

NBER WORKING PAPER SERIES

THE IMPACT OF SOCIOECONOMIC AND CULTURAL DIFFERENCES ON ONLINE
TRADE

Daniel W. Elfenbein
Raymond Fisman
Brian McManus

Working Paper 26197
<http://www.nber.org/papers/w26197>

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

We are grateful for eBay's support for this research, especially Sharat Raghavan, Dimitriy Masterov, and Tom Blake. Eli Locicero provided excellent research assistance. We thank Rachel Croson, Dave Donaldson, Liran Einav, Ali Hortascu, and Avri Ravid, as well as seminar participants at the University of Southern California and the ASSA 2019 meetings, for helpful 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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.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 Daniel W. Elfenbein, Raymond Fisman, and Brian McManus. 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.

The Impact of Socioeconomic and Cultural Differences on Online Trade
Daniel W. Elfenbein, Raymond Fisman, and Brian McManus
NBER Working Paper No. 26197
August 2019
JEL No. D12,D91,F14,L15,R12

ABSTRACT

We use U.S. eBay data to investigate how trade is influenced by differences in socioeconomic characteristics, tastes, and trust. States' similarity in cultural characteristics (ethnicity, religious affiliations, and political behavior) is predictive of online trade; cultural similarity similarly predicts trade between finer (three-digit zip code) geographies. The culture-trade relationship is mediated in part by consumers' tastes, and is stronger for transactions with sellers who lack extensive reputations or certification, suggesting that consumers infer seller trustworthiness from cultural similarity. There is no correlation between cultural similarity and buyer satisfaction, consistent with perceived differences in trustworthiness not being validated by actual transactions.

Daniel W. Elfenbein
Olin School of Business
Washington University in St. Louis
St Louis, MO 63130
elfenbein@wustl.edu

Brian McManus
Department of Economics CB 3305
University of North Carolina
Chapel Hill, NC 27599
mcmanusb@unc.edu

Raymond Fisman
Department of Economics
Boston University
270 Bay State Road, 304A
Boston, MA 02215
and NBER
rfisman@bu.edu

1. Introduction

Trade requires a match between a seller's production or inventory and a buyer's demands, as well as buyer-seller trust. The match in product characteristics may be higher and the information frictions lower for more similar buyer-seller pairs, which would imply higher trade flows between market participants that share socio-cultural attributes that may be drivers of taste and (perceived) trust.

In this paper, we use data from eBay transactions in the U.S. in the years 2015 and 2016 to document socioeconomic and cultural similarity as drivers of online trade, and we also present evidence suggesting that both shared tastes and greater trust underlie the culture-trade relationship. Several key features of our setting allow us to focus on a subset of potential drivers of trade flows, particularly for online platforms that constitute a sizeable and ever-larger share of retail sales worldwide, growing 2.7 times faster than retail in the U.S. overall during 2009-2018.¹ First, because all of the transactions take place on the same platform and are governed by the same regulations (both national laws and platform-specific rules), we hold constant many of the institutional factors that might complicate the interpretation of earlier studies linking similarity to trade based on country-pair analyses (e.g., see Anderson and Van Wincoop's (2004) survey of the gravity model trade literature). Furthermore, as a platform that directly connects individual buyers to sellers, eBay allows us to focus directly on *consumer* responses to socioeconomic and cultural similarities and differences.² Prior work investigating these relationships has, for the most part, examined the combined trade in intermediate and final goods, even when implemented within a single country as in Hillberry and Hummels (2008), and has generally focused on trade between countries (e.g., Guiso, Sapienza and Zingales, 2009).

¹ Calculations based on data from the St. Louis Federal Reserve, <https://fred.stlouisfed.org/series/ECOMSA> (accessed May 15, 2019).

² An eBay seller's location is a prominent feature of each eBay product listing. Although information about the seller's buying and selling history on the eBay platform is available to buyers (and summarized in several reputation metrics), information about the seller's gender, ethnicity, or socioeconomic status is generally unavailable. Sellers use pseudonyms on the platform, making it difficult to infer these attributes from names. Furthermore, nearly all contact with sellers occurs via electronic communication. We believe, then, that any inferences buyers make about socioeconomic and cultural proximity or distance with sellers come from their beliefs about the expected characteristics of sellers at the designated location.

Specific features of the eBay setting also allow us to distinguish, at least partially, between overlap in buyer-seller inventory and demand versus buyer-seller trust as underlying drivers of trade. In particular, eBay’s product category structure, together with the detailed transactional data we employ, enables us to develop measures of congruence between the product preferences of consumers in a given geographic location and the products supplied by sellers in other locations. We may then use this measure to examine the extent to which the culture-trade relationship is mediated by similarity in “product taste” or, to be more specific, overlap between the inventory in one location and the purchases in another. Furthermore, because eBay has several well-established (and well-documented) mechanisms for identifying trustworthy sellers, we can explore how socioeconomic and cultural distance influence trade involving sellers that vary in their performance records, and hence the perceived riskiness of the transaction from the consumer’s perspective. These features allow us to partially disentangle the degree to which cultural or socioeconomic distance reflect a buyer-seller divergence in product preferences rather than perceived advantages from dealing with more trustworthy counterparties.

We employ eBay trade data at two levels of geographic aggregation in our analysis: state-pair and 3-digit zip-code-pair, aggregating pairwise trade to the year-level in each case.³ As proxies for cultural distance (i.e., differences in shared values and beliefs), we employ differences in state ethnic composition, voting behavior, and religion; we also account for differences in a range of socioeconomic characteristics, such as income, urban share, home values, and average age, as well as geographic distance. All specifications also include buyer-location and seller-location (whether state or zip code) fixed effects, by year, to account for differences in the scale of trade originating in and destined for particular regions.

For our state-pair data, we find that greater cultural distance – for all three “culture” measures – between two states is associated with significantly lower trade. Our estimates imply, for example, that a one standard deviation increase in ethnic similarity is associated with an 8 percent increase in state-pair

³ A 3-digit zip-code is the first 3 of 5 digits in a conventional U.S. zip code.

trade.⁵ Or to be more concrete, if the ethnic similarity of Oklahoma and New York (the 75th percentile of ethnic distance) were as similar as that of Massachusetts and Ohio (the 25th percentile distance), trade between these states would be 11 percent greater. (Other socioeconomic characteristics are also predictive of trade: our estimates suggest that if the gap in home values between Arkansas and Colorado were equal to that of Maryland and Washington state, for example, trade between Arkansas and Colorado would be 3 percent greater.)

While state-level attributes may be salient for prospective buyers, there is a great deal of within-state variation that may also be discernable to trading partners: the “cultural” characteristics of Philadelphia, PA, and Austin, TX, for example, differ markedly from their surrounding states. When we repeat our analysis at the 3-digit zip-code-pair level, the patterns are very similar to the state-pair results.⁶ This is true even when we focus on within-state trade only, which removes any cross-state variation from the analysis.

As we noted at the outset, these strong associations between socioeconomic and cultural variables and trade could be due to greater trust between buyers and sellers, a greater “inventory-preference overlap” between them, or a combination of the two.

To address the role of inventory-preference overlap between two geographies (pairs of states or 3-digit zip codes), we construct two measures that aim to capture the congruence between the products that eBay users in one area generally sell, and the products that users in another area generally buy. Focusing on state-level overlap, to construct these measures we calculate the cross-category distribution of goods sold and the cross-category distribution of goods purchased in each state. Our first measure is based on differences between buyer states’ and seller states’ shares of activity in each of 33 individual product categories. This distance measure decreases, for example, as the fraction of purchases in the electronics

⁵ The quantitative results we discuss here are from a classic log-linearized gravity equation, which we estimate with OLS. The results are qualitatively the same when we employ Pseudo Poisson Maximum Likelihood (PPML) estimation, as suggested by Silva and Tenreyro (2006).

⁶ As we discuss in more detail below, there are two potentially offsetting effects in moving to more disaggregated data. First, it may lead to less measurement error since it allows us to account for different buyer-level assessments of, say, Philadelphia versus Allentown. On the other hand, to the extent that state-level rather than zip-level variation is salient to consumers, we may be adding noise.

category in the buyer's state approaches the share of sales of electronics from the seller's state. The second measure is (one minus) the cross-product of the two category shares, which captures the likelihood that a randomly selected buyer from the buyer state is looking to purchase a product listed by a randomly selected seller from the seller state. We also calculate these inventory-preference overlap measures at the zip-pair level. The measures are related to Finger and Kreinin's (1979) similarity index, which aims to capture overlapping exports, whereas our indexes capture the congruence between consumers' purchases and sellers' offerings.⁷

We find that both of our inventory-preference overlap measures are predictive of state-pair trade. Furthermore, the inclusion of one or both overlap measures has a large effect on the estimated coefficients for our cultural and socioeconomic variables. For example, the coefficient on our measure of states' ethnic similarity falls by 38 percent when we include the category differences-based overlap measure. A second measure of cultural distance – vote share for the 2016 Presidential election winner – shrinks by more than half. Some of the reluctance to trade between dissimilar partners, therefore, is due to simple asymmetries in what sellers offer and what buyers demand. As an alternative way to account for compositional differences in what states buy and sell, we also conduct our analysis by examining state-to-state trade *within* individual product categories; the culture-trade relationship is attenuated only marginally in this analysis.

We next explore whether the correlation between sociocultural proximity and trade is the result of greater trust. We are motivated to do so by the large body of research on in-group favoritism dating back at least to Tajfel et al. (1971), and more recently work which shows that individuals place greater trust in, and are more apt to cooperate with, others with whom they share common traits (e.g., Foddy et al. 2009, Buchan

⁷ Guiso, Sapienza, and Zingales (2009) take a distinct approach to accounting for taste overlap that focuses instead on the demand side, by incorporating correlations between consumption patterns in buyer and seller locations, rather than correlating the offerings of sellers and the purchases of buyers.

et al. 2002, DeBruine 2002).^{8,9} Our test is based on the insight that, if social and cultural similarity leads to greater trust, then quality-assurance mechanisms – which serve as an alternative source of trust – may moderate the impact of cultural and socioeconomic similarity. We explore this possibility by examining whether the effect of cultural and socioeconomic variables on trade is different for sellers with eBay’s Top Rated Seller (eTRS) designation, which serves as an alternative signal of trustworthiness. Intuitively, we expect that trust between groups will matter less if there is a separate source of quality assurance (see Elfenbein, Fisman, and McManus 2012; 2015). Consistent with cultural similarity serving as a source of trust, we find that the correlations between trade and our cultural similarity variables (e.g., similarities in ethnicity and political affiliation) are substantially diminished for eTRS listings, whether we look at state-pair or zip-pair levels of aggregation. We find a similar pattern when we separate sellers by their level of buyer-provided feedback,¹⁰ which reinforces the notion that cultural similarity is especially important for trade when sellers do not yet have track records of previous success to reassure buyers of their reliability.

Finally, we examine the relationship between cultural similarity and buyer satisfaction with completed transactions. We measure satisfaction using two standard metrics: the fraction of all transactions that lead to positive feedback, and the fraction of feedback that is negative. While prior studies have documented discrimination in online sales (e.g., Ayres, Banaji, Jolls 2015, Doleac and Stein 2013), if the sort of “animus” between groups described by Becker (1957) were to play a major role in shaping trade patterns online, we might expect to see a negative association between favorable feedback and our measures

⁸ Guiso, Sapienza, and Zingales (2006: 29) argue that “the opening through which culture entered the economic discourse was the concept of trust” and conceptualize trust as a feature of culture. An alternative perspective is that trust, or cooperative behavior, may be easier to sustain among socially or culturally similar actors.

