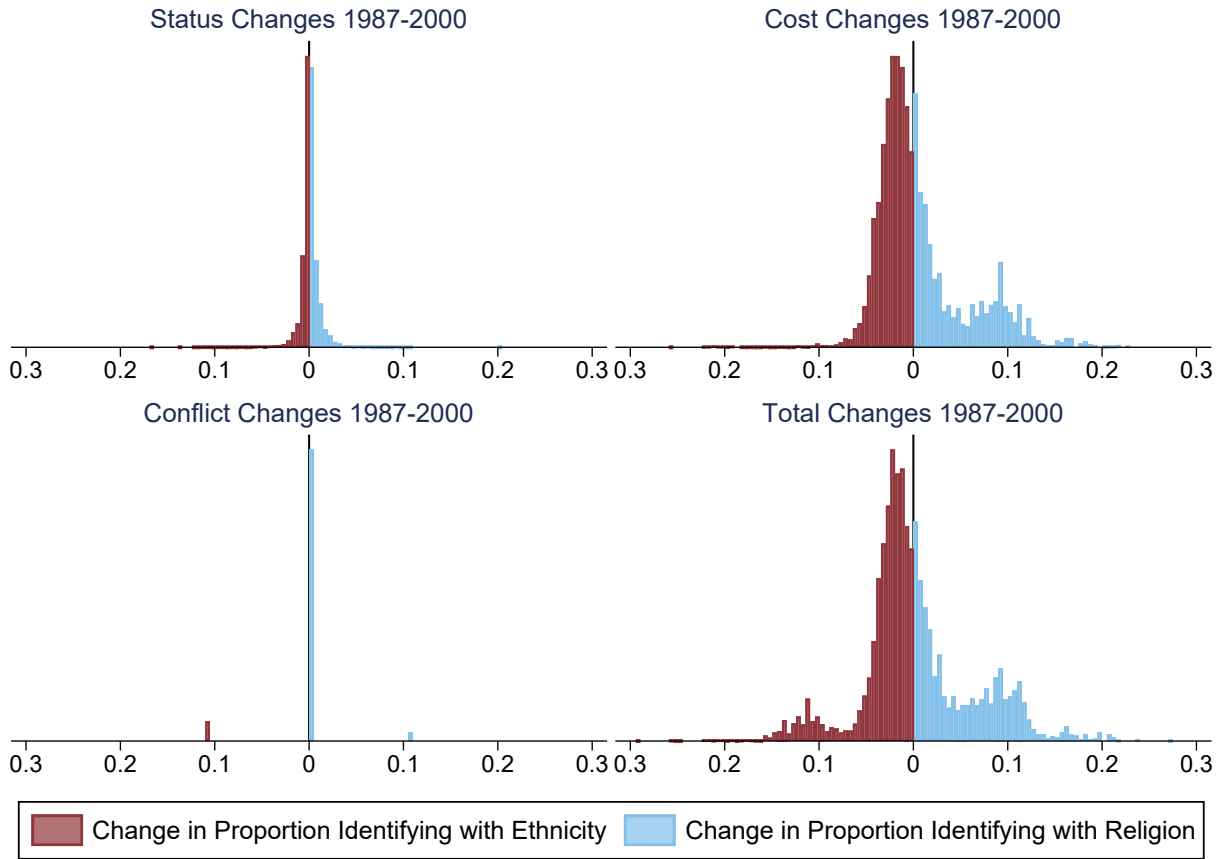
We first explore the changes in identity induced by changes in cost, status and salience. Drawing on the local linear approximation in equation (13), the change in the proportion of households in a religion-ethnicity-district cell that identifies with their religion,  $dP_r$ , approximately equals

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

Thus we can use the  $\eta$  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. Note that such an exercise abstracts from potential general equilibrium effects on prices. As discussed in Section 4.4,  $\eta_1$  is likely biased towards zero due to positive shocks to a particular identity raising the prices of that group’s bundle. Relatedly, the long changes in costs we feed into the counterfactual are probably smaller than the underlying costs changes absent this GE feedback. Given these forces, our counterfactuals likely underestimate the magnitude of the costs channel.

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. 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).

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



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), more than ten percent of households switched identity. Second, there is substantial heterogeneity across district-religion cells. Over this period there were more cells where households were shifting from their religious to ethnic identity than vice versa. But on average there was a net movement of 0.3 percent of the population in the other direction—from ethnic to religious identity—as cells moving in this direction had more households shifting identity. Appendix Figure J.1 breaks out these changes by religious group, aggregating over all districts using population weights. Interestingly, while the general movement among upper-caste Hindus in this period was towards a religious identity, Muslims tended to move towards their ethnic identity.

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. 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 six to nine 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.<sup>52</sup> Furthermore, we are only identifying the local effect of conflict. Thus, our estimates do not pick up the possibility that, for example, Hindu-Muslim conflict in northern India affects identification choices in the south.

The fact that the economic costs due to price changes have such large impacts is striking given that this channel has been largely overlooked in both public discussion and scholarly work. Indeed, it offers a new perspective on recent and much-discussed efforts by the Hindu-nationalist BJP party to raise the effective price of beef through bans and legislation.<sup>53</sup>

## 6.2 Voting and Identity Change, 1987-2000

We now explore whether these identity changes, derived from household consumption decisions, are associated with an alternative proxy for identity choices—changes in voting behavior. A word of caution: voting is driven by many factors other than identity, and many people vote for parties that are not running on an ethnic or religious agenda. Furthermore, identity shifters, such as changes in prices and occupational returns, may affect voting via non-identity channels. Nonetheless, since identity is often invoked as an important factor determining voter behavior, especially given the rise of the Hindu nationalist vote in recent decades, a positive association serves as a useful check on our revealed preference approach to uncovering identity choices.

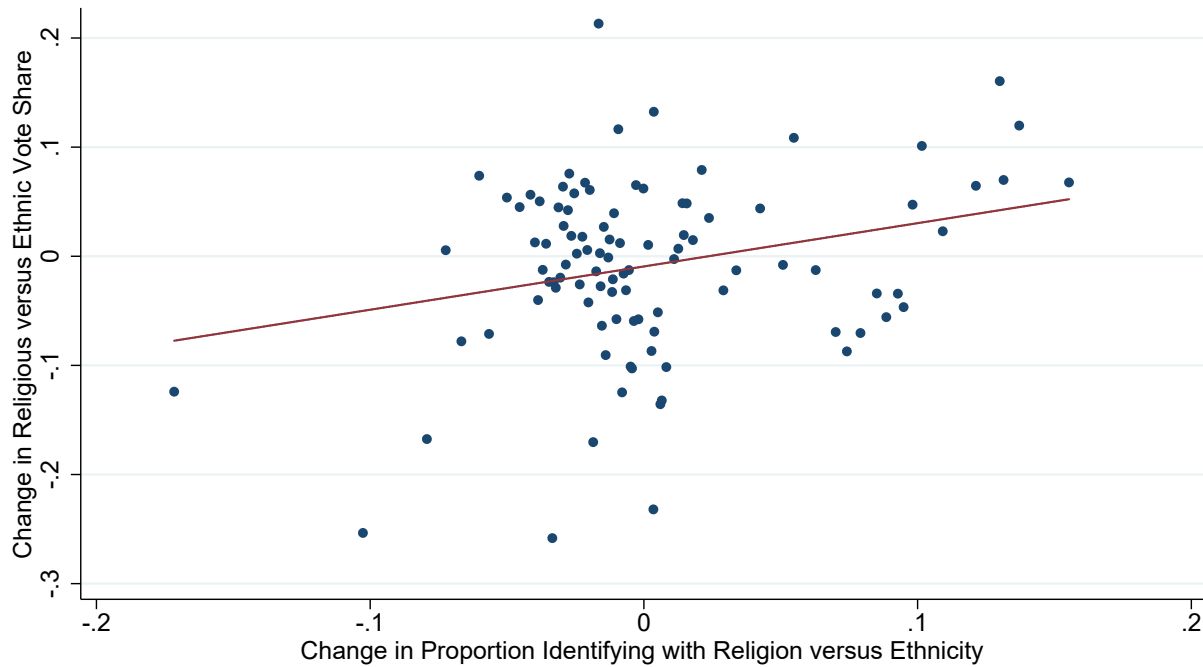
Our data on vote shares for religious versus ethnic parties comes from the Indian State Assembly Election and Candidates Database (Jensenius and Verniers 2017). We classify parties as religious or ethnic based on both the party manifesto and media reports on the party's political philosophy and traditional support base. For ethnic parties we also use a second classification, obtained from Election Commission reports which classify certain parties as "State Parties". Of a total of 457 parties listed in the database for the relevant years, we classify 6 as religious parties and 30 as ethnic parties, and the Election Commission classifies 49 as State Parties (29 of which overlap with our ethnic party classification). Note that many parties are neither ethnic nor religious according to this classification, including the Indian National Congress Party, which accounts for around one third of the total state assembly votes in our sample period. To compute

---

<sup>52</sup>Note that a similar conclusion holds if we just focus on the three states—Gujarat, Maharashtra and Uttar Pradesh—where most of the conflicts occurred (see Appendix Figure J.2).

<sup>53</sup>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.

Figure 4: Recovered Identity Changes and Changes in Vote Shares, 1987-2000



Notes: Binned scatter-plot of changes in religious versus ethnic vote shares on changes in proportion of households in a district-religion cell that identify with their religion. Linear regression line shown (slope = 0.397, s.e. = 0.102).

vote shares for a given identity at the district-identity-round level, we aggregate constituencies up to the district level and use the state assembly election closest to each NSS round.<sup>54</sup> Appendix K provides further details on our classifications and Appendix Table K.3 provides descriptive statistics for vote shares by identity group and round.

Figure 4 shows a binned scatter-plot of vote shares on identity changes. On the x-axis we plot the net change between 1987-2000 in the proportion of households in a district-religion cell that identify with their religion, calculated from our consumption data in Section 6.1 above. On the y-axis we plot the change in the difference between the religious and ethnic vote shares in that cell over the same time period. As we saw in Figure 3, many district-religion cells had small net changes in household identification. However, in those cells where a significant share of households switch to identify with their religion ( $dP_{rd} > 0$ ), the vote share of the religious parties relative to ethnic parties also tends to increase. Similarly, the ethnic vote share tends to increase relative to the religious in cells where more households identify ethnically ( $dP_{rd} < 0$ ).

Table 8 explores this relationship in more detail. Column 1 regresses the change in the religious versus ethnic vote share between 1987-2000 on the estimated change in the proportion

<sup>54</sup>State Assembly elections are organized every five years and are not held at the same time across India.

Table 8: Identity Changes and Vote Shares, 1987-2000

	LHS Variable: Change in Vote Shares, 1987-2000				
	Religious vs. Ethnic		Religious	Ethnic	State
	(1)	(2)	(3)	(4)	(5)
change in proportion religious <sub><i>r,d</i></sub>	0.397*** (0.102)		0.113* (0.0581)		
change in cost <sub><i>r</i></sub> - cost <sub><i>s</i></sub>		-0.145* (0.0748)			
change in status <sub><i>r</i></sub> - status <sub><i>s</i></sub>		-0.203 (0.175)			
change in conflict <sub><i>r</i></sub>		0.0931*** (0.0150)			
change in proportion religious <sub><i>d</i></sub>				-0.321** (0.127)	-0.566*** (0.176)
Observations	729	729	795	334	358
Adjusted $R^2$	0.019	0.056	0.004	0.016	0.025

*Notes:* Dependent variable in columns 1-2 is the double difference in vote share for religious parties versus ethnic parties between 1987 and 2000 by district-religion. Dependent variable in column 3 is the difference in vote share for religious parties by district-religion over the period. Dependent variable in columns 4-5 is the difference in vote share for ethnic parties (two definitions) by district over the period. Vote shares are computed for State Assembly elections closest to the 1987-1988 and 1999-2000 NSS rounds. See text for a detailed description of the party categorization. Change in proportion religious is the recovered change in proportion identifying with religion versus ethnicity over the period as revealed by food consumption. Differences in costs, status and conflict are long differences between the 1987-1988 and 1999-2000 NSS rounds within district-religion-ethnicity cells. Cost is the mean cost in the district of the entire prototypical bundle. Status is the difference between religious and ethnic status measured by national returns to the initial local occupational mix of religion and ethnicity. Regressions weighted by survey population weights. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

religious from Section 6.1. A 10 pp increase in the share of households who identify religiously rather than ethnically is associated with a 4 pp increase in the vote share of religious versus ethnic parties in that district-religion cell (significant at the 1 percent level).

Column 2 explores the ‘reduced-form’ relationship between vote shares and the three identity shifters that we have been exploring. Cells in which the prototypical religious bundle became more expensive relative to the prototypical ethnic bundle exhibit a reduction in the religious vote share relative to the ethnic vote share. Consistent with Iyer and Shrivastava (2018), increased religious conflict is associated with an increase in the vote share of religious versus ethnic parties. The association with status changes is statistically insignificant.

The last three columns show the association between recovered identity changes and the change in the religious vote share (column 3), and with the change in the ethnic vote share measured either using our classification (column 4) or the classification of “State Parties” in the Election Commission reports (column 5).<sup>55</sup> In all three specifications, the correlation is in the expected direction and is particularly strong for the vote share of ethnic/state parties.

Taken together, the association between our revealed preference measures and voting-based

<sup>55</sup>Notice that for the ethnic vote share regressions, there is now only one observation per district. Hence the number of observations drops and “change in proportion religious” refers to all religions in the district.

measures corroborates the view that identity changes are important for explaining shifts in voting behavior. More generally, food consumption patterns might be useful for studying other identity-driven behaviors like occupational choice or female labor force participation. Additionally, this exercise provides some reassurance that our measures are uncovering changes in identity choices, particularly since voting is done in private and so is less likely to merely reflect social pressure, social connections and signaling to others.

### 6.3 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.<sup>56</sup> The change in caloric intake per capita can be calculated for each religion-district cell by taking the change in identity choices and multiplying by the difference in the caloric intake of households under their two possible identities:

$$dCalories_{ih} \approx calories\_per\_kg_i \times \frac{foodexp_h}{p_{ih}} (\bar{x}_{ir} - \bar{x}_{is}) dP_r.$$

The left panel of Figure 5 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 5 percent 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).

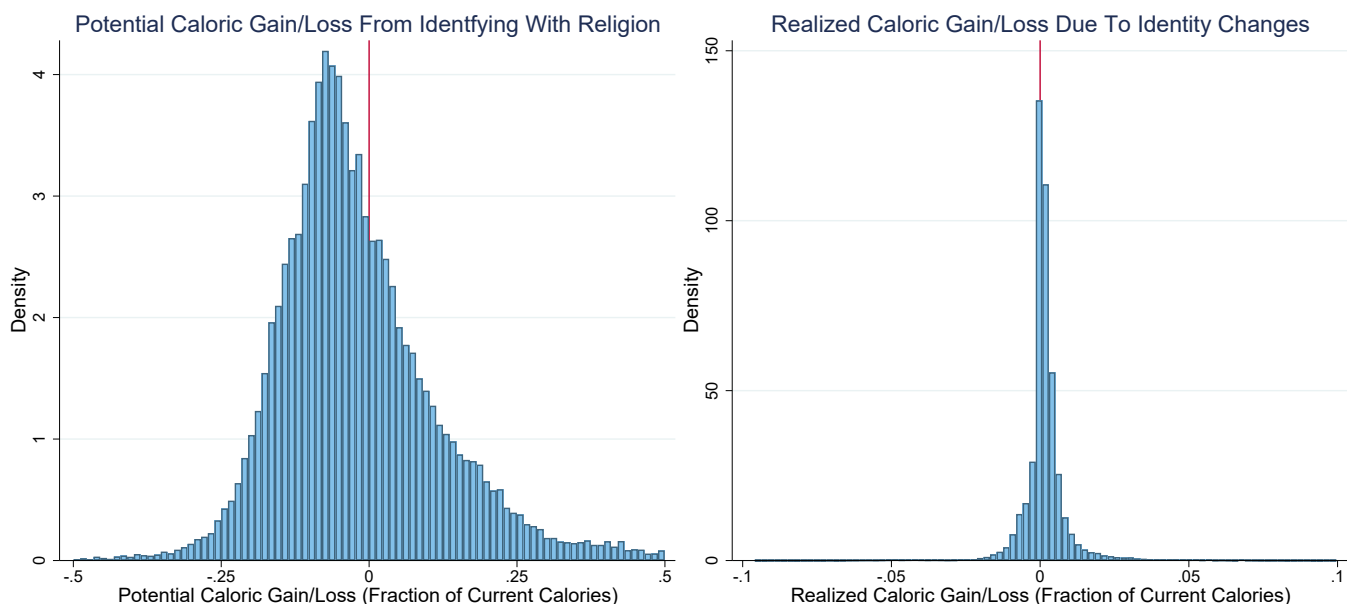
The right panel of Figure 5 plots the distribution of realized caloric gains and losses due to identity changes over this period. As with the choices themselves, the distribution is reasonably symmetric. On average, caloric consumption declined by 1.4 percent due to identity shifts. Note that if this average comes from a binary identity choice, only a fraction of households are switching identity but those few that are lose or gain 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, 1.4 percent is equal to the average decline in caloric intake per capita between 1987 and 2000 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

---

<sup>56</sup>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 5: Caloric Gains from Identity Changes, 1987-2000



identity changes may have played a role.