⁹ Research on trust games more generally is more ambiguous on the relationship between social distance and trust. Buchan et al. (2006), for example, find that this relationship depends on cultural context, with strong in-group effects in the U.S. but not in Asian countries, while Fershtman and Gneezy (2001) find that both Ashkenazi and Eastern Jews in Israel exhibit lower trust of Eastern Jews. Glaeser et al. (2000) find no statistically significant relationship between demographic similarity and trust.

¹⁰ The feedback score is one of the principal reputation mechanisms that eBay has used since its early days, and has featured in many studies; see Bajari and Hortacsu (2004) for survey. The literature has generally concluded that buyers see greater risk in dealing with sellers with low feedback scores, because their trustworthiness is more uncertain.

of cultural difference; however, we find no such association. Rather, we find that, conditional on trade, the relationship between our measures of buyer satisfaction and cultural similarity is not economically significant. We recognize, however, that this does not fully rule out the possibility that buyer selection adjusts trade volume in a precise way across state- or zip-code-pairs, so that consummated trades might capture only the instances in which buyers find sufficiently appealing or trustworthy sellers. Still, these results are not consistent with perceptions of seller trustworthiness (untrustworthiness) in culturally similar (different) locations being validated by executed transactions.¹¹

Our work contributes first and foremost to the literature that studies the determinants of online trade. The clearest antecedent is Hortacsu, Martinez-Jerez, and Douglas (2009), who study the effect of buyer-seller geographic distance on trade for eBay transactions, and also for a similar Latin American platform, Mercado Libre. Like us, they find a much smaller distance effect relative to gravity models estimated using trade in both intermediate and final goods (whether for cross-state or cross-country trade). They further provide an indication that tastes play some role for “local” preferences, by showing that the distance effect is particularly prominent for sports memorabilia and tickets. Our agenda is distinct from Hortacsu, Martinez-Jerez, and Douglas (2009) in that our aim is not to understand the geographic distance effect but rather how trade is affected by a broader set of similarities and differences between populations. A number of studies explore the roles of trust and/or quality certification, and generally emphasize *either* seller attributes (such as charitable contributions, as in Elfenbein, Fisman, and McManus, 2012) or buyer attributes (such as experience, as in Perez-Truglia, 2017) to understand the drivers of trust. By contrast, we emphasize the importance of the buyer-seller match, which implies a difference in (perceived) opportunism among culturally proximate trading partners. That is, our approach emphasizes the role of the buyer-seller match, rather than the attributes of the seller or attributes of the buyer.

¹¹ This is also consistent with the findings of DeBruine (2002), who finds that, while individuals exhibit greater trust in others with similar facial features to themselves, individuals are no less likely to betray facially similar partners.

Also closely related to our paper, Bailey et al. (2018) examine social media friendships on platforms such as Facebook, and they document that trade – online and otherwise – between counties correlates with measures of social connectedness. Our work explores similar socio-cultural determinants of commercial interactions, but allows us to explore overlapping product interests versus trust as distinct mechanisms.

Previous researchers have also considered the effects of cultural and social distance on trade across countries. Guiso, Sapienza, and Zingales (2009), in particular, study the link between survey-based measures of bilateral trust among European nations and economic activity such as trade and investment. They find that, particularly for “trust-sensitive” products, trust impacts trade flows. More broadly, an extensive literature documents that cross-border exchange of goods and services is greater between counterparties when they are closer together, whether in terms of physical, cultural, legal, or other “distance” (see Anderson and van Wincoop, 2004). While these cross-country patterns are helpful in guiding our understanding of the determinants of trade, as noted earlier, we are able to hold many more factors constant in looking at trade on a particular platform within a single country. One attempt at looking at the determinants of trade across regions within a single country comes from Hillberry and Hummels (2008), who look at the flow of goods within the U.S. They focus on understanding a very strong geographic distance effect, and their explanation centers on production co-location and trade in intermediate goods. In that sense, we see our focus on consumer goods and preferences as complementary to their work.¹²

2. Data

We use two broad types of data in our analysis. The first is sales data from eBay, at varying degrees of aggregation. These include data on trade between geographies at the state-to-state and zip-to-zip level, as well as state-to-state trade disaggregated by eBay’s 33 top-level categories. Throughout, when we refer to

¹² Our work also connects, albeit more indirectly, to a relatively new literature on habit formation, taste, and the geography of consumption, pioneered Bronnenberg, Dhar, and Dubé (2007) in a study showing that brand preferences are heavily influenced by the distance of consumers from where the company was founded. Bronnenberg, Dubé, and Gentzkow (2012) and Atkin (2013) extend this framework to show that consumers maintain loyalty to a brand (and in the case of Atkin, food type) even when they migrate. Both of these latter papers naturally have taste-based implications for the flow of goods within a country.

zip-code-level data, we refer to 3-digit zip-code areas (e.g., the area “275” includes zip codes 27514, 27516, and others), of which there are 887 in the U.S. with civilian populations and complete demographic data. We also use the distribution of U.S. buyer and seller transactions across product categories to construct our measures of taste overlap by state- and zip-code-pair. The second is state- and zip-code-level demographic data, drawn from a variety of sources we describe below, which capture cultural and socioeconomic characteristics of individual areas.

2.1 eBay data

Our eBay data come from the firm’s U.S. platform, which hosted over \$36 billion worth of transactions during 2017.¹³ eBay.com is the eleventh-most frequently visited website in the U.S., behind only Amazon among e-commerce platforms.¹⁴ eBay offers its users the opportunity to sell items in several formats, and it attracts sellers who vary widely in their engagement with the platform. Some sellers offer items for sale rarely, while other sellers are professionals who create dedicated “eBay Stores.” Seller quality and other attributes are tracked in a variety of ways. All sellers have at least a small amount of information visible to consumers, including a feedback score and the location (city and state) from which an item will ship. The feedback score reports the sum of individual-transaction feedback (+1, 0, or -1) that an eBay user has received as a seller on the platform. In addition, a small fraction of sellers (who represent a disproportionately large proportion of sales volume) earn an eTRS badge, which indicates that they have cleared specified thresholds for sales volume and customer satisfaction. These reputation mechanisms have been studied extensively in the economics literature. Cabral and Hortascu (2010), Hui, Saeedi, Shen, and Sundaresan (2016), and Elfenbein, Fisman, and McManus (2015), provide empirical evidence of the impact of these mechanisms on sales probability and price, while Hui, Saeedi, Spagnolo, and Tadelis (2018)

¹³ eBay US platform revenue from <https://www.digitalcommerce360.com/article/ebays-sales/>.

¹⁴ Web traffic statistics are from the marketing firm Alexa (an Amazon.com subsidiary), accessed on May 10, 2019.

examine the mechanisms' impact on seller entry, and Nosko and Tadelis (2015) study how buyers learn from experience about the overall quality to be expected on the platform.

When an eBay consumer searches for a product, the platform provides the user with a collection of listings sorted by several factors, including product characteristics, price, sellers' quality ratings, and sellers' physical proximity to the consumer. The precise algorithm used to return search results is proprietary and ever-evolving. The consumer may elect to re-sort the listings based on a single factor, including proximity. (Other than geography, consumers cannot filter or sort listings based on the difference measures we use in our analysis below.) eBay sellers can affect their positions in search results through their choices of prices or shipping fees, but they are largely unable to affect the geographic distribution of consumers who evaluate their products. eBay consumers, therefore, have ample opportunities to inspect products from all over the United States, and interstate trade patterns are likely to reflect consumers' *preferences* over trading partners rather than *awareness* of partners, which can be an important driver of geographic trade patterns in other contexts.

We collect data on eBay transactions in which both the buyer and seller identify themselves as located in the U.S. In addition, we limit the population of sellers to those who do not operate eBay Stores. We apply this filter on sellers for two reasons. First, sellers with eBay Stores may ship from warehouses that are not in the same zip code or indeed the same state as the seller, which creates uncertainty about how to classify the seller's location. Second, we conjecture that buyers are more likely to use location-based characteristics to infer seller attributes when the seller provides only the sparse personal information offered in eBay's standard format, whereas sellers with Stores often use the interface to provide additional information about themselves. While the subset of sellers we study are likely less professional than those with eBay Stores, many sellers in our sample have earned eTRS status or have relatively large feedback counts. We discuss below the characteristics of the sellers in our sample.

2.1.1 Measures of inter-state trade and customer satisfaction

Our primary eBay data comprise a comprehensive record of transactions that occurred between January 2015 and December 2016, aggregated to the year level. In the description that follows we provide an overview of the state-pair data; an identical process is used to generate the zip-to-zip data, just at a finer level of disaggregation.

For the U.S.-based buyers and sellers described in the preceding section, we observe the total quantity of items sold and total dollar revenue from product sales, excluding shipping fees, between each pair of U.S. states; we also observe buyer and seller transactions for Washington D.C., which we treat as a separate (51st) state in our analysis. In addition to the aggregate annual transactions, we observe transactions categorized according to the sellers' eTRS statuses (badged or not) and whether their feedback scores were above or below 200 at the time of the transaction.¹⁵ We define *High feedback* to denote a seller that has feedback of at least 200. This is a relatively low feedback threshold, as we aim to distinguish the sample split based on feedback from the split based on eTRS.¹⁶ Roughly half of all sales revenue (48%) and transactions (45%) associated with feedback above 200 are from sellers without eTRS certification. Thus, these two different sample splits provide somewhat correlated, but not identical, tests.

Finally, we analyze feedback provided by buyers, aggregated to the state-pair-year level. Our main measure is Effective Percent Positive (EPP_{bst}), the fraction of total transactions between buyers in state b and sellers in state s during year t that leads to positive feedback. This measure is proposed by Nosko and Tadelis (2015) to deal with the fact that, conditional on feedback being provided, it is almost always (99.3%) positive. They show that there is information on seller reliability in the fraction of buyers that provide any feedback (which averages approximately 65% in their sample). We use the fraction of transactions with

¹⁵ For example, we may observe that, during 2016, Arizona sellers sold 2000 items for \$20,000 (in total) to Missouri buyers. Within those 2000 items, 800 with a value of \$7000 were from eTRS sellers, and 1000 items with a value of \$14,000 were from sellers with feedback greater than 200.

¹⁶ We have also performed the analysis after separating sellers by whether their feedback is greater than 1000. Our results are very similar to those reported below for a feedback threshold of 200.

negative feedback (*Negative feedback_{bst}*), conditional on feedback being provided, as an alternative measure of buyer (dis-)satisfaction.

In preparing the data, we exclude observations on trade that occurs within states because these cases have zero “distance” in many of the measures that we introduce below. We provide summary statistics on state-to-state annual trade in Table 1. Trade between states (exclusive of shipping) has a mean value of \$4,381,590 per year in total for 109,447 items (medians of \$1,342,110 and 35,342, respectively), summing to nearly \$22 billion in sales on over 550 million items during the two-year period we study. In our sample, which conditions on less-professional sellers by excluding data from eBay Stores, state-level sales by non-eTRS sellers averages 65% more than that of eTRS sellers; the quantity sold by non-eTRS sellers is 24% greater. The total sales by low-feedback sellers is about half of that associated with sellers with feedback above 200. The average EPP_{bst} value across state dyad-year combinations is just over one half, and the mean rate of negative feedback is 0.36%.¹⁷

As expected, pairs of large states have the greatest transaction volume. In 2016, the top 162 state dyads in total revenue involve buyers or sellers from California, Florida, New York, or Texas; the top 10 is almost entirely comprised of these states trading with each other (with one appearance each by New Jersey and Illinois). States with low trade volumes include Alaska, the Dakotas, Washington D.C., Rhode Island, Vermont, and Wyoming.

Our process for generating the zip-to-zip annual transaction dataset is virtually identical, and we include summary statistics for this dataset in Appendix Table A1. As expected, average transaction volume is much lower at the zip-to-zip level.

¹⁷ The difference between our mean EPP value and the 65% rate reported by Nosko and Tadelis (2015) may be due to differences in our respective samples’ seller populations; they take a platform-wide average that includes professional sellers that we largely omit.

2.1.2 Measures of geographic differences in buyer and seller activity across product categories

We supplement the trade flow data with additional information from eBay on the distribution of transactions across 33 high-level product categories (e.g., Consumer Electronics, Collectables and Art, Home and Garden, etc.). We once again focus on state-pairs in describing the construction of our data; however, the approach extends straightforwardly to measuring zip-pair inventory-preference overlap.

For each combination of buyer state (b), seller state (s), and product category (c), we obtain the total quantity of items sold and the associated revenue during 2015-16. Let $q_{b,s,c}$ represent the quantity of items sold in a (b, s, c) combination. For each b and s pair, we calculate:

$$q_{b,-s,c} = \sum_{s' \in S \setminus s} q_{b,s',c} \quad \text{and}$$

$$q_{b,-s} = \sum_c q_{b,-s,c}.$$

The notation $S \setminus s$ indicates the set of all states excluding state s . We use these objects to calculate

$$\sigma_{b,-s,c}^q = \frac{q_{b,-s,c}}{q_{b,-s}}$$

as the share of b 's quantities that is in category c , excluding any trade between b and s . The vector $\sigma_{b,-s}^q$ contains $\sigma_{b,-s,c}^q$ values for each category. For an individual buyer state we obtain 50 versions of $\sigma_{b,-s}^q$ (one for each $s \neq b$, including Washington D.C.), but they are all quite similar to each other because removing a single trading-partner seller-state (which we do to avoid a mechanical correlation between our taste overlap measures and realized buyer-seller state transactions) generally has a small impact on the distribution of buyer activity across categories within a state. We produce analogous measures for three additional variables: buyer-state revenue (r) across categories, which generates $\sigma_{b,-s}^r$ values, and for quantities and revenues while focusing on the distribution of items sold from state s . Let the seller-focused share vectors be $\sigma_{s,-b}^q$ and $\sigma_{s,-b}^r$.

We use these distributions to construct inventory-preference difference measures that aim to capture the divergence between states in the products they buy and sell. For each pair (b, s) and each

transaction variable ($j = q$ or r), we compute a “cross-product category difference” measure, P , as one minus the cross-product of category shares:

$$P_{b,s}^j = 1 - \sum_c \sigma_{b,-s,c}^j \sigma_{s,-b,c}^j.$$

$P_{b,s}^j$ is similar to a Herfindahl index in which (b, s) pairs replace squared market shares; we subtract the summed products from one so that it is increasing in distance. Intuitively, this measure captures the likelihood that, within a buyer-seller state dyad, a buyer, chosen at random, would be searching in the category that a seller, also chosen at random, had decided to offer a product. When the buyer and seller states both have large fractions of their activity in different categories (i.e., one state tends to sell a lot in categories that are little demanded by the other state), $P_{b,s}^j$ will be relatively large. In the regression analysis below, we rescale the measure to have mean zero and standard deviation one, to make its estimated impact easily comparable to other variables.

We use a Euclidean distance calculation to develop an alternative measure of the differences between items that are sold in a seller state and items that are purchased in a buyer state. The variable D captures “Euclidean category distance,” and we compute it as

$$D_{b,s}^j = \sqrt{\sum_c (\sigma_{b,-s,c}^j - \sigma_{s,-b,c}^j)^2}.$$

As in the case of $P_{b,s}^j$, we calculate $D_{b,s}^j$ at the quantity ($j = q$) and revenue ($j = r$) level, and in our analysis we normalize the variable so that it has mean of zero and standard deviation of one. The measures $P_{b,s}^q, P_{b,s}^r, D_{b,s}^q$ and $D_{b,s}^r$ capture different aspects of trade and handle between-state differences in distinct manners, particularly in how they are affected by buyer and seller category dispersion. A buyer-seller pair with high, but similar, dispersion across categories will have a low value for the Euclidean category distance but a large value for the cross-product category difference. Thus, while both measures capture some intuitive element of overlapping category interests, in some cases the two measures may be negatively correlated. In

practice, however, once we control for differences in average levels of $P_{b,s}^j$ and $D_{b,s}^j$ across states, the correlation between the two measures is strongly positive.

2.2 Demographic data

The underlying data for calculating group-pair similarity comes from several datasets which we query at the state and county levels. We use the state-level demographic data directly in our analysis of between-state trade, and we use the county-level data to construct zip-code level demographic measures using a weighting procedure described below.

We draw data on *Median household income*, *Share with bachelor's degrees and above*, *Median age*, *Share of males*, *Share in urban areas*, *Home ownership share*, and *Median home value* from the 2017 American Community Survey, which is organized by the US Census. The ACS interviews a representative sample of over two million Americans each year, providing a high-quality source of data with broad geographic coverage. We measure similarity in political attitudes using voting patterns from the 2016 U.S. presidential election, taken from the Federal Election Commission. We use the numbers of votes cast in each state or county for the two major parties to calculate the winning candidate's fraction of these votes at the state and zip-code level; we label this variable *Winner vote share*.¹⁸

Table 2 reports summary statistics for the state-level aggregations of these variables, which we use to create measures that capture differences in socioeconomic and cultural characteristics across states. In Table 3 we report the cross-state differences of each variable, which is simply the absolute value of the difference between the buyer and seller state values.

¹⁸ We focus on 2016 vote shares to measure political preferences because it is based on real-stakes decisions during the same period covered by our eBay transactions data. For robustness, we considered two alternative measures of between-state differences in political preferences. The first is based on the winner vote share in 2012. The second is constructed using survey responses from a 2015 Gallup Poll of 175,000 adult Americans on their political views. The raw correlations among our three state-level measures of political preferences are each above 0.93; if we use 2012 election data or Gallup survey responses to measure political differences across states, our results are qualitatively similar to those based on the most recent election data.

We report three additional measures at the state dyad level in Table 3: *Distance*, and *Ethnic difference*, and *Religious difference*. *Distance* is the “shortest curve” distance (in kilometers) between the state population centers, which we obtain from the US Census.¹⁹ For *Ethnic difference*, we use the county-level ACS data from 2017 on ethnicity, aggregated to the state level to construct a measure of ethnic similarity between two states. The variable we construct is analogous to a between-group measure of ethnic fractionalization (see, e.g., Alesina et al. 2003) which aims to capture the likelihood that a person chosen at random from the seller state is in the same ethnicity category as a person chosen at random from the buyer state. Let S_E^i be the share of individuals in the ethnic category E of state i , which is drawn from a set of mutual exclusive and collectively exhaustive categories provided by the ACS.²⁰ We then calculate the ethnic dissimilarity between seller state s and buyer state b as $Ethnic\ difference_{bs} = -\sum_{all\ E} S_E^b S_E^s$. A higher level of this measure indicates a greater likelihood that two people drawn at random from each state are of different ethnicities. *Religious difference* is calculated similarly, using data on religious affiliations from the 2010 U.S. Religion Census, which provides the number of adherents for 236 distinct faith groups by state and county. We aggregate adherents of distinct faiths into five main categories, with the residual state population categorized as non-adherents.²¹

¹⁹ Alternative distance measures yield near-identical results, which is unsurprising given the very high correlation across such measures. For example, the correlation between the log of the distance between capitals and the log of the distance between population centers is 0.995.

²⁰ We use the shares of non-Hispanic individuals who are white, African-American, Asian, Pacific Islander, other (single) race, and two or more races; in addition we use the share who are Hispanic or Latino.

²¹ The aggregate categories are Evangelical Protestant, Black Protestant, Mainline Protestant, Catholic, and Other. As an alternative state-level measure of religious similarity, we may use the results of a survey conducted by Pew Center’s US Religious Landscape Center Research, which was reported in 2016. See <http://www.pewresearch.org/fact-tank/2016/02/29/how-religious-is-your-state> [accessed 2/23/2018]. Pew reports data on the percentage of adults in a state who say that religion is very important in their lives, the percentage who say they pray daily, the percentage who say they attend worship services at least weekly, and the percentage who say they believe in God with absolute certainty. This would be closest to using simply adherents versus non-adherents from the Religion Census. For our state-level analyses, using the Pew-based measure yields quite similar results.

The summary statistics in Table 3 are raw difference measures. However, to make the effect sizes more easily comparable across distance measures in our empirical analysis in the next section, we normalize all cultural and demographic distance variables to have a mean of zero and standard deviation of one.

To construct zip-code-level values of the demographic variables, we take a weighted average of all counties that contribute to the 3-digit zip code. We know the share of each county's land area that is within each zip code, and we assume that a county's population is distributed uniformly within its boundaries. For each county, we obtain the fraction of its population that is in zip code z , and calculate the county's weight with the fraction of z 's population that comes from the county.²² We construct our zip code demographic variables by using these weights with the county-level versions of the demographic variables described above.²³ In Appendix Tables A2 and A3 we present summary statistics for zip-code-level and zip-code-pair attributes.

3. Empirical Analysis and Results

Our main state-level specification takes the form:

$$\log(Trade_{bst}) = \Gamma C_{bs} + \Theta X_{bs} + \Upsilon T_{bs} + \beta \log(Distance_{bs}) + u_{bt} + v_{st} + e_{bst} \quad (1)$$

where C is a vector of cultural similarity variables, X a vector of socioeconomic similarity variables, T contains our inventory-preference overlap variables (P and D), $\log(Distance)$ is logarithm of the distance between the two states' population centers, and u and v are buyer-state \times year and seller-state \times year fixed effects, respectively. Throughout, we focus our discussion on results using a revenue-based measure of annual $Trade_{bst}$ between seller state s and buyer state b during year t . For robustness, we also present

²² For example, zip code z may contain 100% of 80,000-person county A, 20% of 100,000-person county B and 60% of 200,000 person county C. We conclude that z 's total population is 220,000 ($=80,000+0.2\times 100,000+0.6\times 200,000$), and the counties' weights in z are their population shares of z 's total population, e.g. 0.36 for county A.

²³ This procedure will introduce some measurement error in the zip-code demographic variables, but we believe it is minor. Potential sources of error include: the assumption that county population is uniformly distributed within its geographic boundaries, our use of weighted averages to calculate county-level median ages and house values, and the construction of a zip code geographic center using the average of Census-reported county population centers.

findings based on quantity-based measures, and also with Poisson pseudo-maximum likelihood (PPML) specifications rather than OLS, as suggested by Silva and Tenreyro (2006), which use the levels of trade and transactions as the dependent variables, rather than log transformations of them.²⁴

As described in Section 2, the set of variables we use to capture cultural/attitudinal factors include differences in ethnic and religious composition as well as the pairwise absolute difference in the fraction of each state’s vote share in the 2016 Presidential election. The set of socioeconomic variables we use are (absolute) differences in median income, share with a bachelor’s degree, median age, male share, median home value, and shares of urban residents and owner-occupied housing. All covariates are normalized so that the mean difference between states is 0 and the standard deviation is 1. We use two-way clustering by seller state and buyer state to calculate the standard errors. For our zip-code-level analyses, the specification is virtually identical to (1), with 3-digit zip codes defining b and s throughout.

3.1 Impact of cultural and socioeconomic differences on trade

We begin by presenting our baseline state-pair trade findings in Table 4, for both revenue- and quantity-based measures of trade. Throughout, we present our state-to-state trade results in the main text and zip-to-zip findings in a set of appendix tables.