#### 6.4 Effects of Identity Changes on Welfare 1987-2000

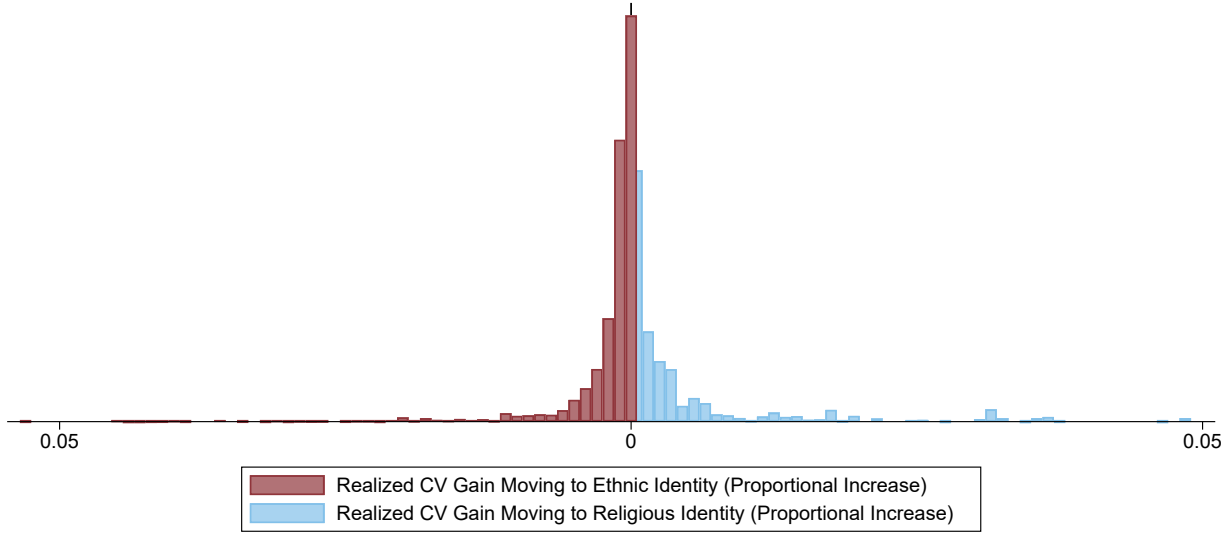
Finally, if we take our conceptual framework literally, we can say something about the additional welfare gains made possible by the fungibility of identity. In our framework, households are more flexible than consumer theory typically assumes, as they have an additional means to respond to shocks to their environment. By switching identity, households are essentially able to change the set of parameters that enter their (indirect) utility function  $V_{hJ}(p, m_h, y_J, \kappa_{hJ}; \bar{x}_J)$ , 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 do so, first 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  $J$ :

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

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

Figure 6: Realized Compensating Variation Gains from Identity Changes, 1987-2000



identity compared to switching identity is:

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

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  $\eta$ s from the linear approximation above.<sup>57</sup> 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 due to the fungibility of identity. Figure 6 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 0.03 and an increase of 0.01 for cells moving towards their religious identity. As above, these cell-level averages are small as only a fraction of households switch. Appendix Figure J.4 shows numbers many 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.

<sup>57</sup>Specifically,  $\ln z_{JJ} - \ln z_{JJ'} \approx \sum_i (\bar{x}_{iJ} - \bar{x}_{iJ'}) \ln p_{i1} - \Pi_i p_{i1}^{\beta_i} E\left[\frac{1}{\Pi_i p_i^{\beta_i}}\right] \left(\frac{\eta_2}{\eta_1} (y_{J1} - y_{J'1}) + \frac{\eta_3}{\eta_1} (\kappa_{J1} - \kappa_{J'1})\right)$ .

## 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.

More broadly, our revealed preference approach draws on consumption survey data that are widely available for many countries and time periods. It can therefore be fruitfully used in other contexts to provide a deeper understanding of both identity choice and its ramifications not only for voting but also for other identity-driven behaviors like schooling, occupational choice and female labor force participation.

## 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.
- Aghion, P., R. Burgess, S. J. Redding, and F. Zilibotti, "The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India," *American Economic Review*, September 2008, 98 (4), 1397–1412.
- 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.
- Bankert, Alexa, Leonie Huddy, and Martin Rosema, "Measuring partisanship as a social identity in multi-party systems," *Political behavior*, 2017, 39 (1), 103–132.
- 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, C. and E. Colson-Sihra, "The Conspicuous Consumption of the Poor: Forgoing Calories for Aspirational Goods," *Working Paper*, SSRN 3270814, 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.
- , —, and J. Pan, "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 2015, 130 (2), 571–614.
- 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.
- Bursztyl, 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.
- Charness, Gary and Yan Chen, "Social Identity, Group Behavior and Teams," *Annual Review of Economics*, 2020, 12.
- 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, “Building Nations Through Shared Experiences: Evidence from African Football,” *Working Paper*, 2019.
- 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.
- Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, “Global evidence on economic preferences,” *The Quarterly Journal of Economics*, 2018, 133 (4), 1645–1692.
- 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,” *Review of Economic Studies*, Forthcoming.
- 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.
- Haslam, S Alexander and Naomi Ellemers, “Social identity in industrial and organizational psychology: Concepts, controversies and contributions,” *International Review of Industrial and Organizational Psychology*, 2005, 20 (1), 39–118.
- 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.

- Horowitz, Donald, *Ethnic groups in conflict*, University of California Press, 1985.
- Iyer, Sriya and Anand Shrivastava, "Religious riots and electoral politics in India," *Journal of Development Economics*, 2018, 131, 104–122.
- Jensenius, Francesca R and Gilles Verniers, "Studying Indian politics with large-scale data: Indian election data 1961–today," *Studies in Indian Politics*, 2017, 5 (2), 269–275.
- 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.
- Klor, Esteban F and Moses Shayo, "Social identity and preferences over redistribution," *Journal of Public Economics*, 2010, 94 (3-4), 269–278.
- 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.
- Martin, L. A., S. Nataraj, and A. E. Harrison, "In with the Big, Out with the Small: Removing Small-Scale Reservations in India," *American Economic Review*, February 2017, 107 (2), 354–86.
- 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.
- Sen, Colleen Taylor, *Feasts and fasts: A history of food in India*, Reaktion Books, 2014.
- 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.
- , "Social Identity and Economic Policy," *Annual Review of Economics*, 2020, 12.
- 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.
- Somanathan, Rohini, "The Measurement and Mismeasurement of Social Difference," *Centre for Development Economics*, 2018, *Working Paper No. 294*.
- 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.
- Topalova, P., "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India," *AEJ: Applied*, October 2010, 2 (4), 1–41.
- 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.
- Varshney, Ashutosh, "Ethnic conflict and civil society: India and beyond," *World politics*, 2001, 53 (3), 362–398.
- 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.
- Young, H Peyton, "The evolution of social norms," *Annual Review of Economics*, 2015, 7(1), 359–387.

# How Do We Choose Our Identity?

## A Revealed Preference Approach Using Food Consumption

David Atkin, Eve Colson-Sihra and Moses Shayo

### Online Appendices

#### Contents

<b>A Data</b>	<b>2</b>
<b>B Religious Conflict</b>	<b>8</b>
B.1 Conditional Event Study . . . . .	8
B.2 Event Study: Other Tests . . . . .	10
B.3 Religious Conflict and Taboo Adherence Regressions . . . . .	12
<b>C State Splits</b>	<b>22</b>
<b>D Status Shocks</b>	<b>26</b>
<b>E Cost of Identity</b>	<b>31</b>
<b>F Regressions with Household Controls</b>	<b>32</b>
<b>G Regressions without NSS Round 50</b>	<b>36</b>
<b>H Baseline Taboo Regressions with Ovo-Pesco Vegetarianism</b>	<b>40</b>
<b>I Linear Approximation of Identity Choice</b>	<b>43</b>
<b>J Counterfactuals</b>	<b>48</b>
<b>K Ethnic and Religious Parties</b>	<b>51</b>



# A Data

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

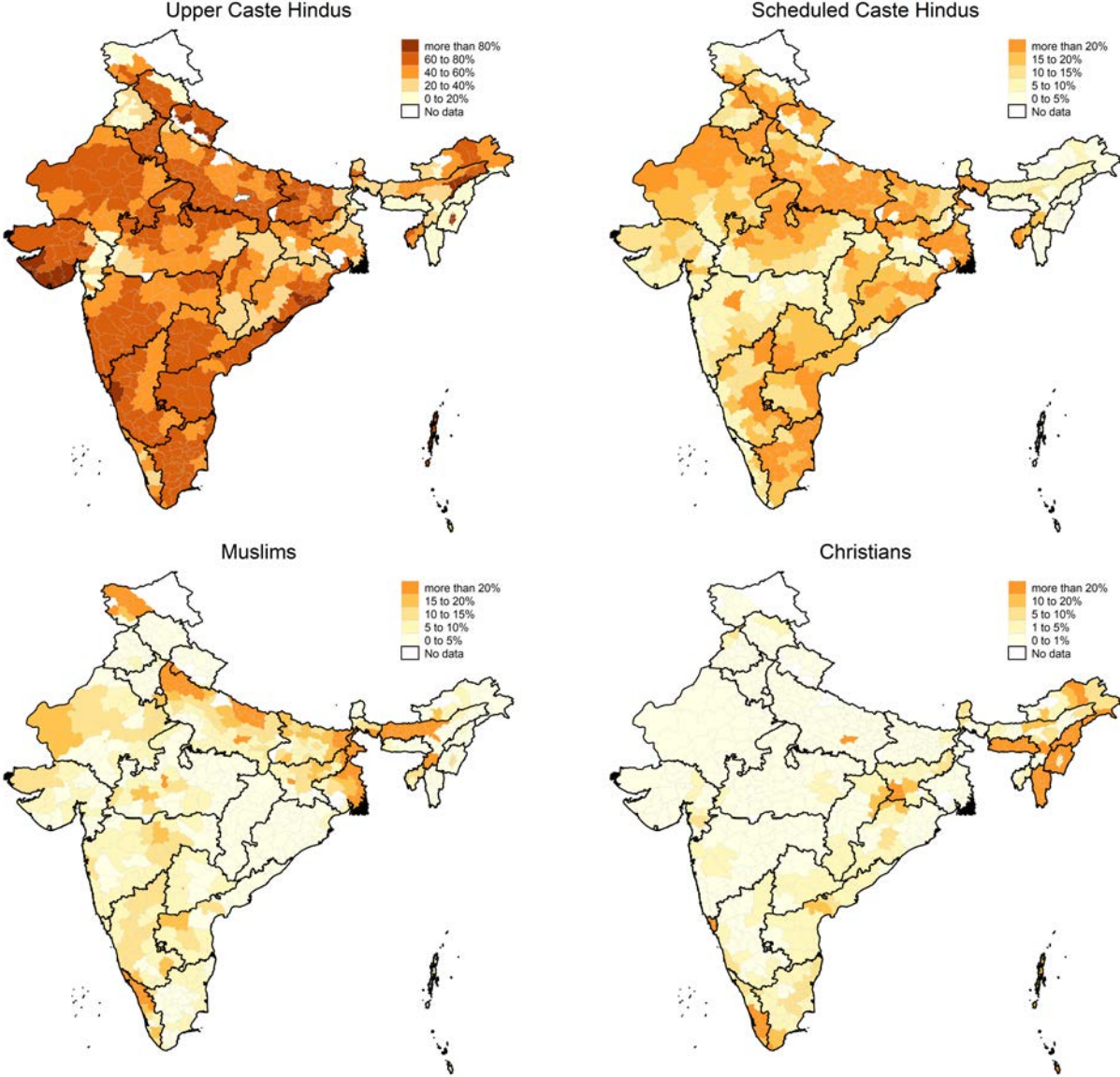
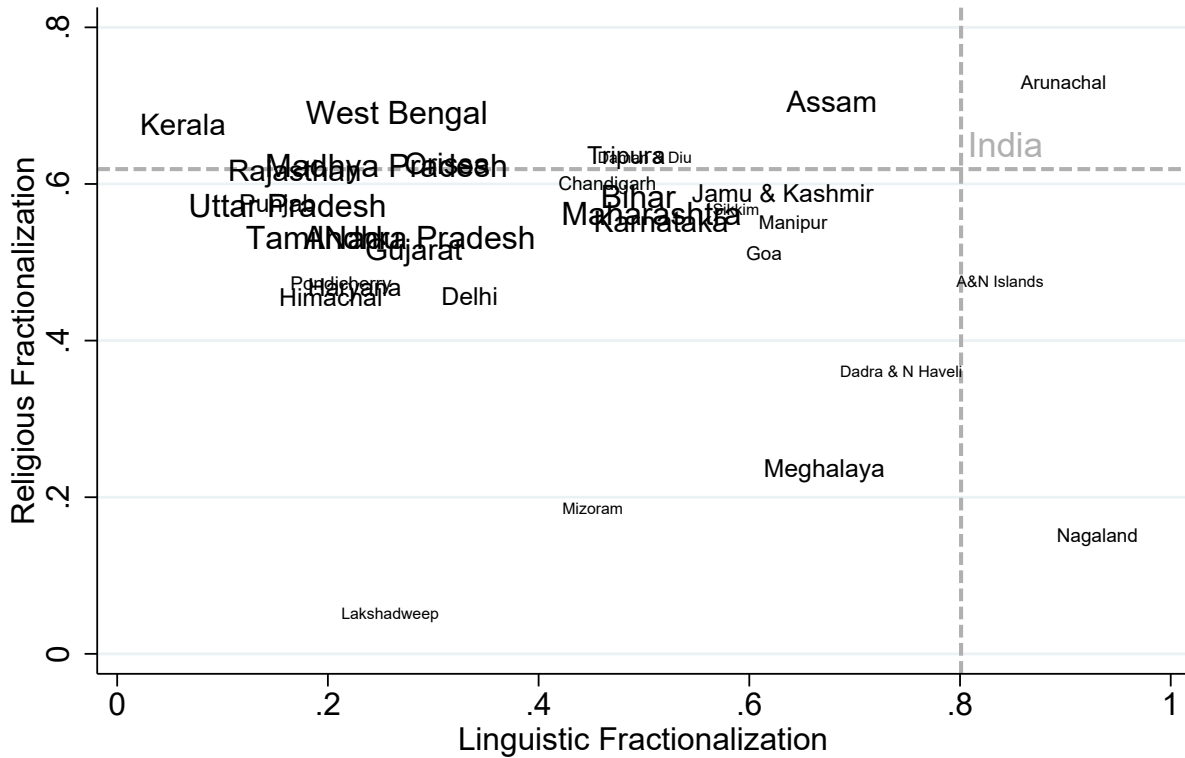


Figure A.2: Linguistic and Religious Fractionalizations by State



Notes: Religious fractionalization based on all religions, NSS 55.  
 Linguistic fractionalization based on mother tongue, 2001 Census.  
 Size of State label proportional to State population.