In the first three columns, we present results that use the log of total sales revenue as our measure of trade. In columns 1 and 2 we include the “culture” variables and the socioeconomic variables separately, and we include both groups of covariates in column 3. Focusing first on the set of measures that reflect cultural differences, we find in column 1 that all three “distance” measures are predictive of trade, in the

²⁴ The PPML specification with buyer-location and seller-location fixed effects can be advantageous in handling zeros, heteroscedasticity, and in solving the “adding up” problem in trade equations. Our dataset has very few zeros. In the state-to-state trade data there are no instances in which there is 0 trade volume in a given year, and at the 3-digit-zip level of analysis only 1.2% of observations are equal to zero. The latter two issues, however, are applicable to the present work, so we include PPML specifications in each of our reported estimates of (1).

expected directions: the coefficients on the ethnic, political, and religious distance measures are all negative and significant at least at the 5 percent level.

Given that these variables are all normalized to have a standard deviation of one, their coefficients are easily interpreted and compared. By far the biggest effect comes from ethnic difference: a one standard deviation increase in ethnic difference is associated with an 11 percent decrease in state-pair trade. The coefficients on religious and political differences imply sizeable but more modest effects.

Turning to the socioeconomic variables, their role in predicting state-pair trade is mixed. While similarity in urban share is predictive of trade in column 2, the inclusion of culture variables reduces the size of its coefficient by more than half. Four of the socioeconomic variables remain significant predictors of trade in column 3: differences in median age, male share, urban share, and home values. (While we group these variables with other demographic measures, there are no clear guidelines on whether buyers view particular measures as reflecting cultural versus other social differences.) None of these results appear to be driven by outlier observations – we observe similar patterns if we omit Alaska and/or Hawaii (both are obvious geographic outliers, and Alaska is also an outlier in male share).

We note that, as expected, all regressions produce coefficients on $\log(\textit{Distance})$ that are negative, economically important, and precisely estimated. The estimated coefficients that we find are larger, but similar in order of magnitude, to the estimates produced by Hortacsu, Martinez-Jerez, and Douglas (2009) for eBay trade. (By construction, however, our analysis differs from Hortacsu et al. (2009) as we exclude within-state (or within-3-digit zip code) transactions, which Hortacsu et al. (2009) show are particularly important for certain categories of transactions, like event tickets).

In column 4 we present results from a specification that employs all culture and socioeconomic difference variables, using the logarithm of annual sales quantity as the dependent variable. The patterns are broadly similar though in some cases marginally weaker than the results that measure trade in terms of total transaction value. Finally, in columns 5 and 6 we present PPML specifications for state-pair annual transaction revenue and quantity, respectively, using the full set of covariates. Most of the PPML models' point estimates are slightly smaller in magnitude than the corresponding coefficients in the OLS models

(including the geographic distance elasticity), but the results are broadly similar across estimation approaches.

In Appendix Table A4, we show analogous results for zip-to-zip annual transaction data, which allows us to account for the very different perceptions that buyers may have of, say, Manhattan versus upstate New York, or Austin versus rural Texas. The first three columns exactly parallel the specifications in columns 1-3 of Table 4, and similarly show a consistent and robust relationship between culture proxies and trade at the zip-code-pair level. In general, the magnitudes of the coefficients on *Religious difference* and *Voting Difference* are larger, and all variables are (unsurprisingly) more precisely estimated than in the state-to-state trade analyses. These specifications exploit both cross- and within-state variation in estimating the culture-trade relationship. In column 4, we limit the sample to within-state zip code pairings to focus on variation that is distinct from the between-state results. The patterns are quite similar to those based on the full set of zip-code-pairs. Finally, we show results based on transaction quantities (column 5) and using PPML estimation (columns 6 and 7); the patterns are again quite similar to (and more precisely estimated than) those based on state-pair data.

In a set of figures, we next illustrate the culture-trade relationship graphically, using binned scatter plots (Hao et al., 2010) to display the relationship between the three main proxies for culture and trade. Focusing first on state-pair trade, to capture the role of all three culture variables simultaneously, we use the sum of (normalized) state-pair ethnic, religion, and voting differences. We group the data into 50 bins with equal numbers of observations, though in practice the patterns are essentially the same if we use coarser (e.g., 25 bins) or finer (e.g., 100 bins) groupings. Finally, we residualize the data, netting out the effects of all socioeconomic and control variables included in our main specification (Table 4, column 3), including buyer-state \times year and seller-state \times year fixed effects, u_{bt} and v_{st} , respectively.

We present the resulting binned scatter plot in Figure 1, which shows a clear negative correlation between our summary measure of state-pair cultural distance and interstate trade.²⁵ The relationship is roughly linear, and it is in line with the estimates presented in Table 4. In Appendix Figures A1 – A3, we present binned scatterplots for each individual cultural difference variable, generated while including all other controls from Table 4, column 3 (e.g., in Figure A1, when we illustrate the role of ethnicity differences, we control for differences in religion and voting). As expected, given the more prominent role of ethnicity in our regression results, the plotted pattern is clearest and the best-fit slope is steepest in Figure A1, which uses ethnicity to measure cultural differences. In Appendix Figures A4 – A6 we show analogous figures to illustrate the zip-code-level relationships between our culture variables and trade. We once again use 50 bins to plot the relationship. Given the vast number of observations underlying each data point (over 30,000), it is perhaps unsurprising that the graphs show a tight fit around the regression line; our main takeaway from these figures is that the log-linear model is again a good fit for the data.

3.2 Accounting for inventory-preference overlap between buyer-seller groups

As noted from the outset, there are two primary reasons we might expect culturally or socioeconomically similar locations to trade more: greater trust and more overlap between sellers’ inventories and buyers’ demands. We now turn to exploring the extent to which these factors account for the correlation between state-pair and zip-code-pair similarity and trade.

Focusing first on state-pair results, in Table 5, we present results that build on those in Table 4, including one or both of our inventory-preference overlap measures as controls. Intuitively, if the inclusion of these direct overlap measures attenuates the effects of cultural or socioeconomic closeness, then we may infer that the initial correlation was capturing, at least in part, a concordance between sellers’ goods and buyers’ preferences. Focusing on total annual state-pair sales revenue as our measure of trade, we explore

²⁵ The scatter plot is quite similar if we include only the fixed effects u_{bt} and v_{st} , as suggested by the stability of the coefficients across columns 1 and 3 of Table 4.

how the inclusion of our direct measures of Euclidean category distance (D) and/or cross-product category differences (P) affect the estimates in Table 4. We begin in column 1 by showing the results from Table 4, Column 3, for comparison. We add Euclidean category distance as a covariate in column 2. Our category difference measure is itself highly predictive of trade. A one standard deviation increase in Euclidean category distance is associated with a 27 percent reduction in state-pair trade. Moreover, the inclusion of category differences has a very large effect on the remaining predictive power of our cultural variables: the coefficient on *Ethnic difference* falls by more than 35 percent and the coefficient on *Voting difference* falls by more than half. One exception is the coefficient on *Religion difference*, which increases marginally in absolute value. The coefficient estimates for most of the socioeconomic variables are also attenuated. Column 3 uses our cross-product category difference measure (P) to capture state-pair differences in sellers' inventories versus buyers' preferences. The coefficients on our cultural difference variables are again attenuated, although to a more modest degree in the case of *Ethnic difference*. In column 4 we include both inventory-preference difference measures as controls. Interestingly, while the coefficient on each is smaller than its counterpart in columns 2 and 3, both measures remain significant at the 1 percent level, suggesting that these two measures effectively capture different types of variation in inventory-preference overlap. Again, most coefficients on our cultural similarity variables are markedly lower than in column 1. Finally, in column 5 we show a PPML specification with the full set of covariates. Comparing these results to the PPML estimates in Table 4, column 5, confirms that the effect of introducing inventory-preference overlap measures is similar between OLS and PPML approaches to estimation.

In Appendix Tables A5 and A6, we repeat the analysis reported on Table 5, but using alternative dependent variables (quantity rather than revenue) and a different approach to geographical boundaries (zip codes rather than states). The results are broadly similar to those presented in Table 5, with two exceptions: the attenuation of the *Voting difference* coefficients in the quantity regressions is much smaller, and the attenuation of the *Religious difference* coefficients is about one-fifth in quantity-based regressions, whereas the *Religious difference* coefficients increased modestly in the revenue-based specifications. In Table A5

and A6, all of our sociocultural coefficients remain statistically significant when inventory-preference overlap controls are added.

An alternative approach to accounting for differences in sellers' inventories and buyers' preferences is to disaggregate our data by category. This allows us to control for category-level differences in inventory-preference overlap with fixed effects at the level of buyer-state \times category \times year and seller-state \times category \times year, rather than the constructed variables D and P . If D and P fail to capture how inventory-preference overlap plays a primary role in driving the link between state-pair cultural differences and trade, we would expect that within-category analyses could generate markedly different results. Furthermore, if differences in trade patterns were *entirely* based upon differences in category-level preferences across locations (e.g., Texans have a greater propensity to shop for clothing and Arizonans sell a greater proportion of clothing), then the fixed effects should fully capture these differences, and the estimated coefficients for sociocultural and socioeconomic similarity should be much reduced. We report these results in Appendix Table A7. The patterns are quite similar to those in Tables 4 and 5, suggesting that while category-level tastes partially explain the relationship between trade patterns and cultural similarity, they do not fully explain it.²⁶

3.3 Quality assurance and the impact of cultural differences on trade

We next explore whether buyers deem sellers from places that are more culturally or socioeconomically different to be less trustworthy, and hence the buyers trade less with them. Since we lack a direct measure of state-pair trust (the approach utilized by Guiso, Sapienza, and Zingales (2009) in studying European trade flows), we use a different empirical approach: we investigate how trade patterns are different with and without additional mechanisms that build trust. We first analyze whether trade volume varies with whether the seller has earned a certification label that may substitute for trustworthiness, namely the eTRS

²⁶ The coefficients in Table A7 are not directly comparable to the corresponding estimates in the state-level aggregate regressions in Tables 4 and 5, as they represent the (unweighted) average effect of similarity across 33 different-sized categories.

designation. Intuitively, little trust is required on the buyer's part to purchase from a seller that has an excellent, certified track record for quality and promptness. Similarly, we posit that questions of trustworthiness are more salient for evaluating sellers with little feedback, who may be new to the platform and have a limited observable track record. To test whether cultural similarity matters more when dealing with these less-established sellers, we allow for different impacts depending on whether sellers have feedback scores greater or (weakly) less than 200 at the time of the transaction.

Table 6 contains results from two models based on our state-pair data. For the sake of comparison, we reproduce Table 5 column 4 as column 1 in Table 6. In the next pair of columns (labeled 2a and 2b), we report results from a single model that includes two observations per year and state-pair: one for aggregate sales by non-eTRS sellers and the other for aggregate sales from eTRS sellers. We allow measures of socioeconomic differences, inventory-preference overlap, and geographic distance to affect each type of seller's transactions differently, and we implement this with a full set of interactions with an indicator (*eTRS*) for whether trade involves an eTRS seller.²⁷ Table 6, column 2a shows the coefficient estimates for the direct effects of all state-pair variables, while column 2b shows the coefficient estimates for all variables' interactions with *eTRS*. Thus, the coefficients in column 2a show the effects of the state-pair differences when sellers lack eTRS certification, which forces buyers to depend on other information or characteristics to infer seller reliability. The coefficients in column 2b show the incremental effect for *eTRS* = 1 transactions, so that the total effect for *eTRS* = 1 transactions is the sum of the coefficients in columns 2a and 2b. Columns 3a and 3b repeat this exercise, using *High feedback* rather than *eTRS* to divide sellers by their performance records.

While the estimates of the interaction terms in column 2b are mixed in sign, for our cultural difference measures for two cases (ethnic and voting differences) they go in the direction of implying that

²⁷ We also interact each state-year fixed effect with the eTRS indicator to account for different overall trade volumes flowing out of and into states by eTRS status. For example, if California consumers, on average, prefer eTRS items regardless of the item's origin, then this tendency is captured by the additional fixed effect for California buyers by eTRS status. Similarly, if Florida has a greater than average fraction of eTRS sellers shipping to all destinations, then the additional eTRS-specific seller-state fixed effect will account for this pattern.

these differences matter less when buyers consider purchases from eTRS-badged sellers; the religion difference interaction goes in the opposite direction, but does not approach statistical significance. As an intuitive test of whether the three variables collectively differ for eTRS versus non-eTRS sellers, we provide an F-test of the sum of their coefficients, which has a p -value of 0.051. (An F-test of the joint significance of the three separate coefficients has a p -value of 0.049). The eTRS interactions with the socioeconomic variables are of inconsistent sign and an F-test of their sum has a p -value of 0.224. Together, these estimates suggest that cultural differences have a greater dampening effect on trade between buyers and sellers who have not earned quality certification, relative to the effect of cultural differences on trade between buyers and sellers with quality certification.

The results in columns 3a and 3b corroborate those in 2a and 2b. In particular, the joint effect of cultural differences is significantly smaller for *High feedback* sellers: an F-test of the summed coefficients has a p -value < 0.001 . Comparing the coefficients across columns 3a and 3b, our estimates suggest that the effect of ethnic differences is more than halved for high feedback (relative to low feedback) sellers, while the effect of voting differences disappears entirely. The estimates in 3a and 3b also show similar patterns for socioeconomic differences as those in columns 2a and 2b, with the effect of these variables of inconsistent sign in both columns. Given that our data do not contain (near-) experimental variation in eTRS or feedback, we interpret these variables as capturing the effects of a range of markers or achievements that distinguish well-established sellers.

Collectively, we interpret the estimates in Table 6 as suggesting that cultural similarity supports trust, so that cultural distance becomes a greater determinant of trade when quality certification or extensive seller reputations are absent. This finding echoes that of Guiso, Sapienza, and Zingales (2009), who find that bilateral trust is a stronger predictor of trade for quality-differentiated products relative to commodities.

As with our previous results, we repeat the specifications here using quantity-based trade data, zip-code-pair data, and PPML estimation. All of these sets of analyses show that cultural variables' impact on

trade is consistently attenuated for eTRS-badged and high feedback sellers.²⁸ Furthermore, in each of these models the coefficient on the interaction term with *Religious difference* is positive, consistent with *Religious difference* mattering less when alternative signals of trustworthiness are available. These results are in Appendix Tables A8 – A10.

3.4 Impact of cultural and socioeconomic differences on buyer satisfaction

In our final set of results, we examine whether, conditional on a transaction having taken place, cultural distance predicts buyer satisfaction. One possibility raised by the literature on economic discrimination, but difficult to explore directly in patterns of trade conducted above, is that buyers simply have preferences to do business with people like themselves, or conversely have an “animus” (Becker, 1957) toward interacting with others unlike themselves that is distinct from expectations about quality or trustworthiness. While both animus and perceptions of untrustworthiness may reduce the likelihood of transactions between dissimilar parties, conditional on the transaction occurring, buyers should be less inclined to provide positive feedback, and more inclined to provide negative feedback, to sellers towards whom they feel animus. We posit that, at least in some instances, these transactions would occur due to favorable product matches or idiosyncratically low prices, despite animus that a buyer may feel. Nonetheless, we recognize that those with the greatest animus may not engage in trade at all with dissimilar partners. Thus, our examination of customer feedback only rule *in* animus, rather than fully ruling it out

In Table 7 we provide results based on a variant of specification (1), in which we use effective percent positive (*EPP*) and fraction negative feedback (*Negative feedback*) as measures of buyer satisfaction. Geographic distance between states is negatively and significantly related to *EPP*, and positively and significantly related to *Negative feedback*; which may reflect shipping-related issues that increase with distance, such as damage or delays. In column (1), which uses *EPP* as a satisfaction measure,

²⁸ We do not have data on zip-code level trade by sellers with feedback scores above versus below 200, so at the zip-code level we divide sellers by eTRS status only.

none of the cultural or economic similarity measures are statistically significant. We obtain similar results using *Negative feedback* as an outcome measure in column (2) – though *Voting difference* is in this case a significant predictor of negative feedback. A one standard deviation increase in *Voting difference* is associated with a .01 percentage point rise in negative feedback, a small impact given the average negative feedback rate of 0.34 percent across the sample. We repeat the analysis with 3-digit zip codes and report the results in Table A11. The geographic distance measures in these regressions continue to have a strong association (both in economically and statistically) with dissatisfaction. By contrast, although the estimated coefficients on cultural and economic similarity in these regressions are statistically significant in several instances – sometimes in the opposite direction of what we would predict from animus-based models – they remain economically very small. In summary, our results on feedback are hard to reconcile with animus-based models.

Considering our feedback results through the lens of perceptions of trustworthiness suggests several possible interpretations. One potential explanation (incorporating both the trade and satisfaction results) is that, while buyers' perceptions of trustworthiness are associated with cultural similarities or differences, these perceptions reflect excessively strong generalizations (e.g., Bordalo et al. 2016) which are not validated in practice. According to this interpretation, perceptions shape the propensity to transact, but do not reflect actual performance conditional on the transaction being executed. A second explanation is that selection drives the observed satisfaction results. Buyers choose to interact only with sellers that they deem sufficiently trustworthy, and make accurate inferences of trustworthiness based upon cultural similarities or differences. Observed transactions then lead to levels of customer satisfaction that are independent of buyer-seller cultural similarity. This possibility raises the question of whether sellers do, in fact, behave in a more trustworthy manner when interacting with buyers from states with culturally similar populations; this question is beyond the scope of the current paper.

4. Conclusions

Our paper provides three contributions. First, we document that cultural distance – as measured by differences in ethnicity, religiosity, and voting behavior – is negatively associated with trade patterns in an online market in which buyers choose among sellers with observable locations but whose personal identities are effectively concealed. These differences are robust to controlling for socioeconomic differences and measures of inventory-preference overlap. This, in and of itself, is somewhat surprising as buyers and sellers in this marketplace do not meet directly, see pictures of one another, or engage in voice communication. Our second contribution is to show that socio-cultural similarity matters more for trade when there is more uncertainty about the quality of the seller, suggesting that buyers associate socio-cultural similarity with greater trustworthiness. In examining customer satisfaction with executed transactions, we raise the question of whether these perceptions of trustworthiness are accurate or driven by mistaken beliefs about the impact of buyer-seller differences. Collectively these findings suggest that cultural similarity drives trade both because of greater inventory-preference overlap and higher trust. These findings contribute to a small but growing body of work that examines the relationship between socio-cultural similarity and trade, such as Guiso, Sapienza, and Zingales (2009) and Bailey et al. (2018). Finally, we see our inventory-preference overlap measures, which explain a significant share of interstate trade on eBay, as a methodological contribution that builds on related similarity indices, which could be applied to cross- or within-country trade models.

The fact that we find the trade patterns exist on an e-commerce platform in which face-to-face contact is rare and demographic information about counterparties is concealed is, perhaps, surprising. While discrimination based on observable buyer or seller ethnicity has been documented on e-commerce platforms such as Airbnb (Cui, Li, and Zhang 2016; Edelman, Luca, and Svirsky 2017), on online classified advertisement platforms (Doleac and Stein 2013), and online used-car markets (Zussman 2013), features of the eBay platform make it nearly impossible to discover the ethnicity or other demographic details of a seller. Buyers in our setting must, therefore, make inferences about a seller based upon his or her location. Our results, which suggest a higher level of trust between more similar eBay users, indicate that eBay users

– whether consciously or not – are incorporating such information in their purchasing decisions. Our observations about buyer satisfaction raise an important question about whether eBay buyers incorporate this information in an accurate way.

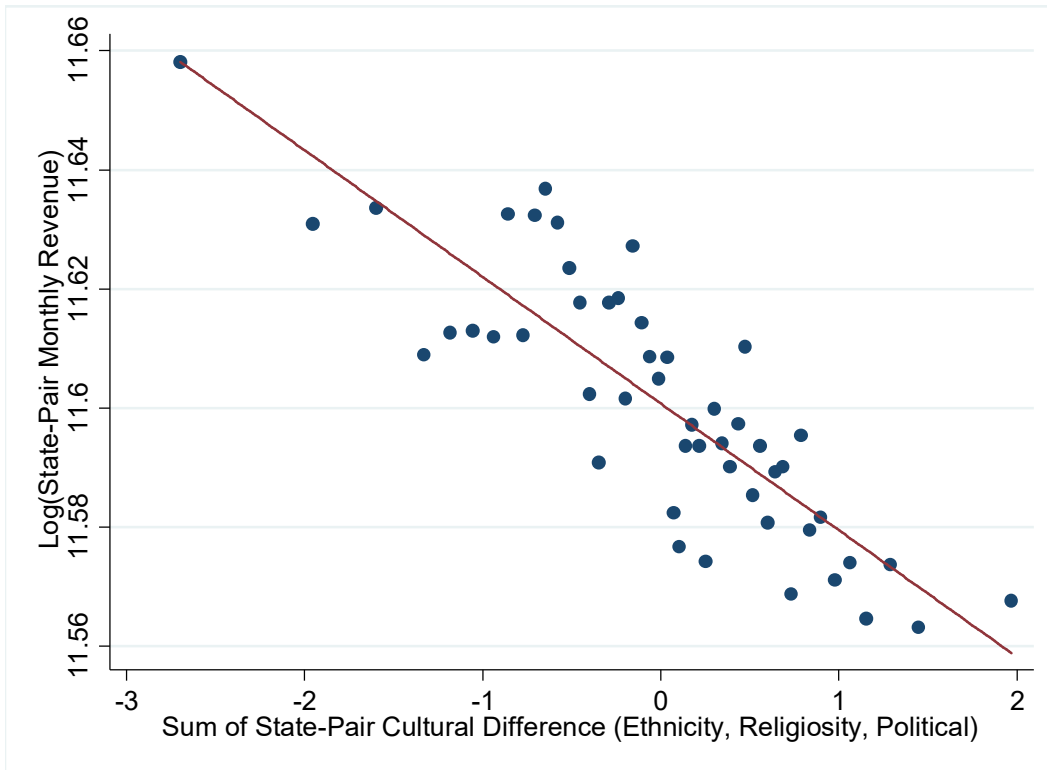
Finally, our results also point to some positive potential measures for stimulating transactions between dissimilar partners: quality-assurance institutions and information about sellers' track records appear to help buyers trust sellers from locations different from their own. Given the emphasis placed by eBay and other two-sided platforms on developing robust customer feedback mechanisms, such markets may hold out the promise of reducing the importance of sociocultural distance in economic transactions.