Figure A.3: Share of Rice and Wheat in Total Cereal Expenditures by District, all NSS Rounds

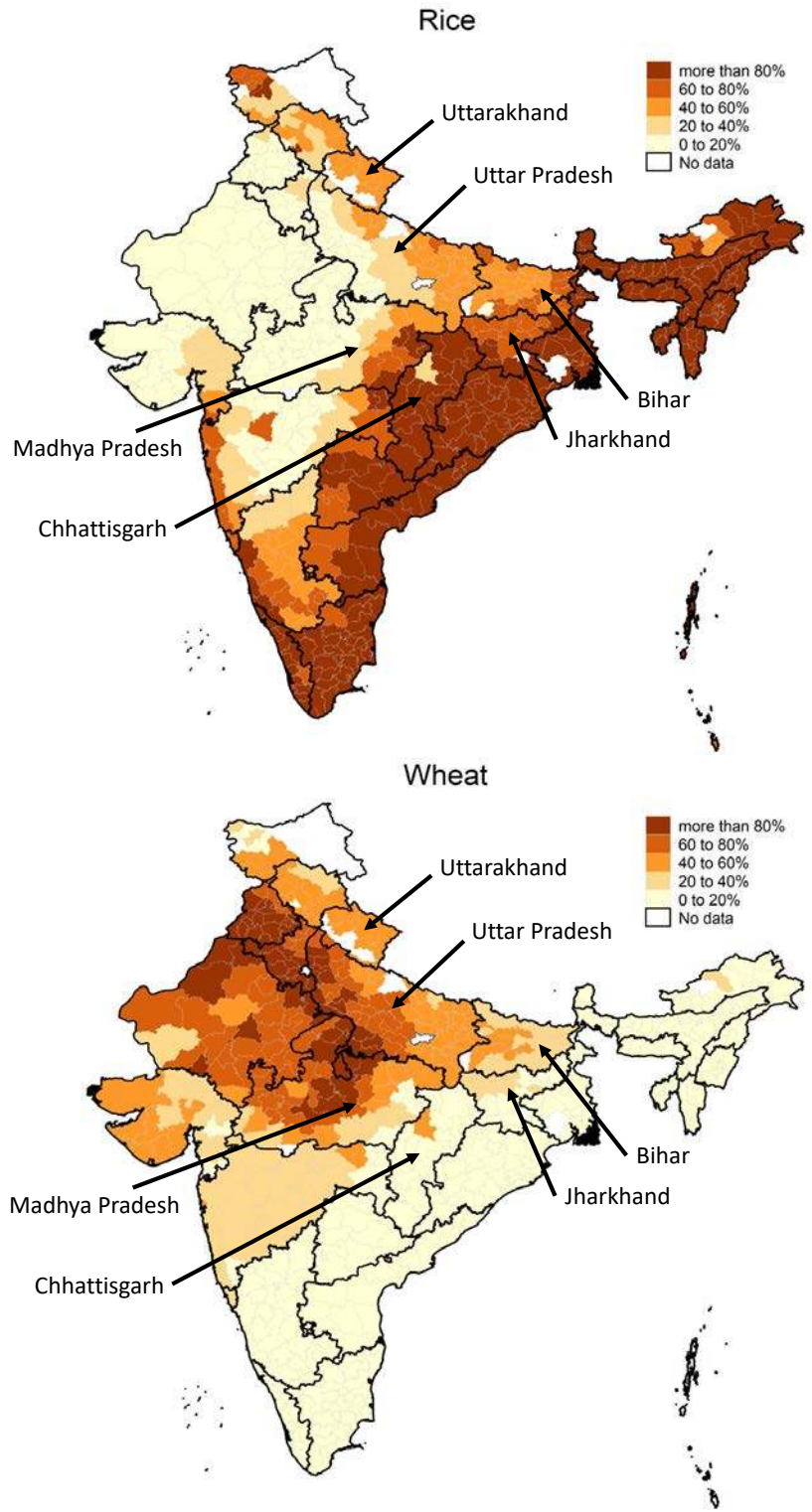


Table A.1: List of Food Items by NSS Categories

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: Causes of Hindu-Muslim Conflict as Reported in the Varshney-Wilkinson Dataset During Period of NSS Rounds 43, 50 and 55

<b>Category</b>	<b>Reported Cause</b>	<b>Percent of Conflicts</b>	
<b>Animal Slaughter</b>	<b>Total Animal Slaughter</b>	<b>1.4%</b>	
<b>Politics and Economics</b>	Economic Interest (Land)	0.7%	
	Economic Interest (Other)	0.3%	
	Political (Agitation)	2.5%	
	Political (Other)	1.0%	
	Political Elections	1.0%	
	Political Rivalry	1.7%	
	<b>Total Politics and Economics</b>	<b>7.2%</b>	
<b>Religion</b>	Desecration of Religious Site/Symbol	23.7%	
	Public Rituals\Festivities	16.3%	
	Religious Pamphlets	0.3%	
	<b>Total Religion</b>	<b>40.3%</b>	
<b>Previous Violence</b>	Previous Violence (Communal)	9.8%	
	Previous Violence (Other)	1.4%	
	Previous Violence (Terrorism, War)	1.4%	
	<b>Total Previous Violence</b>	<b>12.6%</b>	
<b>Other Reported Cause</b>	Allegations of Unpatriotic Acts	1.0%	
	Construction/Attempted Construction	0.3%	
	Cricket Match	1.0%	
	Crime	0.3%	
	Death of Muslim Protestors in Police Firing	0.7%	
	Family Rivalry	0.3%	
	Homicide	2.0%	
	Illegal Attack (Miscellaneous)	0.7%	
	Militant Gun Down by Army	0.3%	
	Petty Issue	1.0%	
	Private Quarrel	3.4%	
	Protest Against Police Actions	1.0%	
	Protest against Urdu News Telecast	0.3%	
	Rape	0.3%	
	Sexual Harrasment	2.5%	
	Student Clashes over T.V. Channel Surfing	0.3%	
	<b>Total Other Reported Cause</b>	<b>15.4%</b>	
	<b>Cause Insufficiently Reported</b>	<b>Total Cause Insufficiently Reported</b>	<b>23.1%</b>

*Notes:* Table reports causes of Hindu-Muslim conflict for the period six months before, during and six months after each round in the Varshney-Wilkinson Dataset.

Table A.3: Hindu-Muslim Conflict by State and NSS Round

State	1987-88		1993-94		1999-2000	
	Incidence	No. Killed	Incidence	No. Killed	Incidence	No. Killed
Andhra Pradesh	0	0	5	24	2	1
Arunachal Pradesh	0	0	0	0	0	0
Assam	0	0	5	0	0	0
Bihar	3	18	10	44	4	19
Delhi	9	14	3	41	2	1
Goa	0	0	0	0	0	0
Gujarat	49	131	16	316	9	14
Haryana	0	0	1	4	0	0
Himachal Pradesh	0	0	0	0	0	0
Jammu and Kashmir	9	7	0	0	3	9
Karnataka	10	23	20	57	2	6
Kerala	1	0	0	0	2	7
Madhya Pradesh	4	2	4	130	0	0
Maharashtra	28	51	15	796	13	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	1	0	1	0
Punjab	0	0	0	0	0	0
Rajasthan	7	4	3	47	2	14
Sikkim	0	0	0	0	0	0
Tamil Nadu	1	1	3	10	0	0
Tripura	0	0	0	0	0	0
Uttar Pradesh	18	181	13	210	8	13
West Bengal	4	15	3	42	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

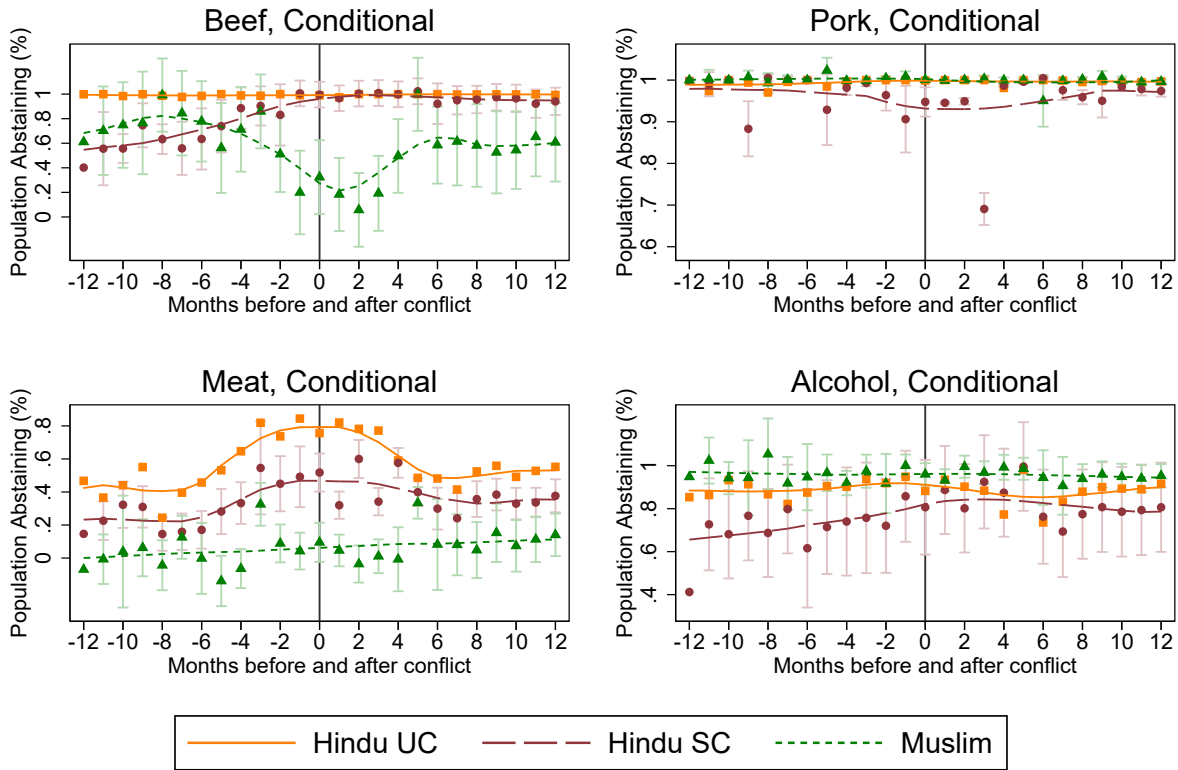
The non-parametric plots that show taboo abstention in the period building up to and after local conflict (Figure 2) do not account for potential confounds coming from price and income changes or other factors. For example, conflicts may be more likely in certain regions (those with different endowments 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:

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

where  $\text{Abstain}_{ihgm}$  is an indicator variable that takes the value 1 if a household does not consume good  $i$ ;  $\text{SC}_h$  and  $\text{Muslim}_h$  are indicators that take the value 1 if a household  $h$  is scheduled-caste Hindu or Muslim (upper-caste Hindu is the reference group);  $\text{Conflict}_{gm}$  is an indicator for being surveyed  $m$  months before or after the first Hindu/Muslim conflict in region  $g$ ;  $\ln \text{price}_{jh}$  is the village median price of good  $j$  that controls for own- and cross-price effects;  $\ln \text{realfoodexp}_h$  is the log of per capita food expenditure deflated by a Stone price index that controls for income effects; and  $\delta_{igm}$  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  $\theta_{im}^r$  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  $\theta_{im}^r$  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)

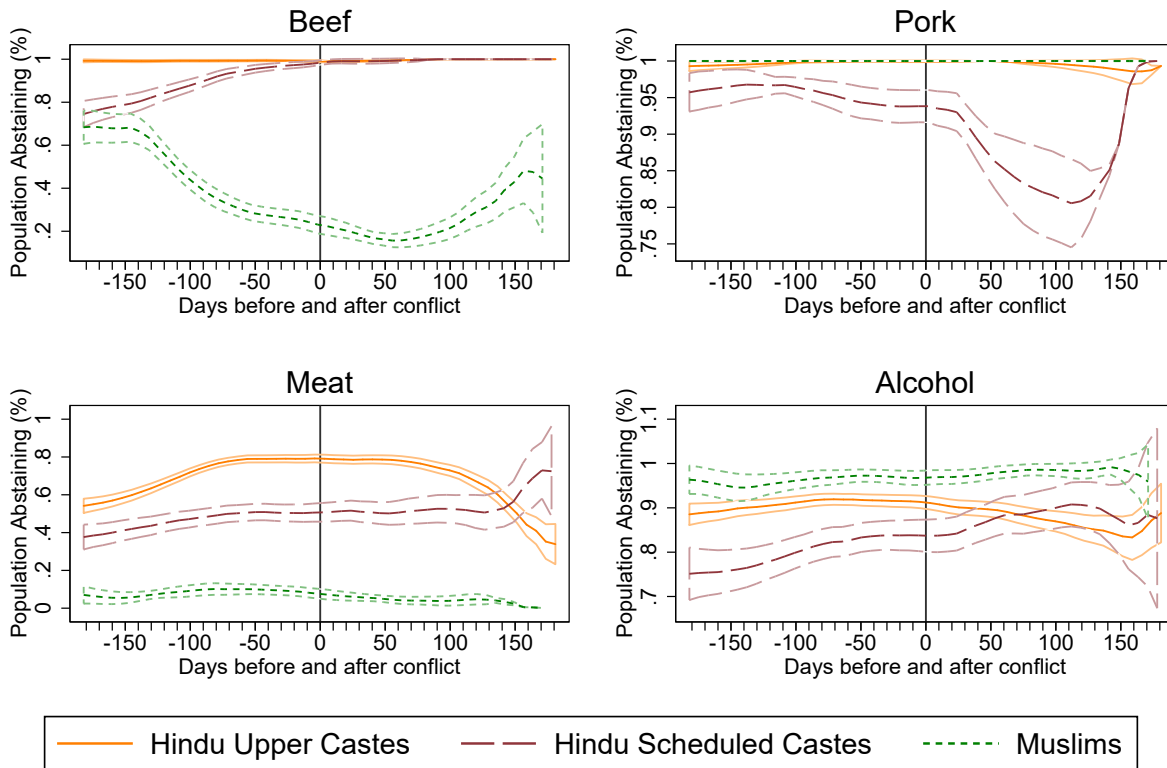


Figure B.2.2: Conflict and Chicken/Mutton Avoidance, NSS 50th Round (1993-1994)

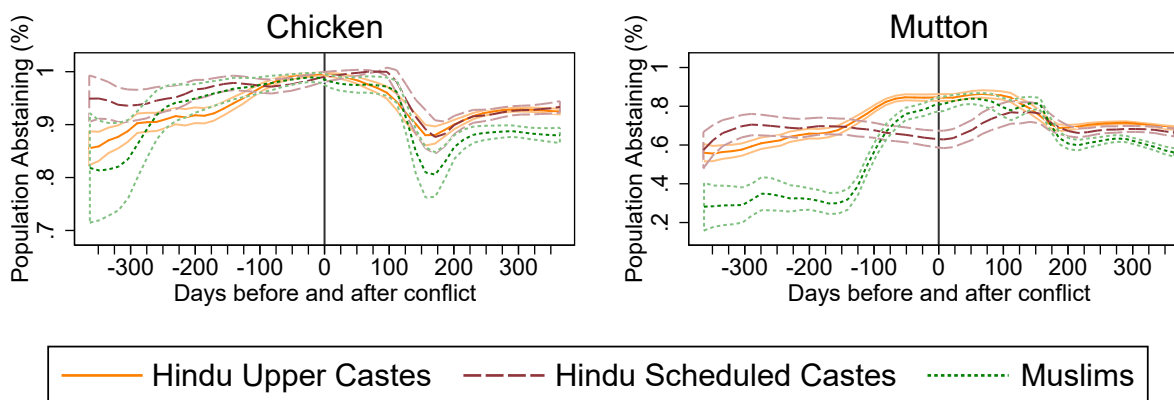
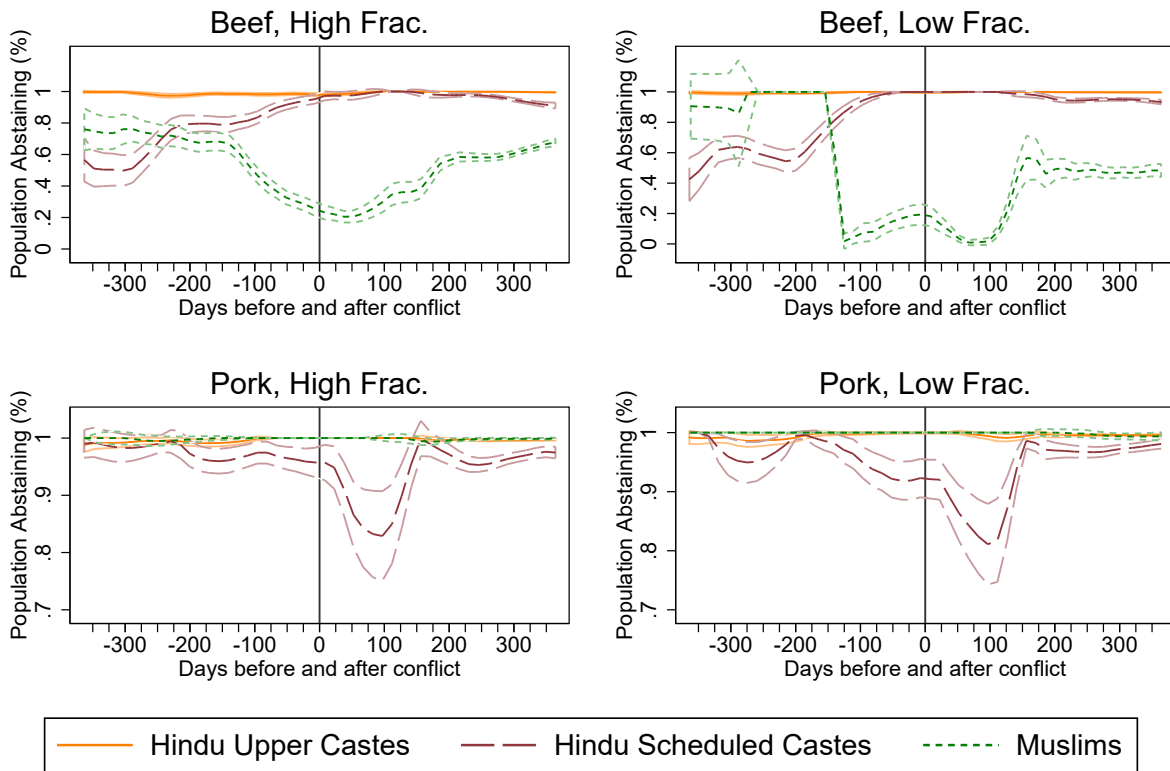