REFERENCES

- Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R., 2003. Fractionalization. *Journal of Economic Growth*, 8(2): 155-194.
- Anderson, J.E. and Van Wincoop, E., 2004. Trade costs. *Journal of Economic Literature*, 42(3): 691-751.
- Atkin, D., 2013. Trade, tastes, and nutrition in India. *American Economic Review*, 103(5): 1629-63.
- Bailey, M., Cao, R.R., Kuchler, T., Stroebel, J. and Wong, A., 2018. Social Connectedness: Measurement, Determinants, and Effects. *Journal of Economic Perspectives*, 32(3): 259-80.
- Bajari, P. and Hortacsu, A., 2004. Economic Insights from Internet Auctions. *Journal of Economic Literature*, 42(2): 457-486.
- Becker, G. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A., 2016. Stereotypes. *The Quarterly Journal of Economics*, 131(4): 1753-1794.
- Bronnenberg, B.J., Dhar, S.K. and Dubé, J.P., 2007. Consumer packaged goods in the United States: National brands, local branding. *Journal of Marketing Research*, 44(1): 4-13.
- Bronnenberg, B.J., Dubé, J.P.H. and Gentzkow, M., 2012. The evolution of brand preferences: Evidence from consumer migration. *American Economic Review*, 102(6): 2472-2508.
- Buchan, N.R., Croson, R.T. and Dawes, R.M., 2002. Swift neighbors and persistent strangers: A cross-cultural investigation of trust and reciprocity in social exchange. *American Journal of Sociology*, 108(1), pp.168-206.
- Buchan, N.R., Johnson, E.J., and Croson, R.T., 2006. Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences. *Journal of Economic Behavior & Organization*, 60(3): 373-398.
- Cabral, L. and Hortacsu, A., 2010. The dynamics of seller reputation: Evidence from eBay. *The Journal of Industrial Economics*, 58(1): 54-78.
- Cui, R., Li, J. and Zhang, D.J., 2016. Discrimination with incomplete information in the sharing economy: Field evidence from Airbnb. Washington University in St. Louis working paper.
- DeBruine, L. M. 2002. Facial resemblance enhances trust. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1498): 1307-1312.
- Doleac, J.L. and Stein, L.C., 2013. The visible hand: Race and online market outcomes. *The Economic Journal*, 123(572): F469-F492.
- Edelman, B., Luca, M. and Svirsky, D., 2017. Racial discrimination in the sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics*, 9(2): 1-22.
- Elfenbein, D.W., Fisman, R. and McManus, B., 2012. Charity as a substitute for reputation: Evidence from an online marketplace. *Review of Economic Studies*, 79(4): 1441-1468.
- Elfenbein, D.W., Fisman, R. and McManus, B., 2015. Market structure, reputation, and the value of quality certification. *American Economic Journal: Microeconomics*, 7(4): 83-108.
- Fershtman, C., and Gneezy, U., 2001. Discrimination in a segmented society: An experimental approach. *The Quarterly Journal of Economics*, 116(1): 351-377.
- Finger, J. M., & Kreinin, M. E. (1979). A Measure of Export Similarity and Its Possible Uses. *The Economic Journal*, 89(356), 905-912.

- Foddy, M., Platow, M.J., and Yamagishi, T., 2009. Group-based trust in strangers: The role of stereotypes and expectations. *Psychological Science*, 20(4): 419-422.
- Glaeser, E.L., Laibson, D.I., Scheinkman, J.A., and Soutter, C.L., 2000. Measuring trust. *The Quarterly Journal of Economics*, 115(3): 811-846.
- Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2006. Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2): 23-48.
- Guiso, L., Sapienza, P. and Zingales, L., 2009. Cultural biases in economic exchange? *The Quarterly Journal of Economics*, 124(3): 1095-1131.
- Hao, M.C., Dayal, U., Sharma, R K., Keim, D.A., and Janetzko, H., 2010. Visual analytics of large multidimensional data using variable binned scatter plots. In *Visualization and Data Analysis 2010* (Vol. 7530, p. 753006). International Society for Optics and Photonics.
- Hillberry, R. and Hummels, D., 2008. Trade responses to geographic frictions: A decomposition using micro-data. *European Economic Review*, 52(3): 527-550.
- Hortaçsu, A., Martínez-Jerez, F. and Douglas, J., 2009. The geography of trade in online transactions: Evidence from eBay and Mercadolibre. *American Economic Journal: Microeconomics*, 1(1): 53-74.
- Hui, X., Saeedi, M., Shen, Z. and Sundaresan, N., 2016. Reputation and regulations: evidence from eBay. *Management Science*, 62(12): 3604-3616.
- Hui, X., Saeedi, M., Spagnolo, G. and Tadelis, S., 2018. Certification, Reputation and Entry: An Empirical Analysis. National Bureau of Economic Research working paper w24916.
- Nosko, Chris, and Steven Tadelis, 2015. The limits of reputation in platform markets: An empirical analysis and field experiment. National Bureau of Economic Research working paper w20830.
- Perez-Truglia, Ricardo, 2018. Markets, trust and cultural biases evidence from eBay. *Journal of Behavioral and Experimental Economics* 72: 17-27.
- Silva, J.M.C.Santos, and Silvana Tenreyro, 2006. The log of gravity. *Review of Economics and Statistics* 88(4): 641-658.
- Tajfel, H., Billig, M.G., Bundy, R.P., and Flament, C., 1971. Social categorization and intergroup behaviour. *European journal of social psychology*, 1(2): 149-178.
- Zussman, A., 2013. Ethnic discrimination: Lessons from the Israeli online market for used cars. *The Economic Journal*, 123(572): F433-468.

Figure 1: Relationship between cultural distance and interstate trade



Note: The figure shows a binned scatter plot to illustrate the relationship between state-pair cultural differences and interstate trade. The horizontal axis is the sum of the three measures of cultural distance: ethnicity, religiosity, and voting. We use 50 bins, and residualize the data to account for all control variables included in Table 4, column 3.

Table 1: State-to-state annual sales and feedback

	Mean	Median	SD	Min	Max
Revenue	4,381,590	1,342,110	10,474,446	10,011	178M
Quantity	109,447	35,342	247,015	245	4M
Revenue, non-eTRS sellers	2,727,223	833,795	6,504,506	6,839	109M
Revenue, eTRS sellers	1,654,367	475,990	4,139,427	937	87M
Revenue, low feedback sellers	1,403,531	440,778	3,308,409	4,432	51M
Revenue, high feedback sellers	2,978,059	872,455	7,244,180	5,579	130M
Effective pct. positive feedback	50.87	51.01	4.72	21.77	70.36
Negative feedback (percent)	0.36	0.34	0.15	0	2.1
Observations	5100				

Table 2: State characteristics

	Mean	Median	SD	Min	Max
Winner vote share 2016	0.52	0.52	0.13	0.04	0.76
Median income (1000)	58.24	56.57	9.85	42.01	78.92
Bachelors share	0.31	0.30	0.06	0.20	0.57
Median age	38.37	38.30	2.37	31.00	44.60
Male share	0.49	0.49	0.01	0.47	0.52
Median home value (1000)	211.86	178.70	97.81	109.30	563.90
Urban share	0.73	0.73	0.16	0.33	1.00
Owner-occupied share	0.67	0.69	0.06	0.44	0.75
Observations	51				

Note: Includes all 50 states and the District of Columbia.

Table 3: State pair characteristics

	Mean	Median	SD	Min	Max
Ethnic difference	0.50	0.50	0.14	0.12	0.87
Religious difference	0.33	0.33	0.06	0.10	0.52
Voting Difference, 2016	0.14	0.12	0.11	0.00	0.71
<i>Difference in ...</i>					
Median income (10,000)	1.12	0.95	0.83	0.00	3.69
Bachelors share	0.07	0.05	0.06	0.00	0.37
Median age	2.58	2.00	2.15	0.00	13.60
Male share	0.01	0.01	0.01	0.00	0.05
Median home value (100,000)	0.98	0.68	0.98	0.00	4.55
Urban share	0.18	0.15	0.13	0.00	0.67
Owner-occupied share	0.06	0.04	0.06	0.00	0.31
Dist. between capitals (km)	1964.91	1600.78	1468.42	31.73	8226.99
Observations	2550				

Note: See text for variable definitions

Table 4: The impact of cultural and socioeconomic differences on interstate trade

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	OLS	OLS	OLS	OLS	PPML	PPML
Dependent variable	log(revenue)	log(revenue)	log(revenue)	log(quantity)	Revenue	Quantity
Ethnic difference	-0.110*** (0.0206)		-0.0798*** (0.0167)	-0.0635*** (0.0123)	-0.0647*** (0.0116)	-0.0721*** (0.0111)
Religious difference	-0.0286** (0.0129)		-0.0327** (0.0136)	-0.0410*** (0.0101)	-0.00986 (0.0155)	-0.0277*** (0.0103)
Voting difference	-0.0403*** (0.00801)		-0.0262*** (0.00631)	-0.0120** (0.00491)	-0.0215*** (0.00658)	-0.00901** (0.00362)
Median income (000)		0.0122* (0.00715)	0.00980 (0.00713)	0.00140 (0.00540)	0.00737 (0.00751)	0.00560 (0.00714)
Bachelors share		-0.0210** (0.00991)	-0.0133 (0.00992)	-0.00886 (0.00642)	0.000356 (0.00945)	-0.00327 (0.00652)
Median age		0.0104* (0.00576)	0.0119** (0.00532)	0.00741 (0.00521)	0.00860*** (0.00239)	0.00506** (0.00216)
Male share		-0.0257*** (0.00859)	-0.0237*** (0.00797)	-0.0108** (0.00450)	-0.0107 (0.00937)	-0.00425 (0.00835)
Median home value (000)		-0.0340** (0.0160)	-0.0197 (0.0156)	-0.00789 (0.00826)	-0.0147* (0.00857)	-0.00778 (0.00543)
Urban share		-0.0261*** (0.00715)	-0.0117* (0.00653)	-0.00866* (0.00462)	-0.0171 (0.0116)	-0.0140 (0.0105)
Owner-occupied share		-0.0270*** (0.00614)	-0.0152** (0.00613)	-0.0141*** (0.00437)	-0.0178** (0.00714)	-0.0113** (0.00476)
Log distance between capitals	-0.146*** (0.00955)	-0.145*** (0.0103)	-0.134*** (0.00919)	-0.115*** (0.00754)	-0.0973*** (0.0102)	-0.0849*** (0.00633)
Observations	5,100	5,100	5,100	5,100	5,100	5,100
R-squared (within)	.5702	.5605	.5873	.6308	n/a	n/a

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, all models include buyer-state \times year and seller-state \times year fixed effects. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

Table 5: The impact of taste differences on interstate trade

Estimation method	(1)	(2)	(3)	(4)	(5)
Dependent variable	OLS	OLS	OLS	OLS	PPML
	log(revenue)	log(revenue)	log(revenue)	log(revenue)	Revenue
Ethnic difference	-0.0798*** (0.0167)	-0.0497*** (0.0111)	-0.0554*** (0.0126)	-0.0492*** (0.0115)	-0.0371*** (0.00726)
Religious difference	-0.0327** (0.0136)	-0.0400*** (0.0119)	-0.0446*** (0.0115)	-0.0436*** (0.0111)	-0.0214* (0.0124)
Voting difference	-0.0262*** (0.00631)	-0.0103** (0.00464)	-0.00981** (0.00468)	-0.00800* (0.00446)	-0.00754 (0.00628)
Median income (000)	0.00980 (0.00713)	0.0105* (0.00560)	0.00924 (0.00602)	0.00984* (0.00576)	-0.00292 (0.00663)
Bachelors share	-0.0133 (0.00992)	-0.0156* (0.00902)	-0.0161** (0.00706)	-0.0162** (0.00771)	0.00431 (0.00680)
Median age	0.0119** (0.00532)	0.00503 (0.00364)	0.00525 (0.00359)	0.00430 (0.00341)	0.00293 (0.00236)
Male share	-0.0237*** (0.00797)	-0.0183** (0.00735)	-0.0197** (0.00765)	-0.0184** (0.00752)	-0.00681 (0.00591)
Median home value (000)	-0.0197 (0.0156)	-0.0174 (0.0114)	-0.0186* (0.0105)	-0.0178* (0.0105)	-0.00627 (0.00738)
Urban share	-0.0117* (0.00653)	-0.00767 (0.00539)	-0.00741 (0.00549)	-0.00701 (0.00535)	-0.0103 (0.00629)
Owner-occupied share	-0.0152** (0.00613)	-0.0114** (0.00543)	-0.0155*** (0.00544)	-0.0133** (0.00538)	-0.00810** (0.00372)
Euclidean category distance (revenue)		-0.269*** (0.0147)		-0.146*** (0.0261)	-0.197*** (0.0139)
Cross-product category difference (revenue)			-0.0783*** (0.00958)	-0.0458*** (0.00959)	-0.0410*** (0.00914)
Log distance between capitals	-0.134*** (0.00919)	-0.127*** (0.00836)	-0.126*** (0.00857)	-0.126*** (0.00849)	-0.0937*** (0.00661)
Observations	5,100	5,100	5,100	5,100	5,100
R-squared (within)	.5873	.6621	.6638	.6724	n/a