Figure B.2.3: Conflict and Beef/Pork Avoidance, Round 1993-1994, High vs. Low Local Religious Fractionalization



### B.3 Religious Conflict and Taboo Adherence Regressions

Table B.3.1: 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.3.2: 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	0.181*** (0.00727)		
taboo=1 $\times$ conflict	0.106*** (0.0169)	0.0396*** (0.0108)	0.0599*** (0.00988)
Observations	1,115,640	1,115,292	1,114,116
Adjusted $R^2$	0.560	0.585	0.596
log prices and total expenditure controls	Yes	Yes	Yes
product*district*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 is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. 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.3.3: Religious Conflict and Taboo Adherence, Clustering at Religion-District Level

	LHS Variable: Abstain from Consuming Good $i$		
	Baseline	Cross-section	Panel
	(1)	(2)	(3)
	All	All	All
taboo=1	0.181*** (0.00944)		
taboo=1 $\times$ conflict	0.106*** (0.0254)	0.0396** (0.0158)	0.0599*** (0.0119)
Observations	1,115,640	1,115,292	1,114,116
Adjusted $R^2$	0.560	0.585	0.593
log prices and total expenditure controls	Yes	Yes	Yes
product*district*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 is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. 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 in parentheses. Regressions weighted by survey population weights. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table B.3.4: Religious Conflict and Taboo Adherence, Lags and Leads

	LHS Variable: Abstain from Consuming Good $i$		
	Panel		
	(1)	(2)	(3)
taboo=1 $\times$ conflict t-0 quarter	0.0527*** (0.0172)	0.0522*** (0.0171)	0.0517*** (0.0168)
taboo=1 $\times$ conflict t-1 quarter		0.0381*** (0.0141)	0.0378*** (0.0143)
taboo=1 $\times$ conflict t-2 quarter		0.0468*** (0.0131)	0.0473*** (0.0132)
taboo=1 $\times$ conflict t-3 quarter		0.00630 (0.0131)	0.00712 (0.0129)
taboo=1 $\times$ conflict t-4 quarter		-0.00995 (0.0137)	-0.00918 (0.0138)
taboo=1 $\times$ conflict t+1 quarter			0.0209 (0.0157)
taboo=1 $\times$ conflict t+2 quarter			0.00415 (0.0170)
taboo=1 $\times$ conflict t+3 quarter			0.0237 (0.0151)
taboo=1 $\times$ conflict t+4 quarter			-0.0218 (0.0302)
Observations	1,114,116	1,114,116	1,114,116
Adjusted $R^2$	0.596	0.596	0.596
log prices and total expenditure controls	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes
religion*district*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 effects 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.3.5: Religious Conflict and Beef Abstention in NSS 50 (1993-1994)

	LHS Variable: Abstain from Consuming Beef	
	Baseline	Cross-section
	(1)	(2)
taboo=1	0.325*** (0.0123)	
taboo=1 × conflict	0.407*** (0.0439)	0.343*** (0.0516)
Observations	59,279	59,248
Adjusted $R^2$	0.378	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 is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. 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.3.6: Number of Religious Conflict Fatalities and Taboo Adherence

	LHS Variable: Abstain from Consuming Good $i$		
	Baseline	Cross-section	Panel
	(1)	(2)	(3)
taboo=1	0.185*** (0.00360)		
taboo=1 × log fatalities	0.0439*** (0.00889)	0.00782 (0.00587)	0.0118** (0.00596)
Observations	1,115,640	1,115,292	1,114,116
Adjusted $R^2$	0.560	0.585	0.596
log prices and total expenditure controls	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes
religion*district*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 quarter in which the household is surveyed or in the preceding two quarters. 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.3.7: 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	0.158*** (0.00385)	677438.0 (58582231.3)	8.035 (115479.0)
taboo=1 $\times$ conflict	0.101*** (0.0162)	0.0509*** (0.0107)	0.0597*** (0.00981)
taboo=1 $\times$ conflict, other districts in region	0.106*** (0.00900)	0.0489*** (0.00779)	-0.00581 (0.00501)
Observations	1,115,640	1,115,292	1,114,116
Adjusted $R^2$	0.561	0.585	0.596
log prices and total expenditure controls	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes
religion*district*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 is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. 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.3.8: Religious Conflict and Taboo Adherence, Interaction with Local Abstention

	LHS Variable: Abstain from Consuming Good $i$					
	(1)	(2)	(3)	(4)	(5)	(6)
taboo $\times$ conflict	0.0393*** (0.0106)	0.0442*** (0.0101)	0.0380*** (0.0111)	0.0428*** (0.0115)	0.0334*** (0.0127)	0.0415*** (0.0132)
taboo $\times$ conflict $\times$ $\mathbb{1}[\text{abstention}_{irst} > 0.99]$	-0.0440** (0.0178)					
taboo $\times$ conflict $\times$ $\mathbb{1}[\text{abstention}_{irsdq} > 0.99]$		-0.0119 (0.0151)				
taboo $\times$ conflict $\times$ $\mathbb{1}[\text{abstention}_{irst} > 0.95]$			-0.0414*** (0.0119)			
taboo $\times$ conflict $\times$ $\mathbb{1}[\text{abstention}_{irsdq} > 0.95]$				-0.0185 (0.0143)		
taboo $\times$ conflict $\times$ $\mathbb{1}[\text{abstention}_{irst} > 0.90]$					-0.0310** (0.0139)	
taboo $\times$ conflict $\times$ $\mathbb{1}[\text{abstention}_{irsdq} > 0.90]$						-0.0267* (0.0154)
Observations	1,115,292	1,114,116	1,115,292	1,114,116	1,115,292	1,114,116
Adjusted $R^2$	0.585	0.596	0.585	0.596	0.585	0.596
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	Yes	No	Yes	No	Yes	No
religion*state*product*district*quarter	No	Yes	No	Yes	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 is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters.  $\mathbb{1}[\text{abstention}_{irst}]$  and  $\mathbb{1}[\text{abstention}_{irsdq}]$  are indicators equal to 1 if the local abstention level for good  $i$  is above 0.99 (columns 1-2), 0.95 (columns 3-4) or 0.90 (columns 5-6). They measure cross-sectional (good-religion-state-time) and panel (good-religion-state-district-quarter) abstention levels respectively. All columns include the main effects and interactions of taboo, conflict and local abstention. All columns include the baseline fixed effects, columns 1, 3 and 5 add the fixed effects for cross-sectional identification and columns 2, 4 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 B.3.9: Demand-Side Effects of Conflict on Prices

	LHS Variable: log price by good-district-time							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2SLS	2SLS	2SLS	2SLS	RF	RF	RF	RF
fraction abstaining <sub>idt</sub>	-2.989*** (0.696)	-0.0427 (0.153)	-0.412 (0.431)	-0.757** (0.353)	-2.517*** (0.572)	-0.0704 (0.253)	-0.343 (0.360)	-0.958** (0.447)
conflict				0.0681 (0.0470)				0.146** (0.0590)
Observations	12,368	13,185	12,368	12,368	12,368	13,185	12,368	12,368
Adjusted $R^2$	0.237	0.220	0.525	0.528	0.254	0.219	0.520	0.520
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)	447.9	1408.9	431.7	684.0				
First-stage F-statistic (RKF)	157.1	350.9	154.9	311.6				

Notes: Dependent variable is the log price at the good-district-quarter-round level. Fraction abstaining<sub>idt</sub> 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  $\widehat{Abstain}_{iht} = \widehat{\alpha}_1 \widehat{Taboo}_{ir} + \widehat{\alpha}_2 \widehat{Conflict}_{rdt} + \widehat{\alpha}_3 \widehat{Taboo}_{ir} \times \widehat{Conflict}_{rdt}$  using the estimated  $\widehat{\alpha}$ 's from the baseline regression, and then compute  $\widehat{ShareAbstain}_{idt}$  as the weighted mean of  $\widehat{Abstain}_{iht}$  by product-district-quarter-round. Columns 4 and 8 also include conflict as an independent variable (dummy for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters.). 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table B.3.10: Religious Conflict and Taboo Adherence, without Conflicts Caused by Animal Slaughtering, Political/Economic or Religious Factors

	LHS Variable: Abstain from Consuming Good $i$									
	Baseline		Without Conflicts Caused by				Slaughter/PolEcon			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Cross	Panel	Cross	Panel	Cross	Panel	Cross	Panel	Cross	Panel
taboo=1 × conflict	0.0396*** (0.0107)	0.0599*** (0.00981)	0.0387*** (0.0108)	0.0606*** (0.00983)	0.0401*** (0.0111)	0.0636*** (0.00995)	0.0392*** (0.0112)	0.0660*** (0.0101)	0.0439** (0.0176)	0.0536*** (0.0175)
Observations	1,115,292	1,114,116	1,115,292	1,114,116	1,115,292	1,114,116	1,115,292	1,114,116	1,114,116	1,114,116
Adjusted $R^2$	0.585	0.596	0.585	0.596	0.585	0.596	0.585	0.596	0.585	0.596
log prices and total expenditure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
prod*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*state*prod*round*quarter	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
relig*state*prod*district*quarter	No	Yes	No	Yes	No	Yes	No	Yes	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 is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. In columns 3-4, conflicts caused by animal slaughtering are coded as zero. In columns 5-6, conflicts caused by political or economic events are coded as zero. In columns 7-8, both conflicts caused by animal slaughtering or political/economic events are coded as zero. In columns 9-10, conflicts caused by religious factors, animal slaughtering or political/economic events are coded as zero. All columns include the baseline fixed effects, columns 1, 3, 5, 7 and 9 add the fixed effects for cross-sectional identification and columns 2, 4, 6, 8 and 10 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.3.11: Religious Conflict and Taboo Adherence by Cause of Conflict

LHS Variable: Abstain from Consuming Good  $i$

	Conflicts Caused by					
	Baseline (1)	Religion (2)	Previous Conflict (3)	Other Causes (4)	No Reported Cause (5)	Slaughter/PolEcon (6)
taboo=1 × conflict	0.0599*** (0.00981)	0.0594*** (0.0113)	0.0485 (0.0314)	0.0972** (0.0392)	0.0364* (0.0194)	-0.0288 (0.0267)
Observations	1,114,116	1,114,116	1,114,116	1,114,116	1,114,116	1,114,116
Adjusted $R^2$	0.596	0.596	0.596	0.596	0.596	0.596
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	No	No	No	No	No
religion*state*product*district*quarter	Yes	Yes	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 in the quarter in which the household is surveyed or in the preceding two quarters. Column 1 is the baseline-panel specification. Conflict is restricted to Hindu-Muslim conflict caused by religious factors in column 2, previous violence in column 3, other miscellaneous causes in column 4, unreported cause in column 5, and animal slaughter, political or economic causes in column 6. All columns include the baseline fixed effects and the fixed effects for panel identification. Robust standard errors clustered at religion-district-round-quarter in parentheses. Regressions weighted by survey population weights. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

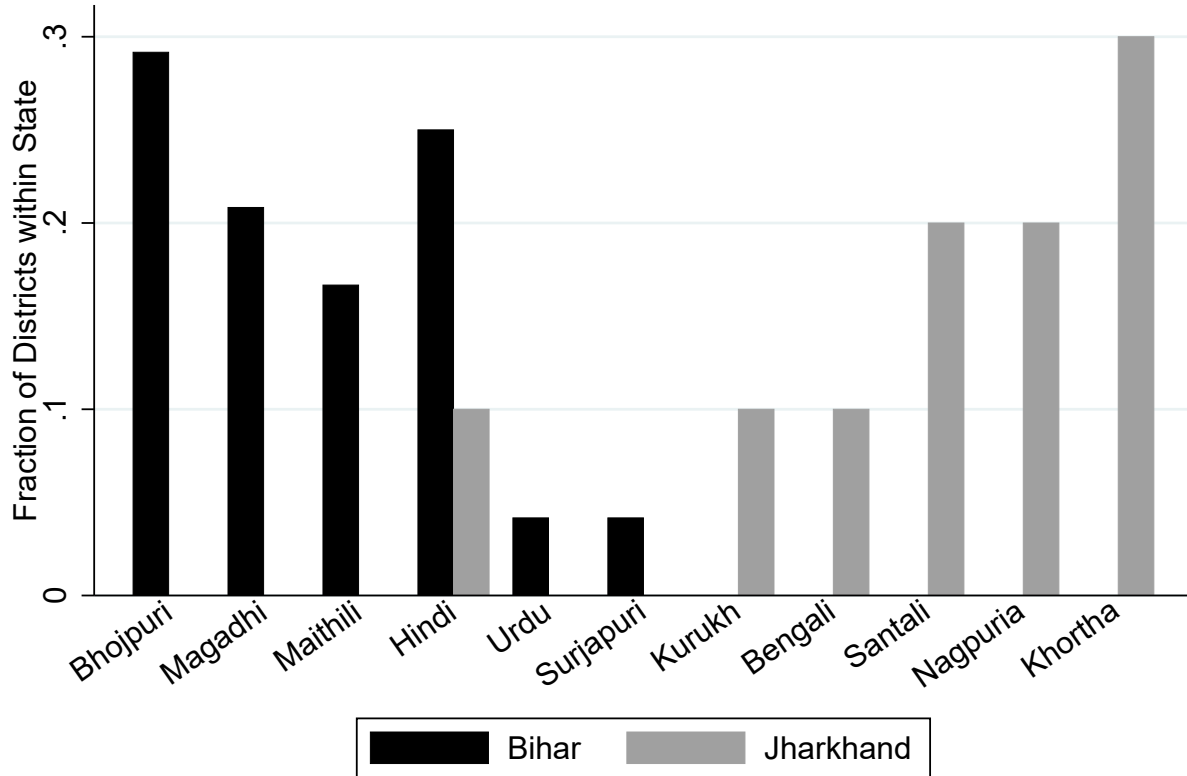
Table B.3.12: Religious Conflict and Taboo Adherence in Survey Rounds 43 and 50, Excluding Households Serving Meals to Guests

	LHS Variable: Abstain from Consuming Good $i$					
	All			W/o Households Serving Meals to Guests		
	Baseline	Cross-section	Panel	Baseline	Cross-section	Panel
	(1)	(2)	(3)	(4)	(5)	(6)
taboo=1	0.176*** (0.00460)			0.168*** (0.00473)		
taboo=1 $\times$ conflict	0.122*** (0.0184)	0.0605*** (0.0145)	0.0483*** (0.0125)	0.132*** (0.0189)	0.0652*** (0.0162)	0.0571*** (0.0145)
Observations	689,004	688,744	687,404	495,696	495,452	494,064
Adjusted $R^2$	0.557	0.582	0.596	0.565	0.589	0.603
log prices and total expend	Yes	Yes	Yes	Yes	Yes	Yes
prod*district*round*quart	Yes	Yes	Yes	Yes	Yes	Yes
relig*district*round*quart	Yes	Yes	Yes	Yes	Yes	Yes
relig*state*prod*round*quart	No	Yes	No	No	Yes	No
relig*state*prod*district*quart	No	No	Yes	No	No	Yes

*Notes:* Dependent variable is an indicator for abstaining from good  $i$ . The sample is restricted to survey rounds 43 and 50 (1987-1988 and 1993-1994) for which households reported meals served to guests. 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 quarter in which the household is surveyed or in the preceding two quarters. Columns 4-6 exclude households reporting to have served meals to guests. 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

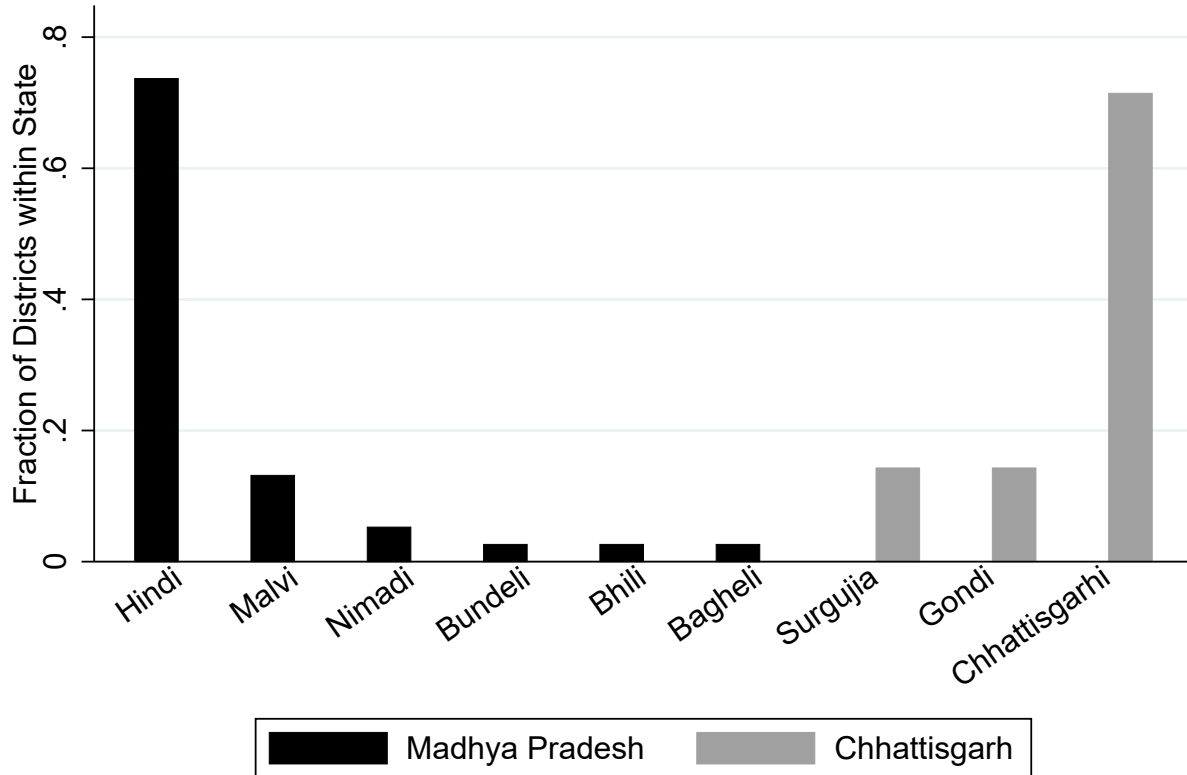
## C State Splits

Figure C.1: Dialects Spoken in Bihar and Jharkhand



Notes: For each district we compute the dialect most commonly spoken. The bars show the fraction of districts within each state by most-spoken dialect. Data on languages from 2001 Indian Census.

Figure C.2: Dialects Spoken in Madhya Pradesh and Chhattisgarh



Notes: For each district we compute the dialect most commonly spoken. The bars show the fraction of districts within each state by most-spoken dialect. Data on languages from 2001 Indian Census.

Figure C.3: Cross-District Migration and State Splits

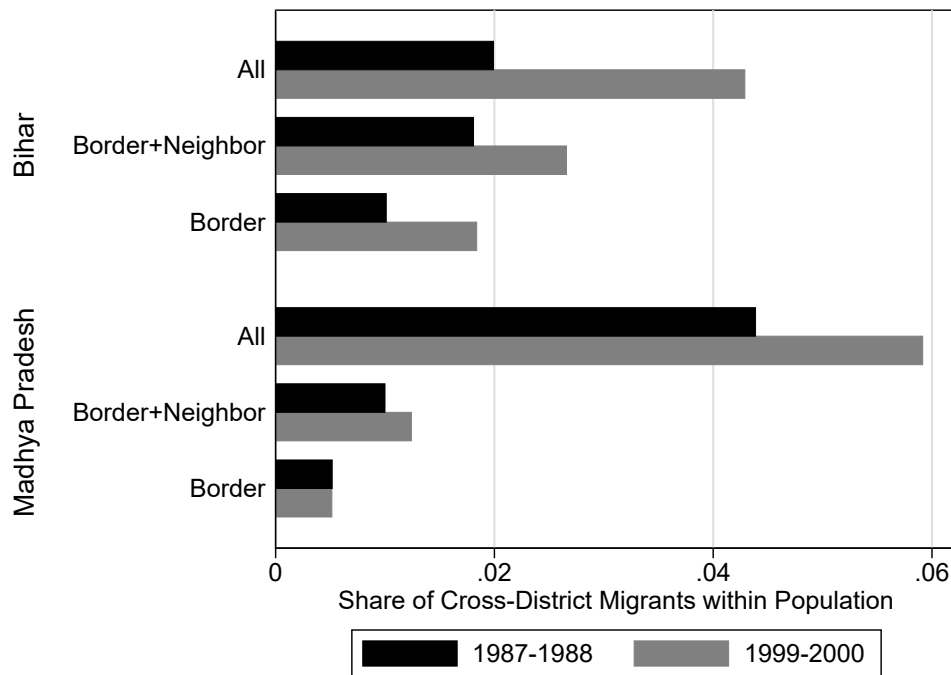


Table C.1: Ethnic Goods and State Splits

	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 $\times$ Ethnic Cereal $\times$ 1993-1994	0.0373** (0.0183)	0.0379*** (0.0138)			
Rice-loving $\times$ Ethnic Cereal $\times$ 1993-1994	0.0227** (0.0103)	0.0369*** (0.0138)			
Wheat-loving $\times$ Ethnic Cereal $\times$ 1999-2000	0.0953*** (0.0174)	0.0724*** (0.0134)	0.0845*** (0.0100)	0.0590*** (0.0129)	0.0787*** (0.0186)
Rice-loving $\times$ Ethnic Cereal $\times$ 1999-2000	0.0428*** (0.00966)	0.0797*** (0.0116)	0.0260*** (0.00866)	0.0656*** (0.0139)	0.107*** (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
original state*product*round*quarter	Yes	Yes	Yes	Yes	Yes
region*product*quarter	Yes	Yes	No	No	No
district*product*quarter	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. 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-product-time) and the fixed effects for panel identification (region- or district-product-quarter). 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 $\times$ 1987-1988	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Ethnic Cereal $\times$ 1993-1994	-0.00271 (0.00663)	-0.00230 (0.00726)			
Ethnic Cereal $\times$ 1999-2000	0.0167** (0.00654)	0.00862 (0.00592)	0.0183* (0.0106)	0.0236 (0.0151)	0.0291 (0.0214)
Observations	2,840	1,452	1,880	696	456
Adjusted $R^2$	0.676	0.702	0.765	0.774	0.768
original state*product*round*quarter	Yes	Yes	Yes	Yes	Yes
region*product*quarter	Yes	Yes	No	No	No
district*product*quarter	No	No	Yes	Yes	Yes

*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-product-time) and the fixed effects for panel identification (region- or district-product-quarter). 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	(5) Border Districts
Taboo Good x 1987–1988	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Taboo Good x 1993–1994	-0.0107 (0.0121)	-0.00137 (0.0135)			
Taboo Good x 1999–2000	-0.0130 (0.0105)	-0.0119 (0.0119)	-0.0147 (0.00948)	-0.0164 (0.0162)	-0.0479** (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
original state*product*round*quarter	Yes	Yes	Yes	Yes	Yes
region*religion*product*quarter	Yes	Yes	No	No	No
district*religion*product*quarter	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-product-time) and the fixed effects for panel identification (region- or district-religion-product-quarter for columns 3-5). Robust standard errors clustered at region-religion-round-quarter (columns 1-2) or district-religion-round-quarter (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.2: Heterogeneity in the Growth of Returns by Occupation, 1987-2000 (49 Most Common Occupations)

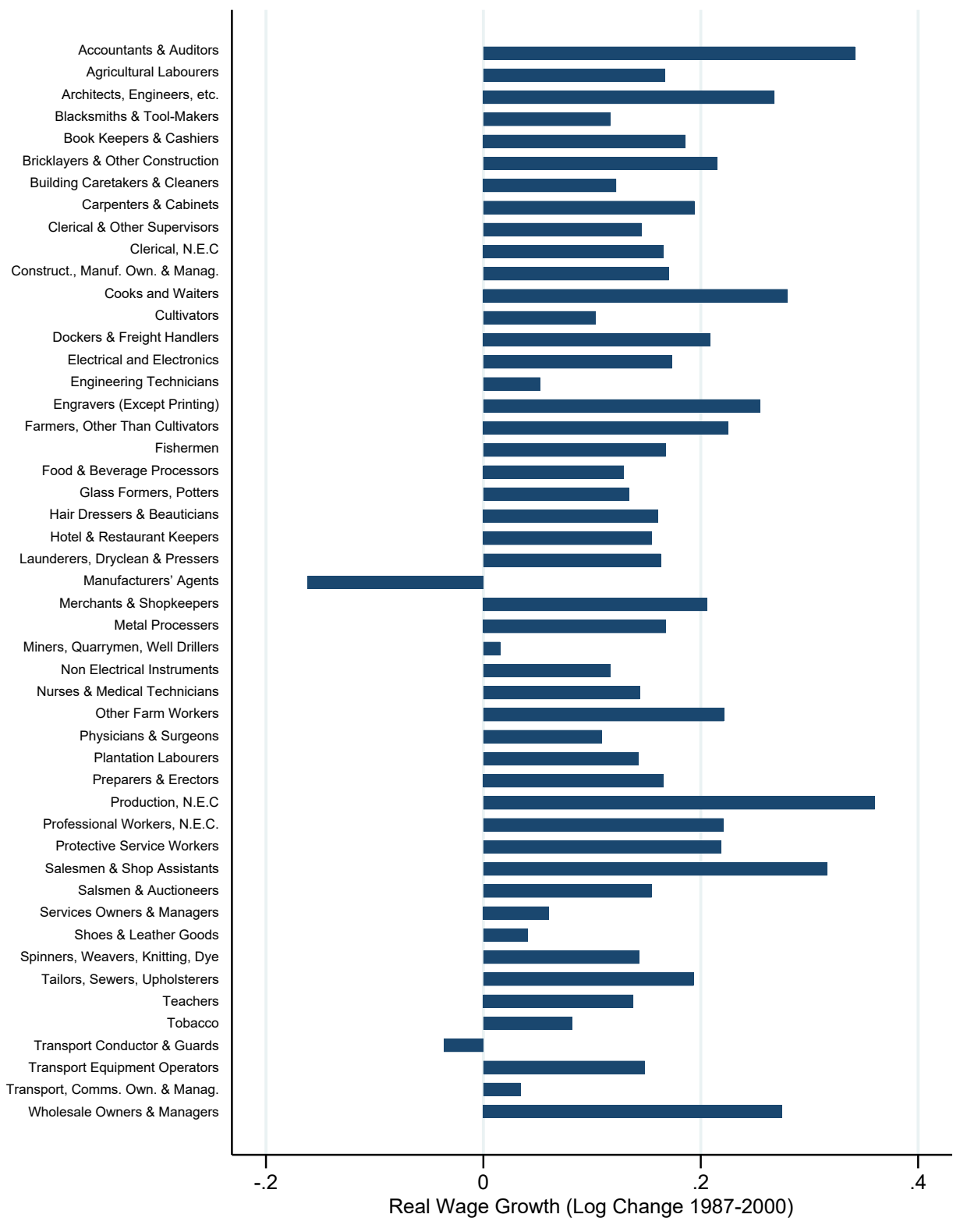


Table D.1: Status and Choice of Identity: First Stage and IV Regressions using status<sub>rdt</sub><sup>national\_occ(r)</sup> as Instrument

	First Stage			OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Baseline	Cross-section	Panel	Baseline	Cross-section	Panel	Baseline	Cross-section	Panel
status <sub>rdt</sub> <sup>national_occ(r)</sup>	1.564*** (0.0487)	1.115*** (0.0856)	1.056*** (0.0921)	0.194*** (0.00376)			0.201*** (0.00404)		
taboo=1									
taboo=1 × ln income <sub>rdt</sub>				0.0423*** (0.00600)	0.0594*** (0.00881)	0.0103*** (0.00293)	0.0910*** (0.00978)	0.0855*** (0.0145)	0.0198*** (0.00463)
Observations	12,998	12,661	11,987	1,115,640	1,115,292	1,114,116	1,111,072	1,110,724	1,109,544
Adjusted R <sup>2</sup>	0.831	0.852	0.885	0.560	0.585	0.595	0.015	-0.013	-0.023
log prices and total expenditure controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
district*round*quarter	Yes	Yes	Yes	No	No	No	No	No	No
religion*state*round*quarter	No	Yes	No	No	No	No	No	No	No
religion*state*district*quarter	No	No	Yes	No	No	No	No	No	No
product*district*round*quarter	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*round*quarter	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*district*quarter	No	No	No	No	Yes	No	No	Yes	No
First-stage F-statistic (CDF)	No	No	No	No	No	Yes	No	No	Yes
First-stage F-statistic (RKF)							620135.4 3102.8	833933.8 4037.3	669060.1 3761.0

Notes: Columns 1-3 present the first stage regressions using status<sub>rdt</sub><sup>national\_occ(r)</sup> as an instrument for religious-group-specific income. Dependent variable in columns 1-3 is the weighted mean ln income<sub>rdt</sub> (monthly per capita expenditure). Columns 4-6 are the OLS regressions of taboo and the interaction of taboo and ln income<sub>rdt</sub> on taboo abstention, and columns 7-9 are the 2SLS regressions using status<sub>rdt</sub><sup>national\_occ(r)</sup> to instrument ln income<sub>rdt</sub>. Columns 1, 4 and 7 include the baseline fixed effects, columns 2, 5 and 8 add the fixed effects for cross-sectional identification and columns 3, 6 and 9 for panel identification. Columns 1-3 adapt the sets of fixed effects to the unit of observation religion-district-time. Robust standard errors clustered at religion-district-round-quarter in parentheses. \* p < 0.10, \*\* p < 0.05, \*\*\* p < 0.01.