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, all models include buyer-state \times year and seller-state \times year fixed effects. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

Table 6: The moderating role of seller quality in the impact of cultural distance on trade

	(1)	(2a)	(2b)	(3a)	(3b)
	<i>base effect</i>	<i>base effect, eTRS = 0</i>	<i>add'l effect for eTRS = 1</i>	<i>base effect, High fdbk = 0</i>	<i>add'l effect for High fdbk = 1</i>
[1] Ethnic difference	-0.0492*** (0.0115)	-0.0640*** (0.0140)	0.0254* (0.0143)	-0.0754*** (0.0178)	0.0436*** (0.0153)
[2] Religious difference	-0.0436*** (0.0111)	-0.0390*** (0.00992)	-0.00949 (0.00833)	-0.0331*** (0.0102)	-0.0190 (0.0130)
[3] Voting difference	-0.00800* (0.00446)	-0.0128*** (0.00413)	0.0114* (0.00669)	-0.0203*** (0.00487)	0.0200** (0.00793)
Test of sum of [1]-[3] = 0	F = 42.64 (p = 0.000)	F = 47.70 (p = 0.000)	F = 4.00 (p = 0.051)	F = 33.81 (p = 0.000)	F = 5.67 (p = 0.021)
[4] Median income (000)	0.00984* (0.00576)	0.00814 (0.00528)	0.00301 (0.00598)	0.00660 (0.00399)	0.00534 (0.00584)
[5] Bachelors share	-0.0162** (0.00771)	-0.0166** (0.00710)	-0.00203 (0.00619)	-0.0131* (0.00670)	-0.00730 (0.00643)
[6] Median age	0.00430 (0.00341)	0.00860** (0.00385)	-0.0111** (0.00517)	0.00972** (0.00405)	-0.00670 (0.00567)
[7] Male share	-0.0184** (0.00752)	-0.0213*** (0.00775)	0.00679 (0.00848)	-0.0301*** (0.00704)	0.0176*** (0.00574)
[8] Med. home value (000)	-0.0178* (0.0105)	-0.0198** (0.00838)	0.00633 (0.0122)	-0.0302*** (0.00546)	0.0201 (0.0136)
[9] Urban share	-0.00701 (0.00535)	-0.0106** (0.00508)	0.00796 (0.00540)	-0.0175*** (0.00413)	0.0125** (0.00506)
[10] Owner-occup. share	-0.0133** (0.00538)	-0.0181*** (0.00660)	0.0151** (0.00652)	-0.0179** (0.00722)	0.00659 (0.00661)
Test of sum of [4]-[10] = 0	F = 26.68 (p = 0.000)	F = 28.61 (p = 0.000)	F = 1.52 (p = 0.2239)	F = 38.12 (p = 0.000)	F = 6.86 (p = 0.012)
Euclidean categ dist (revenue)	-0.146*** (0.0261)	-0.131*** (0.0353)	-0.0624 (0.0554)	-0.0744** (0.0348)	-0.124*** (0.0338)
Cross-product categ diff (revenue)	-0.0458*** (0.00959)	-0.0203*** (0.00562)	-0.0354** (0.0141)	0.00159 (0.00581)	-0.0593*** (0.0155)
Log distance between capitals	-0.126*** (0.00849)	-0.135*** (0.00928)	0.0212*** (0.00714)	-0.149*** (0.00899)	0.0326*** (0.00630)
Observations	5,100	10,200		10,200	
R-squared (within)	.6724	.5681		.6364	

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, the model in columns 2a and 2b includes buyer-state \times *eTRS* \times year and seller-state \times *eTRS* \times year fixed effects; the model in 3a and 3b replaces *eTRS* with *High feedback*. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

Table 7: The impact of cultural distance on customer satisfaction

VARIABLES	(1) EPP	(2) Share neg. fdbk
Ethnic difference	0.0657 (0.152)	-0.00642 (0.0109)
Religious difference	0.110 (0.0676)	0.00375 (0.00572)
Voting Difference	-0.0197 (0.0873)	0.0105*** (0.00364)
Median income (000)	-0.0195 (0.0619)	0.00253 (0.00441)
Bachelors share	0.0617 (0.0740)	-0.00208 (0.00333)
Median age	0.00771 (0.0765)	0.00170 (0.00448)
Male share	0.0556 (0.0490)	0.00209 (0.00490)
Median home value (000)	-0.00119 (0.0854)	-0.0132** (0.00550)
Urban share	0.106 (0.0645)	-0.00219 (0.00252)
Owner-occupied share	-0.0713 (0.0476)	-0.00511 (0.00532)
Log distance between capitals	-0.532*** (0.0848)	0.0150*** (0.00331)
Observations	5,100	5,100
R-squared (within)	0.0295	0.0176

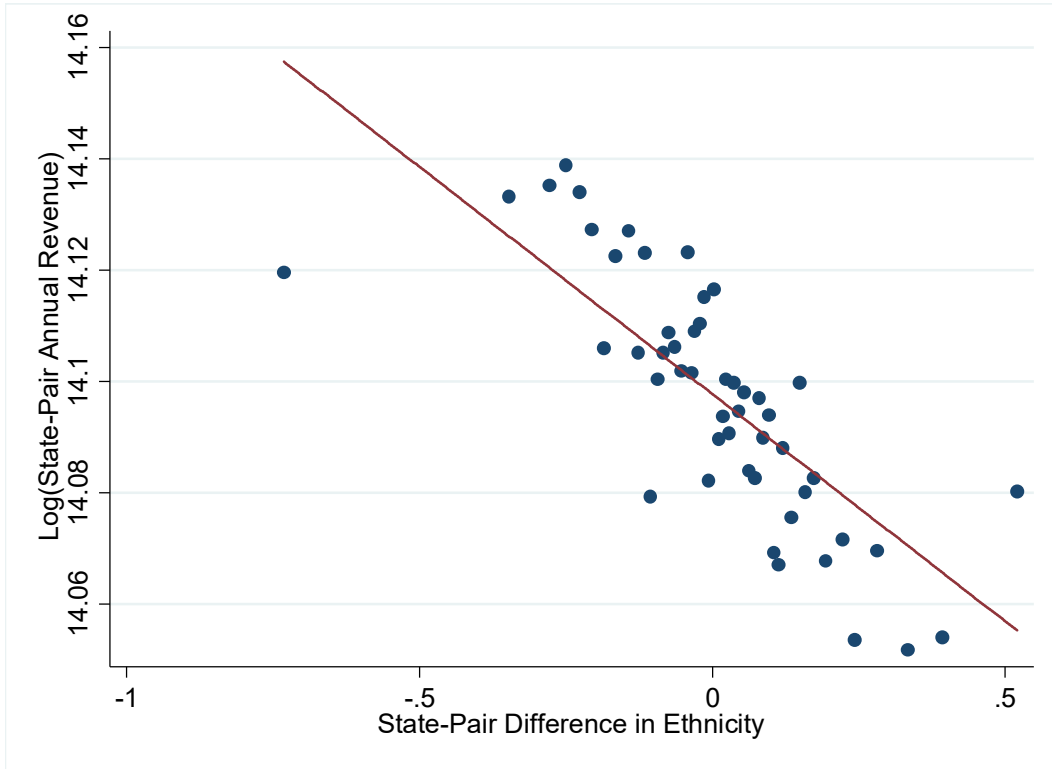
Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: In addition to the listed variables, all models include buyer-state \times year and seller-state \times year fixed effects. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