Table D.2: Status and Choice of Identity: First Stage and IV Regressions using status<sub>rdt</sub><sup>national\_w(o)</sup> as Instrument

	First Stage			OLS			IV		
	(1) Baseline	(2) Cross-section	(3) Panel	(4) Baseline	(5) Cross-section	(6) Panel	(7) Baseline	(8) Cross-section	(9) Panel
status <sub>rdt</sub> <sup>national_w(o)</sup>	0.744*** (0.0301)	0.406*** (0.0372)	0.237** (0.0922)						
taboo=1				0.194*** (0.00376)			0.230*** (0.00490)		
taboo=1 × ln income <sub>rdt</sub>				0.0423*** (0.00600)	0.0594*** (0.00881)	0.0103*** (0.00293)	0.264*** (0.0126)	0.120*** (0.0359)	0.0105*** (0.00399)
Observations	11,627	11,281	11,049	1,115,640	1,115,292	1,114,116	1,089,132	1,088,876	1,088,280
Adjusted R <sup>2</sup>	0.831	0.862	0.892	0.560	0.585	0.595	0.002	-0.013	-0.021
log prices and total expenditure controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
district*round*quarter	Yes	Yes	Yes	No	No	No	No	No	No
religion*state*round*quarter	No	Yes	No	No	No	No	No	No	No
religion*state*district*quarter	No	No	Yes	No	No	No	No	No	No
product*district*round*quarter	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*round*quarter	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	No	No	No	Yes	No	No	Yes	No
religion*state*product*district*quarter	No	No	No	No	Yes	No	No	Yes	No
First-stage F-statistic (CDF)									
First-stage F-statistic (RKF)							374508.0	71061.0	1185250.9
							1796.3	339.5	4116.6

Notes: Columns 1-3 present the first stage regressions using status<sub>rdt</sub><sup>national\_w(o)</sup> as an instrument for religious-group-specific income. Dependent variable in columns 1-3 is the weighted mean ln income<sub>rdt</sub> (monthly per capita expenditure). Columns 4-6 are the OLS regressions of taboo and the interaction of taboo and ln income<sub>rdt</sub> on taboo abstention, and columns 7-9 are the 2SLS regressions using status<sub>rdt</sub><sup>national\_w(o)</sup> to instrument ln income<sub>rdt</sub>. Columns 1, 4 and 7 include the baseline fixed effects, columns 2, 5 and 8 add the fixed effects for cross-sectional identification and columns 3, 6 and 9 for panel identification. Columns 1-3 adapt the sets of fixed effects to the unit of observation religion-district-time. Robust standard errors clustered at religion-district-round-quarter in parentheses. \* p < 0.10, \*\* p < 0.05, \*\*\* p < 0.01.

# E Cost of Identity

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)
$taboo_i=1$	0.188*** (0.00364)	0.167*** (0.0195)	0.225*** (0.0161)						
$\ln p_i$	-0.0331 (0.0407)	0.00183 (0.0415)	-0.0193 (0.0452)	-0.0233 (0.0383)	0.0284 (0.0445)	0.0215 (0.0483)	-0.0136 (0.0399)	-0.00736 (0.0432)	-0.0174 (0.0483)
$\sum \ln p_j$	-0.00251 (0.00173)	-0.0139*** (0.00222)	0.00183 (0.00543)	-0.00125 (0.00172)	-0.00854*** (0.00295)	-0.00367 (0.00547)	-0.000636 (0.00189)	-0.00431* (0.00259)	-0.00710* (0.00430)
$taboo_i=1 \times \ln p_i$		-0.0380*** (0.00343)	-0.0203** (0.00872)		-0.0536*** (0.0106)	-0.0489*** (0.0149)		-0.00692 (0.00658)	0.00228 (0.0128)
$taboo_i=1 \times \sum \ln p_j$		0.0148*** (0.00171)	-0.0297*** (0.00164)		0.00952*** (0.00316)	-0.0114*** (0.00416)		0.00466*** (0.00180)	-0.00155 (0.00339)
$\sum (\ln p_j \times taboo_j)$			-0.0555*** (0.00678)			-0.0164** (0.00791)			-0.000636 (0.00736)
$taboo_i=1 \times \sum (\ln p_j \times taboo_j)$			0.0909*** (0.00175)			0.0348*** (0.00670)			0.0114* (0.00586)
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 $R^2$	0.559	0.560	0.567	0.585	0.585	0.585	0.595	0.595	0.595
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*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)	5698.4	2835.9	2732.5	5587.6	2695.1	2609.2	5027.5	2474.9	2382.7
First-stage F-statistic (RKF)	37.87	18.95	18.88	37.22	18.67	18.56	32.83	16.48	16.43

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 $i$						
				Panel			
	Baseline	Cross-section	Panel	Cause of Conflict			
						Narrow	Relig Only
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
	All	All	All	Urban	Rural	All	All
taboo=1	0.181*** (0.00364)						
taboo=1 $\times$ conflict	0.106*** (0.0162)	0.0396*** (0.0107)	0.0599*** (0.00981)			0.0660*** (0.0101)	0.0594*** (0.0113)
taboo=1 $\times$ conflict past 2 quarters				0.0431** (0.0194)	0.0550*** (0.0134)		
taboo=1 $\times$ conflict present quarter				0.0587** (0.0293)	0.0293 (0.0187)		
Observations	1,112,876	1,112,536	1,111,356	344,880	764,264	1,111,356	1,111,356
Adjusted $R^2$	0.563	0.588	0.599	0.616	0.606	0.599	0.599
log prices and total expenditure	Yes	Yes	Yes	Yes	Yes	Yes	
household controls	Yes	Yes	Yes	Yes	Yes	Yes	
prod*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*state*prod*round*quarter	No	Yes	No	No	No	No	No
relig*state*prod*district*quarter	No	No	Yes	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 quarter in which the household is surveyed or in the preceding two quarters. Columns 4-6 differentiate the effect of a conflict occurring in the quarter of the survey from conflict in the preceding six months. 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. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and columns 3-7 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 E.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	0.196*** (0.00378)			0.212*** (0.00382)		
taboo=1 $\times$ status <sub>rdt</sub> <sup>national_occ(r)</sup>	0.120*** (0.0128)	0.0716*** (0.0120)	0.0444*** (0.0103)			
taboo=1 $\times$ status <sub>rdt</sub> <sup>national_w(o)</sup>				0.489*** (0.0203)	0.0662*** (0.0196)	0.0556*** (0.0213)
Observations	1,108,308	1,107,968	1,106,784	1,086,368	1,086,120	1,085,524
Adjusted $R^2$	0.563	0.587	0.598	0.564	0.587	0.596
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*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 E.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 $_i=1$	0.187*** (0.00359)	0.170*** (0.0184)	0.228*** (0.0158)						
ln $p_i$	0.0134*** (0.00270)	0.0434*** (0.00363)	0.0328*** (0.00484)	0.00962*** (0.00257)	0.0460*** (0.00509)	0.0429*** (0.00560)	0.00798*** (0.00255)	0.0214*** (0.00412)	0.0215*** (0.00523)
sum ln $p_j$	-0.00214 (0.00167)	-0.0132*** (0.00220)	0.00713** (0.00363)	-0.000868 (0.00165)	-0.00808*** (0.00283)	-0.00447 (0.00393)	-0.000323 (0.00165)	-0.00562*** (0.00198)	-0.00296 (0.00374)
taboo $_i=1$ x ln $p_i$		-0.0382*** (0.00278)	-0.0283*** (0.00452)		-0.0455*** (0.00487)	-0.0420*** (0.00553)		-0.0166*** (0.00351)	-0.0169*** (0.00506)
taboo $_i=1$ x sum ln $p_j$		0.0146*** (0.00171)	-0.0294*** (0.00159)		0.00952*** (0.00280)	-0.0118*** (0.00416)		0.00679*** (0.00131)	0.00198 (0.00265)
sum (ln $p_j$ x taboo $_j$ )			-0.0614*** (0.00419)			-0.0139** (0.00621)			-0.00765 (0.00568)
taboo $_i=1$ x sum (ln $p_j$ x taboo $_j$ )			0.0905*** (0.00171)			0.0342*** (0.00659)			0.0103* (0.00568)
Observations	1,112,876	1,112,876	1,112,876	1,112,536	1,112,536	1,112,536	1,111,356	1,111,356	1,111,356
Adjusted $R^2$	0.563	0.564	0.571	0.588	0.588	0.588	0.599	0.599	0.599
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	No	No	Yes	Yes	Yes	No	Yes	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 F4: Linear Approximation of Identity Choice with Cost, Status and Conflict, Household Controls

	LHS Variable: Share Spent on Good $i$		
	(1) Baseline	(2) Cross-section	(3) Panel
$(\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	-0.0351 (0.0535)	-0.696*** (0.0968)	-0.624*** (0.0868)
$(\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	0.479*** (0.0265)	0.246*** (0.0265)	0.234*** (0.0371)
$(\bar{x}_{ir} - \bar{x}_{is}) \times conflict_r$	0.549*** (0.0419)	0.0819** (0.0336)	0.110** (0.0435)
Observations	33,677,408	33,669,720	33,651,244
Adjusted $R^2$	0.767	0.773	0.778
log price and total expenditure controls	Yes	Yes	Yes
household controls	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes
religion*district*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.  $\bar{x}_{ir} - \bar{x}_{is}$  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.  $conflict_r$  is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. 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 Regressions without NSS Round 50

Table G.1: Religious Conflict and Taboo Adherence, Only NSS 43 and 55

	LHS Variable: Abstain from Consuming Good $i$						
	Baseline	Cross-section	Panel	Panel			
				Cause of Conflict			
				Narrow		Relig Only	
(1) All	(2) All	(3) All	(4) Urban	(5) Rural	(6) All	(7) All	
taboo=1	0.187*** (0.00431)						
taboo=1 × conflict	0.0863*** (0.0174)	0.0159 (0.0105)	0.0577*** (0.0136)			0.0694*** (0.0140)	0.0464*** (0.0148)
taboo=1 × conflict past 2 quarters				0.0431** (0.0194)	0.0493** (0.0196)		
taboo=1 × conflict present quarter				0.0587** (0.0293)	0.0222 (0.0300)		
Observations	878,524	878,300	876,956	347,556	527,104	876,956	876,956
Adjusted $R^2$	0.557	0.582	0.594	0.612	0.604	0.594	0.594
log prices and total expenditure	Yes	Yes	Yes	Yes	Yes	Yes	Yes
prod*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*state*prod*round*quarter	No	Yes	No	No	No	No	No
relig*state*prod*district*quarter	No	No	Yes	Yes	Yes	Yes	Yes

*Notes:* All regressions are restricted to NSS rounds 43 and 55 (i.e., without NSS round 50). 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 quarter in which the household is surveyed or in the preceding two quarters. Columns 4-6 differentiate the effect of a conflict occurring in the quarter of the survey from conflict in the preceding six months. Column 5 restricts the analysis to the urban population, and column 6 to the rural population. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and columns 3-7 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.2: Status and Choice of Identity, Only NSS 43 and 55

	LHS Variable: Abstain from Consuming Good $i$					
	Baseline	Cross-section	Panel	Baseline	Cross-section	Panel
	(1)	(2)	(3)	(4)	(5)	(6)
taboo=1	0.199*** (0.00430)			0.207*** (0.00426)		
taboo=1 $\times$ status <sub>rdt</sub> <sup>national_occ(r)</sup>	0.125*** (0.0150)	0.0772*** (0.0137)	0.0508*** (0.0122)			
taboo=1 $\times$ status <sub>rdt</sub> <sup>national_w(o)</sup>				0.471*** (0.0231)	0.0867*** (0.0220)	0.0634*** (0.0242)
Observations	873,956	873,732	872,396	860,272	860,076	859,316
Adjusted $R^2$	0.556	0.581	0.593	0.558	0.580	0.591
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*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: All regressions are restricted to NSS rounds 43 and 55 (i.e., without NSS round 50). 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table G.3: Costs, Price Elasticities and Identity, Only NSS 43 and 55

	LHS Variable: Abstain from Consuming Good $i$								
	Baseline			Cross-section			Panel		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
taboo <sub><math>i</math></sub> =1	0.193*** (0.00422)	0.190*** (0.0193)	0.240*** (0.0165)						
ln p <sub><math>i</math></sub>	0.0151*** (0.00291)	0.0491*** (0.00408)	0.0408*** (0.00545)	0.0107*** (0.00275)	0.0498*** (0.00580)	0.0474*** (0.00633)	0.00938*** (0.00274)	0.0232*** (0.00486)	0.0252*** (0.00604)
sum ln p <sub><math>j</math></sub>	-0.00247 (0.00181)	-0.0135*** (0.00238)	0.00877** (0.00398)	-0.000993 (0.00179)	-0.00620** (0.00313)	-0.000317 (0.00434)	-0.000558 (0.00180)	-0.00579*** (0.00221)	0.0000964 (0.00412)
taboo <sub><math>j</math></sub> =1 x ln p <sub><math>i</math></sub>		-0.0426*** (0.00321)	-0.0361*** (0.00511)		-0.0484*** (0.00560)	-0.0458*** (0.00627)		-0.0169*** (0.00429)	-0.0194*** (0.00587)
taboo <sub><math>j</math></sub> =1 x sum ln p <sub><math>j</math></sub>		0.0145*** (0.00184)	-0.0292*** (0.00174)		0.00694** (0.00311)	-0.0166*** (0.00479)		0.00665*** (0.00152)	0.000727 (0.00310)
sum (ln p <sub><math>j</math></sub> x taboo <sub><math>j</math></sub> )			-0.0641*** (0.00461)			-0.0178** (0.00693)			-0.0125* (0.00653)
taboo <sub><math>i</math></sub> =1 x sum (ln p <sub><math>j</math></sub> x taboo <sub><math>j</math></sub> )			0.0912*** (0.00203)			0.0378*** (0.00751)			0.0126* (0.00681)
Observations	878,524	878,524	878,524	878,300	878,300	878,300	876,956	876,956	876,956
Adjusted R <sup>2</sup>	0.557	0.557	0.565	0.581	0.582	0.582	0.594	0.594	0.594
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*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: All regressions are restricted to NSS rounds 43 and 55 (i.e., without NSS round 50). 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. 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 G.4: Linear Approximation of Identity Choice with Cost, Status and Conflict, Only NSS 43 and 55

	LHS Variable: Share Spent on Good $i$		
	(1) Baseline	(2) Cross-section	(3) Panel
$(\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	-0.172*** (0.0611)	-0.860*** (0.112)	-0.752*** (0.101)
$(\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	0.539*** (0.0309)	0.277*** (0.0305)	0.309*** (0.0434)
$(\bar{x}_{ir} - \bar{x}_{is}) \times conflict_r$	0.545*** (0.0442)	0.0869** (0.0345)	0.144** (0.0563)
Observations	26,668,432	26,662,356	26,638,796
Adjusted $R^2$	0.753	0.759	0.766
log prices and total expenditure controls	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes
religion*district*round*quarter	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No
religion*state*product*district*quarter	No	No	Yes

*Notes:* All regressions are restricted to NSS rounds 43 and 55 (i.e., without NSS round 50). Dependent variable is the share spent on good  $i$  in total food expenditure.  $\bar{x}_{ir} - \bar{x}_{is}$  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.  $conflict_r$  is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## H Baseline Taboo Regressions with Ovo-Pesco Vegetarianism

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

	LHS Variable: Abstain from Consuming Good $i$						
				Panel			
	Baseline	Cross-section	Panel	Cause of Conflict			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
All	All	All	Urban	Rural	All	All	
taboo=1	0.192*** (0.00389)						
taboo=1 × conflict	0.122*** (0.0169)	0.0515*** (0.0107)	0.0715*** (0.0103)			0.0761*** (0.0106)	0.0750*** (0.0123)
taboo=1 × conflict past 2 quarters				0.0448** (0.0205)	0.0714*** (0.0146)		
taboo=1 × conflict present quarter				0.0852** (0.0333)	0.0287 (0.0198)		
Observations	1,115,640	1,115,292	1,114,116	347,556	764,344	1,114,116	1,114,116
Adjusted $R^2$	0.423	0.454	0.467	0.528	0.465	0.467	0.467
log prices and total expenditure	Yes	Yes	Yes	Yes	Yes	Yes	Yes
prod*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
relig*state*prod*round*quarter	No	Yes	No	No	No	No	No
relig*state*prod*district*quarter	No	No	Yes	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 quarter in which the household is surveyed or in the preceding two quarters. Columns 4-6 differentiate the effect of a conflict occurring in the quarter of the survey from conflict in the preceding six months. Column 5 restricts the analysis to the urban population, and column 6 to the rural population. Column 1 includes the baseline fixed effects, column 2 adds the fixed effects for cross-sectional identification and columns 3-7 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: 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	0.210*** (0.00403)			0.227*** (0.00410)		
taboo=1 $\times$ status <sub>rdt</sub> <sup>national_occ(r)</sup>	0.154*** (0.0134)	0.0750*** (0.0125)	0.0341*** (0.0109)			
taboo=1 $\times$ status <sub>rdt</sub> <sup>national_w(o)</sup>				0.554*** (0.0212)	0.0903*** (0.0206)	0.0141 (0.0232)
Observations	1,111,072	1,110,724	1,109,544	1,089,132	1,088,876	1,088,280
Adjusted $R^2$	0.424	0.454	0.467	0.425	0.452	0.464
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*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 H.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	0.199*** (0.00384)	0.217*** (0.0201)	0.286*** (0.0166)						
ln $p_i$	0.0102*** (0.00280)	0.0164*** (0.00369)	0.0149*** (0.00531)	0.00520** (0.00265)	0.0439*** (0.00546)	0.0429*** (0.00617)	0.00371 (0.00264)	0.0182*** (0.00453)	0.0203*** (0.00589)
sum ln $p_j$	-0.00409** (0.00180)	-0.00473** (0.00233)	0.0230*** (0.00424)	-0.00244 (0.00178)	-0.00923*** (0.00327)	0.000949 (0.00477)	-0.00194 (0.00178)	-0.00591*** (0.00223)	0.00163 (0.00458)
taboo $_i$ =1 x ln $p_i$		-0.00786*** (0.00275)	-0.0106** (0.00500)		-0.0481*** (0.00515)	-0.0472*** (0.00605)		-0.0178*** (0.00391)	-0.0204*** (0.00570)
taboo $_i$ =1 x sum ln $p_j$		0.000886 (0.00173)	-0.0472*** (0.00155)		0.00892*** (0.00315)	-0.0192*** (0.00448)		0.00507*** (0.00147)	-0.00347 (0.00306)
sum (ln $p_j$ x taboo $_j$ )			-0.0825*** (0.00478)			-0.0279*** (0.00787)			-0.0175** (0.00728)
taboo $_i$ =1 x sum (ln $p_j$ x taboo $_j$ )			0.110*** (0.00171)			0.0485*** (0.00783)			0.0189*** (0.00700)
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 $R^2$	0.423	0.423	0.436	0.454	0.454	0.454	0.467	0.467	0.467
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
region*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
region*state*product*round*quarter	No	No	No	Yes	Yes	Yes	No	No	No
region*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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

# I Linear Approximation of Identity Choice

Table I.1: Linear Approximation of Identity Choice , NSS 43, with and without Interstate Migrants

	LHS Variable: Share Spent on Good $i$			
	All		Without Interstate Migrants	
	(1) Baseline	(2) Cross-section	(3) Baseline	(4) Cross-section
$(\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	-0.267*** (0.0753)	-1.051*** (0.150)	-0.216*** (0.0746)	-1.000*** (0.149)
$(\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	0.623*** (0.0407)	0.432*** (0.0407)	0.606*** (0.0412)	0.419*** (0.0423)
$(\bar{x}_{ir} - \bar{x}_{is}) \times conflict_r$	0.545*** (0.0562)	0.102*** (0.0385)	0.538*** (0.0546)	0.0795* (0.0418)
Observations	13,994,020	13,989,680	12,536,772	12,532,184
Adjusted $R^2$	0.740	0.746	0.746	0.752
log prices and total expenditure controls	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes
religion*district*round*quarter	Yes	Yes	Yes	Yes
religion*state*product*round*quarter	No	Yes	No	Yes

*Notes:* All regressions are restricted to the NSS 43rd round – the only round for which we observe interstate migrants. Columns 1-2 are estimated using the entire NSS 43 sample, and columns 3-4 omit households with at least one inter-State migrant. Prototypical bundles are recomputed accordingly. Dependent variable is the share spent on good  $i$  in total food expenditure.  $\bar{x}_{ir} - \bar{x}_{is}$  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.  $conflict_r$  is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. Columns 1 and 3 include the baseline fixed effects, columns 2 and 4 adds the fixed effects for cross-sectional 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 I.2: Linear Approximation of Identity Choice with Cost, Status and Conflict, Including Cross-Price Effects

	LHS Variable: Share Spent on Good $i$					
	(1)	(2)	(3)	(4)	(5)	(6)
$(\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	-0.0644 (0.0600)	-0.646*** (0.109)	-0.635*** (0.0990)	-0.0339 (0.0592)	-0.606*** (0.108)	-0.597*** (0.0981)
$(\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	0.479*** (0.0287)	0.272*** (0.0317)	0.232*** (0.0462)	0.476*** (0.0286)	0.271*** (0.0317)	0.228*** (0.0461)
$(\bar{x}_{ir} - \bar{x}_{is}) \times conflict_r$	0.546*** (0.0461)	0.0473 (0.0395)	0.104** (0.0494)	0.544*** (0.0459)	0.0468 (0.0394)	0.104** (0.0493)
Observations	16,878,102	16,868,583	16,845,763	16,741,240	16,731,823	16,709,148
Adjusted $R^2$	0.770	0.777	0.782	0.772	0.778	0.784
log prices and total expenditure controls	Yes	Yes	Yes	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes	Yes	Yes	Yes
religion*district*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.  $\bar{x}_{ir} - \bar{x}_{is}$  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.  $conflict_r$  is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table I.3: Identity Choice with Cost, Status and Conflict, Relaxing Symmetry across Religious and Ethnic Identities

	LHS Variable: Share Spent on Good $i$		
	(1) Baseline	(2) Cross-section	(3) Panel
$\overline{x_{ir}} \times (cost_r - cost_s)$	0.0816* (0.0469)	-0.359*** (0.0939)	-0.297*** (0.0798)
$\overline{x_{is}} \times (cost_r - cost_s)$	0.0608 (0.0577)	0.619*** (0.0928)	0.568*** (0.0833)
$\overline{x_{ir}} \times (status_r - status_s)$	0.294*** (0.0225)	0.149*** (0.0233)	0.149*** (0.0351)
$\overline{x_{is}} \times (status_r - status_s)$	-0.490*** (0.0258)	-0.258*** (0.0271)	-0.252*** (0.0387)
$\overline{x_{ir}} \times conflict_r$	0.532*** (0.0403)	0.0700** (0.0329)	0.105*** (0.0396)
$\overline{x_{is}} \times conflict_r$	-0.456*** (0.0634)	-0.135** (0.0672)	-0.135 (0.0954)
Observations	33,763,092	33,755,156	33,736,680
Adjusted $R^2$	0.767	0.773	0.778
log prices and total expenditure controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*district*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.  $conflict_r$  is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. 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 I.4: 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 $\times (\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	0.182*** (0.0549)	-0.646*** (0.132)	-0.564*** (0.0874)
Hindu UC $\times (\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	-0.0989* (0.0572)	-0.734*** (0.120)	-0.639*** (0.0958)
Muslim $\times (\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	0.278*** (0.0719)	-0.557*** (0.182)	-0.660*** (0.110)
Christian $\times (\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	-0.136 (0.201)	-0.579 (0.428)	-1.051*** (0.254)
Hindu SC $\times (\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	-0.0429 (0.0426)	0.0560 (0.0426)	0.0675 (0.0610)
Hindu UC $\times (\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	1.585*** (0.0755)	0.827*** (0.0875)	0.759*** (0.123)
Muslim $\times (\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	0.351*** (0.0554)	0.152*** (0.0462)	0.180*** (0.0677)
Christian $\times (\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	0.201** (0.0915)	0.207** (0.0830)	0.374** (0.174)
Hindu SC $\times (\bar{x}_{ir} - \bar{x}_{is}) \times conflict_r$	0.453*** (0.0460)	0.0747** (0.0349)	0.0754* (0.0432)
Hindu UC $\times (\bar{x}_{ir} - \bar{x}_{is}) \times conflict_r$	0.508*** (0.0516)	0.104*** (0.0394)	0.104** (0.0471)
Muslim $\times (\bar{x}_{ir} - \bar{x}_{is}) \times conflict_r$	0.543*** (0.0465)	0.0815** (0.0381)	0.177*** (0.0454)
Christian $\times (\bar{x}_{ir} - \bar{x}_{is}) \times conflict_r$	0 (.)	0 (.)	0 (.)
Observations	33,763,092	33,755,156	33,736,680
Adjusted $R^2$	0.767	0.773	0.778
log prices and total expenditure controls	Yes	Yes	Yes
district*product*round*quarter	Yes	Yes	Yes
religion*district*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.  $\bar{x}_{ir} - \bar{x}_{is}$  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.  $conflict_r$  is an indicator for at least one occurrence of Hindu-Muslim conflict in the district in the quarter in which the household is surveyed or in the preceding two quarters. 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 I.5: Linear Approximation of Identity Choice with Cost and Status, 1987-2010 (NSS 43rd to NSS 66th)

	LHS Variable: Share Spent on Good $i$		
	(1) Baseline	(2) Cross-section	(3) Panel
$(\bar{x}_{ir} - \bar{x}_{is}) \times (cost_r - cost_s)$	-0.0128 (0.0397)	-0.750*** (0.0789)	-0.403*** (0.0466)
$(\bar{x}_{ir} - \bar{x}_{is}) \times (status_r - status_s)$	0.395*** (0.0218)	0.172*** (0.0199)	0.116*** (0.0226)
Observations	56,793,748	56,783,601	56,773,769
Adjusted $R^2$	0.761	0.768	0.771
log prices and total expenditure controls	Yes	Yes	Yes
product*district*round*quarter	Yes	Yes	Yes
religion*district*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.  $\bar{x}_{ir} - \bar{x}_{is}$  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. 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## J Counterfactuals

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

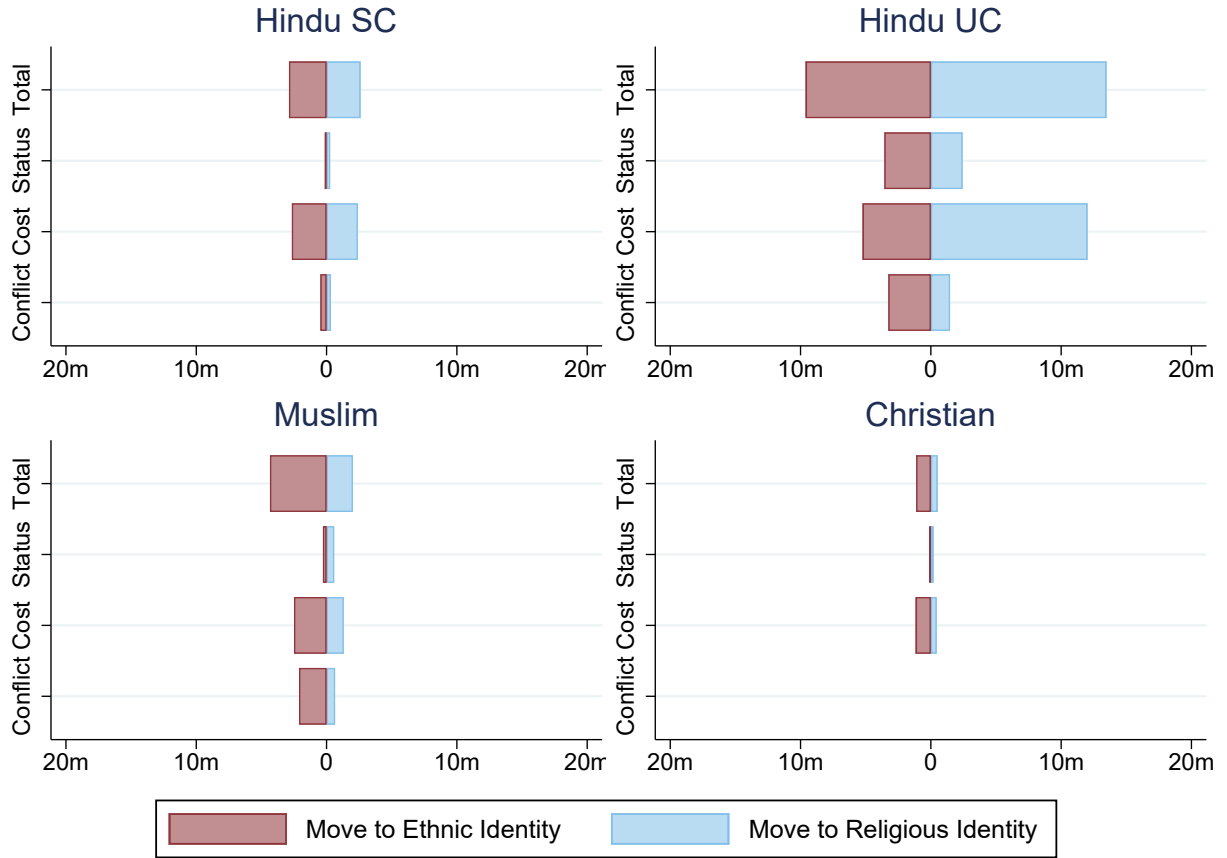


Figure J.2: Proportion of Population Changing Identity across District-Religion Cells in Gujarat, Maharashtra and Uttar Pradesh, 1987-2000

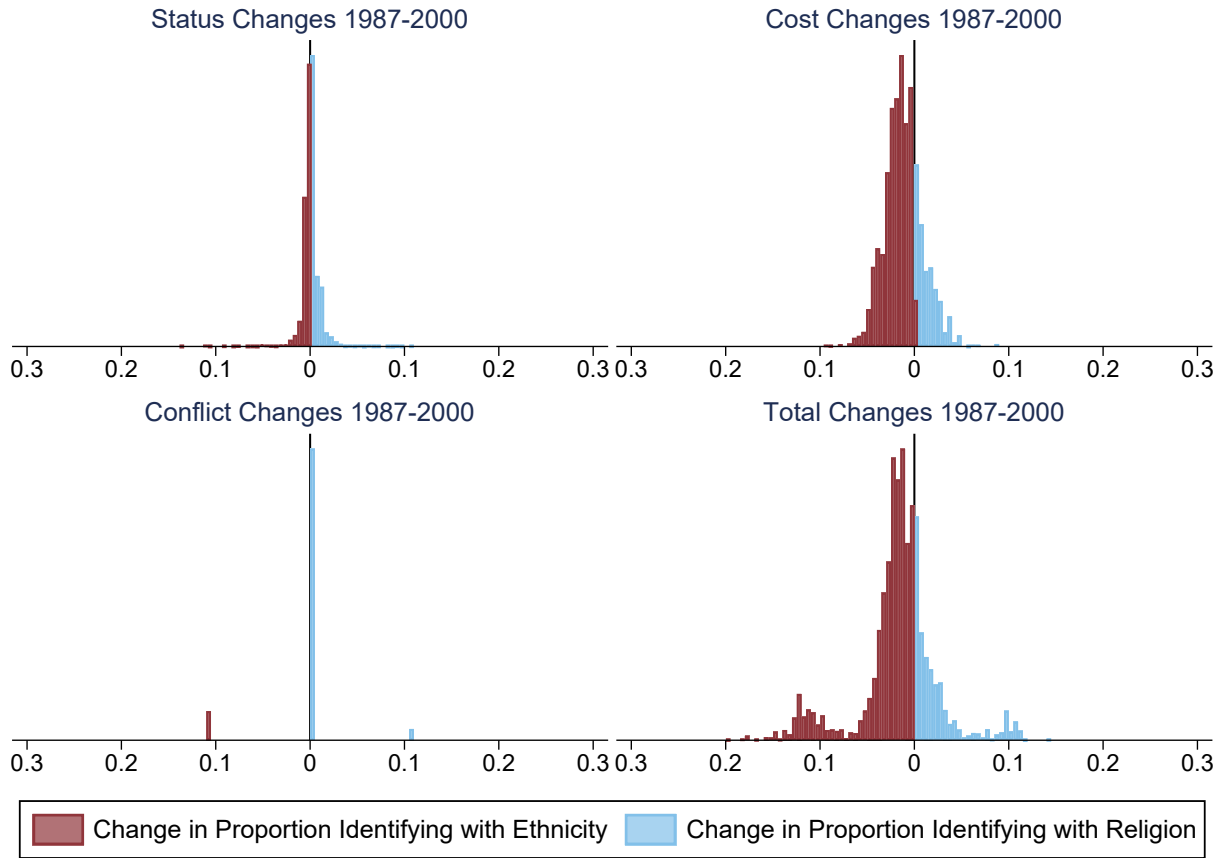




Figure J.3: Realized Compensating Variation Gains from Identity Changes, 1987-2000

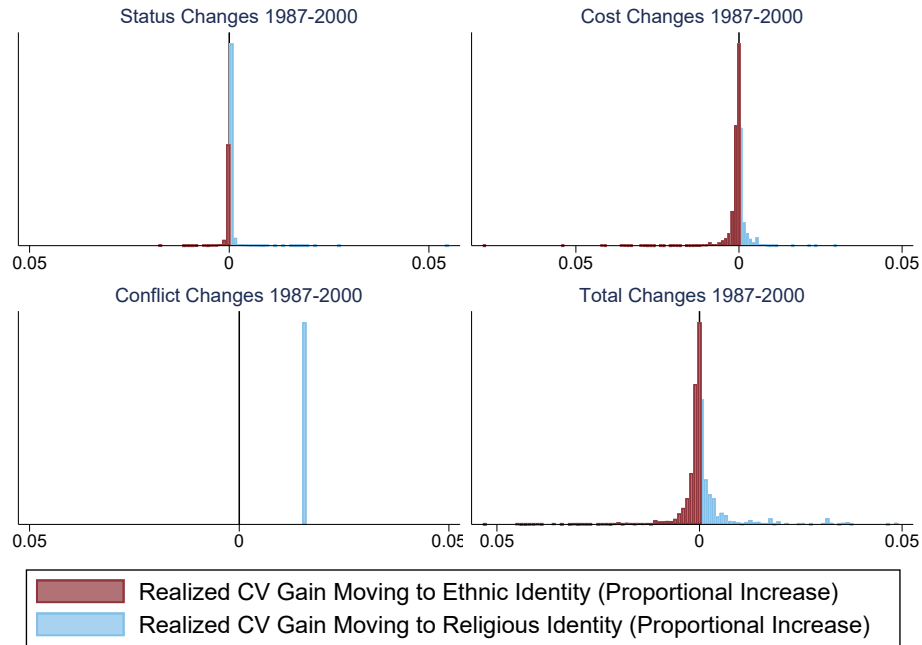
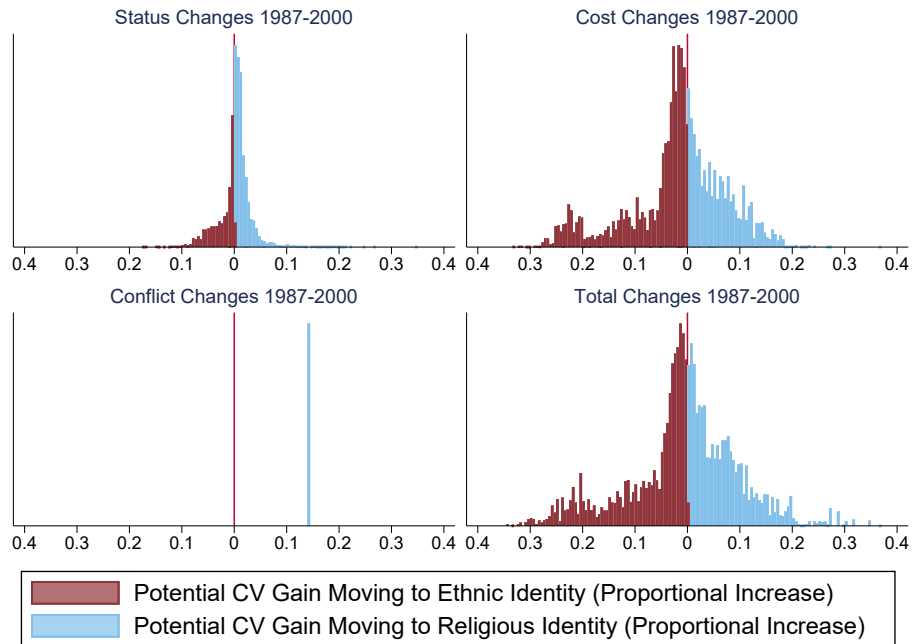


Figure J.4: Potential Compensating Variation Gains from Identity Changes, 1987-2000



## K Ethnic and Religious Parties

Using party names from the Jensenius and Verniers (2017) database, we classify parties into four categories: ethnic/regional; Muslim; Hindu upper caste; and Hindu scheduled caste. This classification is derived from each political party's founding manifesto, the political almanac published by Centre for Public Affairs, and media reports on the party's political philosophy and outreach. Table K.1 below reports the resulting list of religious parties and Table K.2 reports the list of ethnic parties.

An alternative classification of ethnic parties uses the category "State Party" derived from the election commission classification. Every year, the election commission classifies parties to four types: (1) *National Parties* (currently 7); (2) *State Parties* (Currently 49); (3) *Unrecognized regional parties* (1785); and (4) *Independents*. Table K.2 provides the list of State Parties.

Table K.1: Classification of Religious Parties

Religion	Religious Parties
Hindu Upper Caste	Bhartiya Janta Party
Hindu Scheduled Caste	Bahujan Samaj Party Lok Jan Shakti Party Pattali Makkal Katchi
Muslim	Indian Union Muslim League Muslim League Kerala State Committee

Table K.2: Classification of Ethnic Parties

State	Ethnic Parties	State Parties (Election Commission)
Andhra Pradesh	Lok Dal Nationalist Congress Party Ntr Telugu Desam Party (Lakshmi Parvathi) Samajwadi Janata Party Shivsena Sikkim Sangram Parishad Telangana Rashtra Samithi Telugu Desam Party	Indian Union Muslim League Lok Dal Muslim League Kerala State Committee Ntr Telugu Desam Party (Lakshmi Parvathi) Rashtriya Janata Dal Samajwadi Janata Party Samajwadi Party Shivsena Sikkim Sangram Parishad Telangana Rashtra Samithi Telugu Desam Party
Arunachal Pradesh	Arunachal Congress Nationalist Congress Party	Arunachal Congress
Assam	All India Trinamool Congress	All India Trinamool Congress

Continued on next page

Table K.2: Classification of Ethnic Parties (Continued)

<b>State</b>	<b>Ethnic Parties</b>	<b>State Parties (Election Commission)</b>
	Asom Gana Parishad Lok Dal Nationalist Congress Party Plains Tribals Council Of Assam Shivsena United Minorities Front Assam	Asom Gana Parishad Autonomous State Demand Committee Indian Congress (Socialist) Janata Dal (United) Lok Dal Muslim League Kerala State Committee Plains Tribals Council Of Assam Rashtriya Janata Dal Revolutionary Socialist Party Samajwadi Party Shivsena United Minorities Front Assam
Bihar	All India Jharkhand Party Jharkhand Mukti Morcha Lok Dal Nationalist Congress Party Shivsena United Goans Democratic Party	All India Forward Block All India Jharkhand Party Indian Congress (Socialist) Indian Congress (J) Indian Union Muslim League Janata Dal (United) Jharkhand Mukti Morcha Lok Dal Muslim League Kerala State Committee Rashtriya Janata Dal Revolutionary Socialist Party Rld Samajwadi Party Shivsena United Goans Democratic Party
Chhattisgarh	Jharkhand Mukti Morcha Nationalist Congress Party Shivsena	Janata Dal (United) Jharkhand Mukti Morcha Samajwadi Party Shivsena
Delhi	All India Trinamool Congress Indian National Lok Dal Jammu & Kashmir National Panthers Party Jharkhand Mukti Morcha Lok Dal Nationalist Congress Party Shivsena Uttarakhand Kranti Dal	All India Forward Block All India Trinamool Congress Indian Congress (Socialist) Indian National Lok Dal Indian Union Muslim League Jammu & Kashmir National Panthers Party Janata Dal (United) Jharkhand Mukti Morcha Lok Dal Rashtriya Janata Dal

Continued on next page

Table K.2: Classification of Ethnic Parties (Continued)

State	Ethnic Parties	State Parties (Election Commission)
		Rld Samajwadi Party Shivsena Uttarakhand Kranti Dal
Goa	Nationalist Congress Party Shivsena United Goans Democratic Party	Samajwadi Party Shivsena United Goans Democratic Party
Gujarat	Lok Dal Nationalist Congress Party Shivsena	Indian Union Muslim League Janata Dal (United) Lok Dal Rashtriya Janata Dal Revolutionary Socialist Party Samajwadi Party Shivsena
Haryana	Haryana Vikas Party Indian National Lok Dal Lok Dal Nationalist Congress Party Shivsena	All India Forward Block Haryana Vikas Party Indian Congress (Socialist) Indian National Lok Dal Janata Dal (United) Lok Dal Rashtriya Janata Dal Rld Samajwadi Party Shivsena
Himachal Pradesh	Lok Dal Nationalist Congress Party Shivsena	Indian Congress (Socialist) Lok Dal Samajwadi Party Shivsena
Jammu & Kashmir	All India Jharkhand Party Jammu & Kashmir National Conference Jammu & Kashmir National Panthers Party Lok Dal Nationalist Congress Party Samajwadi Janata Party Shivsena	All India Jharkhand Party Jammu & Kashmir National Conference Jammu & Kashmir National Panthers Party Janata Dal (United) Lok Dal Rashtriya Janata Dal Samajwadi Janata Party Shivsena
Karnataka	All India Trinamool Congress Kerala Congress (Pillai Group)	All India Trinamool Congress Indian Congress (Socialist)

Continued on next page

Table K.2: Classification of Ethnic Parties (Continued)

<b>State</b>	<b>Ethnic Parties</b>	<b>State Parties (Election Commission)</b>
	Lok Dal Nationalist Congress Party Samajwadi Janata Party Shivsena	Indian Union Muslim League Janata Dal (United) Kerala Congress (Pillai Group) Lok Dal Muslim League Kerala State Committee Rashtriya Janata Dal Samajwadi Janata Party Samajwadi Party Shivsena
Kerala	Lok Dal Nationalist Congress Party Shivsena	All India Forward Block Indian Congress (Socialist) Janata Dal (United) Kerala Congress Kerala Congress (M) Lok Dal Muslim League Kerala State Committee Rashtriya Janata Dal Revolutionary Socialist Party Shivsena
Madhya Pradesh	Indian National Lok Dal Jharkhand Mukti Morcha Lok Dal Nationalist Congress Party Shivsena	All India Forward Block Indian Congress (Socialist) Indian National Lok Dal Indian Union Muslim League Janata Dal (United) Jharkhand Mukti Morcha Lok Dal Muslim League Kerala State Committee Rashtriya Janata Dal Revolutionary Socialist Party Rld Samajwadi Party Shivsena
Maharashtra	Lok Dal Nationalist Congress Party Shivsena Sikkim Sangram Parishad	All India Forward Block Indian Congress (Socialist) Indian Union Muslim League Janata Dal (United) Lok Dal Muslim League Kerala State Committee Peasant'S & Workers' Party Rashtriya Janata Dal Rld

Continued on next page

Table K.2: Classification of Ethnic Parties (Continued)

State	Ethnic Parties	State Parties (Election Commission)
		Samajwadi Party Shivsena Sikkim Sangram Parishad
Manipur	Federal Party Of Manipur Kuki National Assembly Manipur People'S Party Nationalist Congress Party	Federal Party Of Manipur Indian Congress (Socialist) Janata Dal (United) Kuki National Assembly Manipur People'S Party Rashtriya Janata Dal Revolutionary Socialist Party
Meghalaya	Hill State People'S Democratic Party Manipur People'S Party Nationalist Congress Party	Hill State People'S Democratic Party Manipur People'S Party Rashtriya Janata Dal Samajwadi Party United Democratic Party
Mizoram	Mizo National Front Peoples Conference	Janata Dal (United) Mizo National Front Peoples Conference Rashtriya Janata Dal
Nagaland	All India Trinamool Congress Naga National Democratic Party Naga Peoples Front Nationalist Congress Party	All India Trinamool Congress Janata Dal (United) Naga National Democratic Party Naga Peoples Front Rld
Orissa	All India Jharkhand Party All India Trinamool Congress Jharkhand Mukti Morcha Lok Dal Nationalist Congress Party Shivsena Sikkim Sangram Parishad	All India Forward Block All India Jharkhand Party All India Trinamool Congress Biju Janata Dal Indian Congress (Socialist) Indian Conress (J) Janata Dal (United) Jharkhand Mukti Morcha Lok Dal Rashtriya Janata Dal Revolutionary Socialist Party Samajwadi Party Shivsena Sikkim Sangram Parishad

Continued on next page

Table K.2: Classification of Ethnic Parties (Continued)

<b>State</b>	<b>Ethnic Parties</b>	<b>State Parties (Election Commission)</b>
Punjab	Lok Dal Nationalist Congress Party Shivsena	All India Forward Block Indian Congress (Socialist) Janata Dal (United) Lok Dal Rashtriya Janata Dal Rld Samajwadi Party Shiromani Akali Dal Shivsena
Rajasthan	All India Trinamool Congress Indian National Lok Dal Jammu & Kashmir National Panthers Party Lok Dal Nationalist Congress Party Shivsena Sikkim Sangram Parishad	All India Forward Block All India Trinamool Congress Indian Congress (J) Trikha Group Indian Congress (Socialist) Indian Congress (J) Indian National Lok Dal Indian Union Muslim League Jammu & Kashmir National Panthers Party Janata Dal (United) Lok Dal Muslim League Kerala State Committee Rashtriya Janata Dal Revolutionary Socialist Party Rld Samajwadi Party Shivsena Sikkim Sangram Parishad
Sikkim	Sikkim Democratic Front Sikkim Sangram Parishad	Revolutionary Socialist Party Sikkim Democratic Front Sikkim Sangram Parishad
Tamil Nadu	Dravida Munnetra Kazhagam Lok Dal Manipur People'S Party Marumalarchi Dravida Munnetra Kazhagam Nationalist Congress Party Pondichery Mannila Makkal Munnani Shivsena	All India Forward Block Dravida Munnetra Kazhagam Indian Congress (Socialist) Indian Union Muslim League Janata Dal (United) Lok Dal Manipur People'S Party Marumalarchi Dravida Munnetra Kazhagam Muslim League Kerala State Committee Pattali Makkal Katchi Pondichery Mannila Makkal Munnani Rashtriya Janata Dal

Continued on next page

Table K.2: Classification of Ethnic Parties (Continued)

State	Ethnic Parties	State Parties (Election Commission)
		Revolutionary Socialist Party Rld Samajwadi Party Shivsena Tamil Maanila Congress (Moopanar)
Tripura	All India Trinamool Congress Nationalist Congress Party	All India Forward Block All India Trinamool Congress Janata Dal (United) Rashtriya Janata Dal Revolutionary Socialist Party Tripura Upajati Juba Samiti
Uttar Pradesh	Indian National Lok Dal Lok Dal Nationalist Congress Party Samajwadi Janata Party Shivsena Sikkim Sangram Parishad Uttarakhand Kranti Dal	All India Forward Block Indian Congress (Socialist) Indian Congress (J) Indian National Lok Dal Indian Union Muslim League Janata Dal (United) Lok Dal Muslim League Kerala State Committee Rashtriya Janata Dal Revolutionary Socialist Party Rld Samajwadi Janata Party Samajwadi Party Shivsena Sikkim Sangram Parishad Uttarakhand Kranti Dal
West Bengal	All India Jharkhand Party All India Trinamool Congress Jharkhand Mukti Morcha Lok Dal Nationalist Congress Party Samajwadi Janata Party Shivsena	All India Forward Block All India Jharkhand Party All India Trinamool Congress Indian Union Muslim League Janata Dal (United) Jharkhand Mukti Morcha Lok Dal Muslim League Kerala State Committee Rashtriya Janata Dal Revolutionary Socialist Party Samajwadi Janata Party Samajwadi Party Shivsena



Table K.3: Vote Shares for Religious and Ethnic Parties by Round

	1987–1988	1999–2000
<b>Vote Share Hindu Upper Caste Parties (%)</b>		
Mean	10.92	17.16
Std. Dev.	12.69	12.76
N district-religion cells	325	325
<b>Vote Share Hindu Scheduled Caste Parties (%)</b>		
Mean	3.20	8.18
Std. Dev.	5.96	10.12
N district-religion cells	287	287
<b>Vote Share Muslim Parties (%)</b>		
Mean	2.29	1.63
Std. Dev.	7.34	7.17
N district-religion cells	117	117
<b>Vote Share Ethnic Parties (%)</b>		
Mean	9.02	15.21
Std. Dev.	14.92	18.67
N district-religion cells	729	729
<b>Vote Share State Parties (%)</b>		
Mean	11.04	27.21
Std. Dev.	15.53	16.65
N district-religion cells	729	729

*Notes:* Weighted mean, standard deviation and number of observations of the vote share data underlying Column 1 of Table 8.