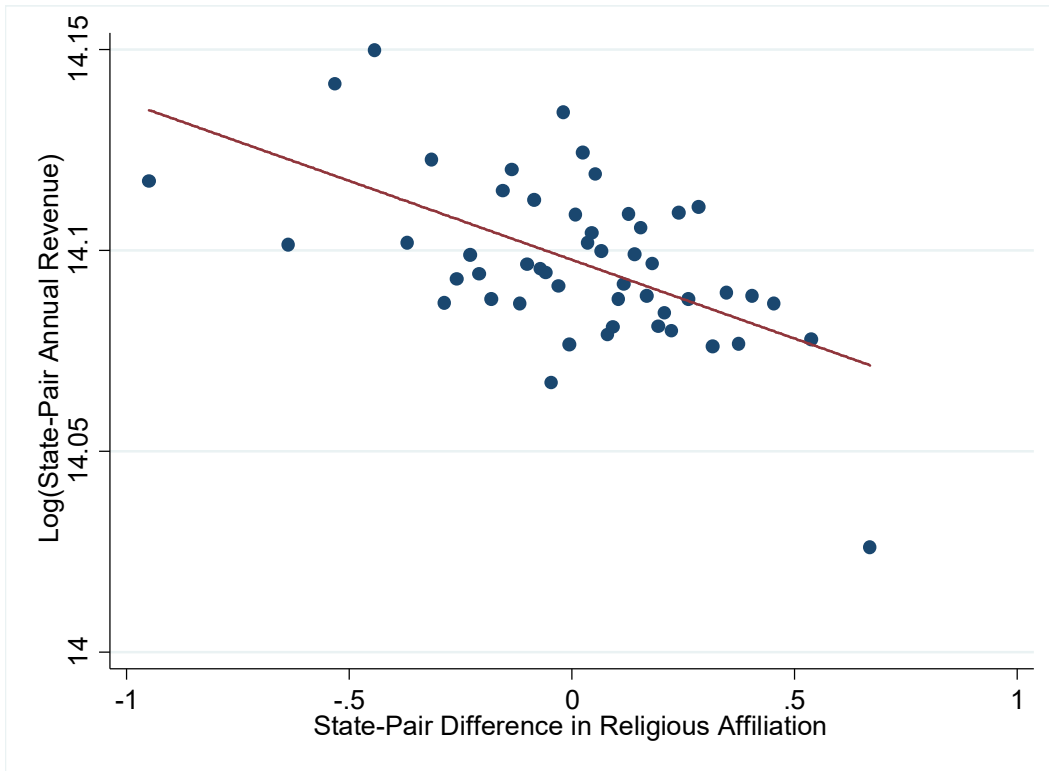
APPENDIX

Figure A1: Relationship between ethnic differences and interstate trade



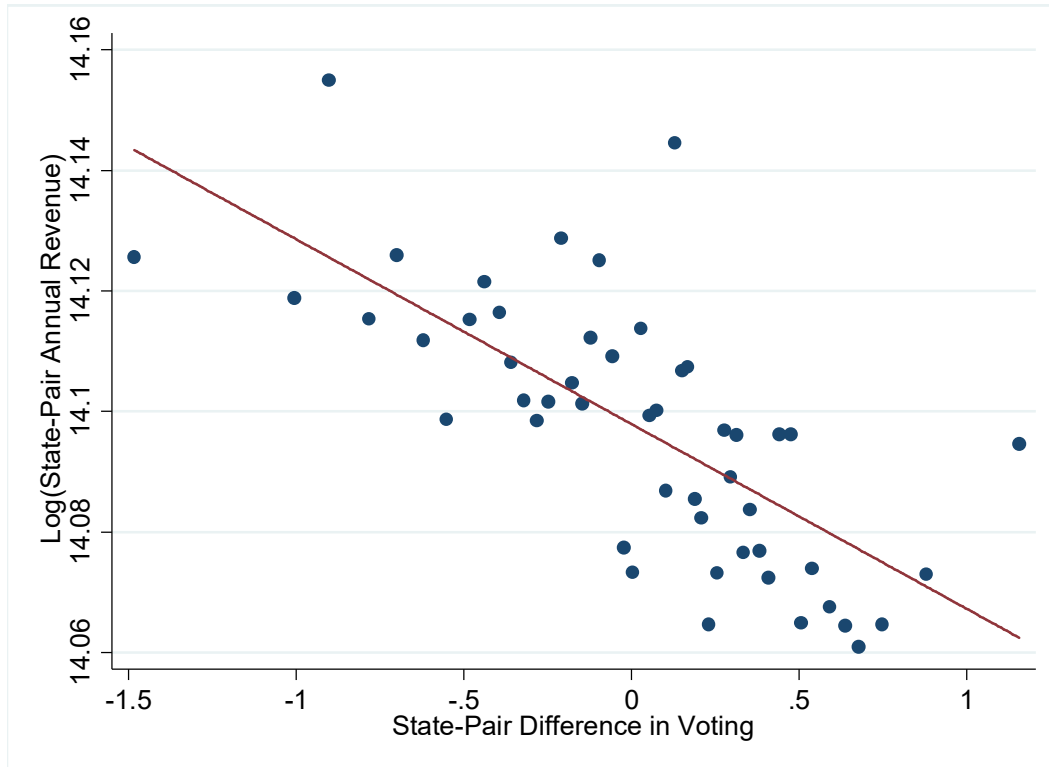
Note: The figure shows a binned scatter plot to illustrate the relationship between state-pair differences in ethnicity and interstate trade. We use 50 bins, and residualize the data to account for all other variables included in Table 4, column 3.

Figure A2: Relationship between differences in religiosity interstate trade



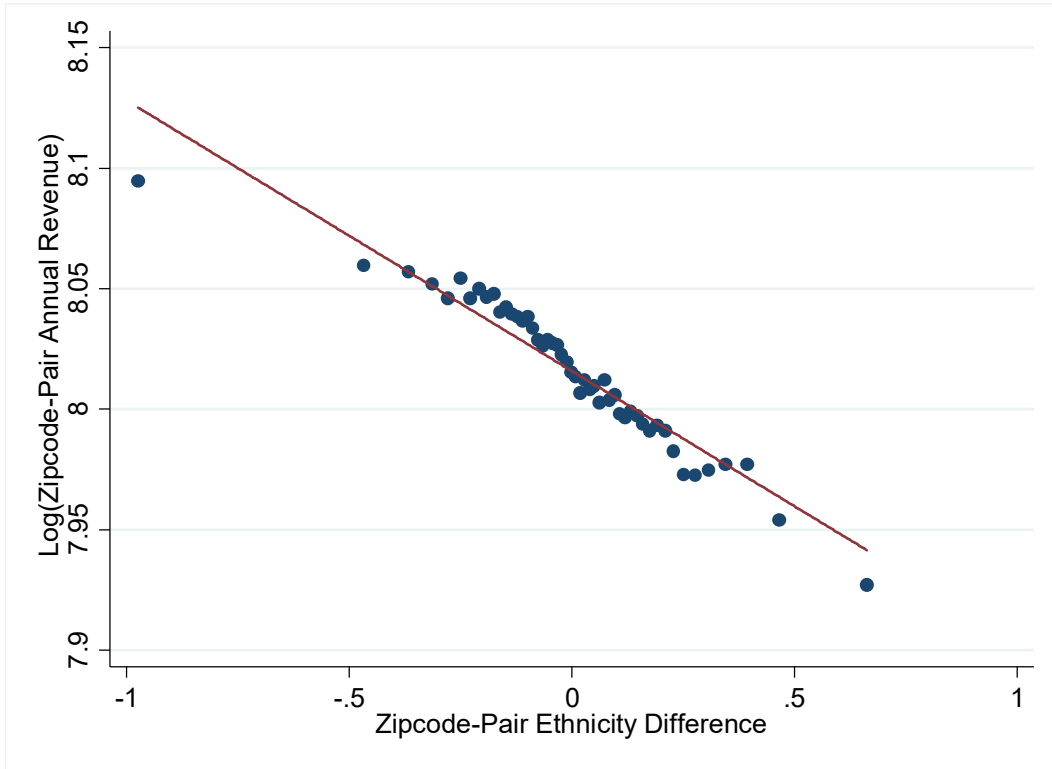
Note: The figure shows a binned scatter plot to illustrate the relationship between state-pair differences in religiosity and interstate trade. We use 50 bins, and residualize the data to account for all other variables included in Table 4, column 3.

Figure A3: Relationship between differences in voting and interstate trade



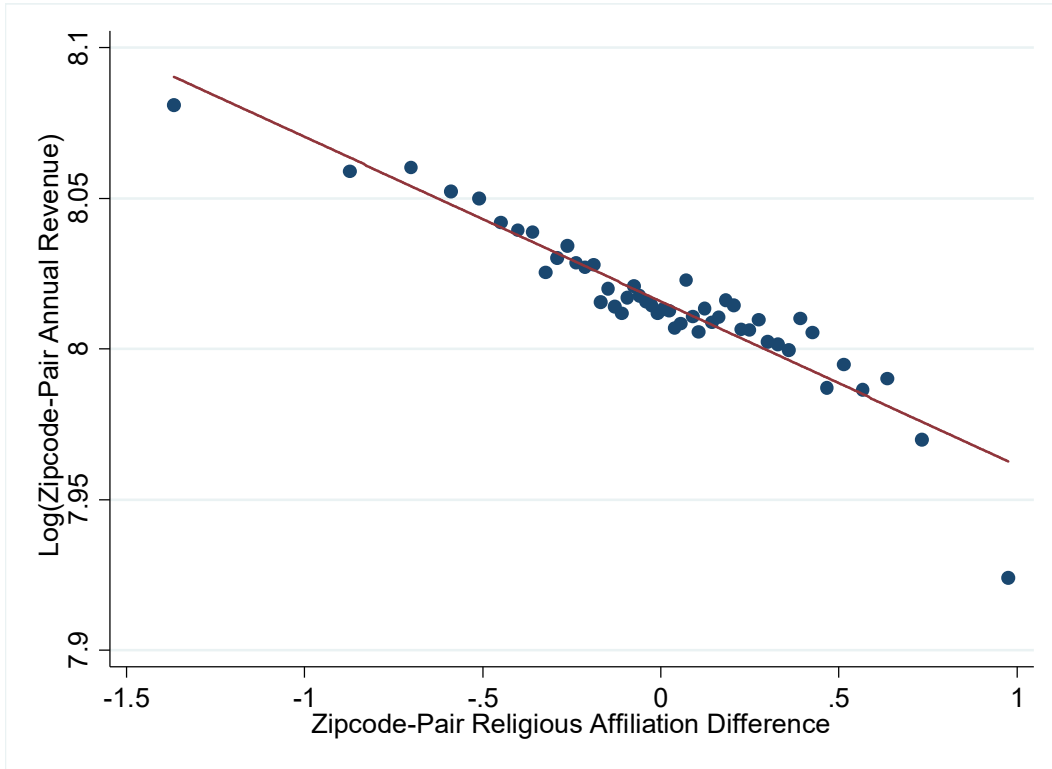
Note: The figure shows a binned scatter plot to illustrate the relationship between state-pair differences in voting in the 2016 presidential election and interstate trade. We use 50 bins, and residualize the data to account for all other variables included in Table 4, column 3.

Figure A4: Relationship between ethnic differences and zip-to-zip trade



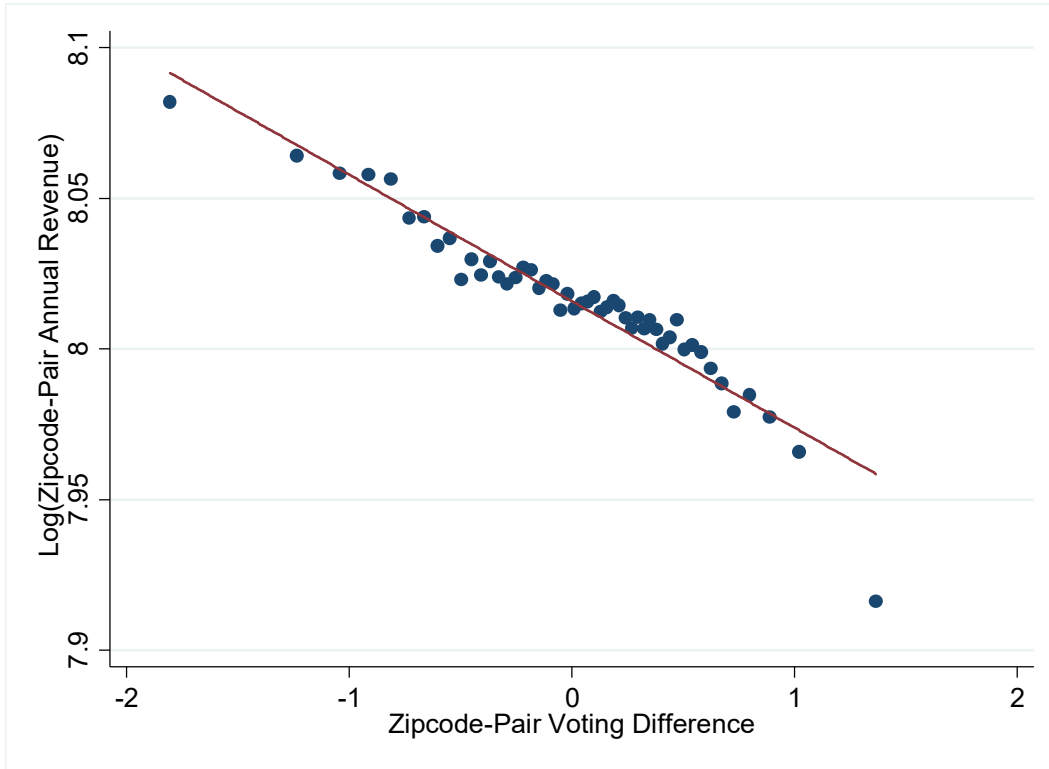
Note: The figure shows a binned scatter plot to illustrate the relationship between zip-code-pair differences in ethnicity and trade. We use 50 bins, and residualize the data to account for all other variables included in Table A4, column 3.

Figure A5: Relationship between religiosity differences and zip-to-zip trade



Note: The figure shows a binned scatter plot to illustrate the relationship between zip-code-pair differences in religiosity and trade. We use 50 bins, and residualize the data to account for all other variables included in Table A4, column 3.

Figure A6: Relationship between voting differences and zip-to-zip trade



Note: The figure shows a binned scatter plot to illustrate the relationship between zip-code-pair differences in voting in the 2016 presidential election and trade. We use 50 bins, and residualize the data to account for all other variables included in Table A4, column 3.

Table A1: Zip code to zip code annual sales and feedback

	Mean	Median	SD	Min	Max
Revenue	15,179	3,524	56,868	0	11,865,290
Quantity	379	101	1,176	0	207,816
Revenue, non-eTRS sellers	9,499	2,296	33,207	0	7,351,905
Revenue, eTRS sellers	5,680	1,042	27,910	0	8,572,258
Effective percent positive feedback	52.49	52.27	13.44	0	100
Negative feedback share	0.37	0	1.76	0	100
Observations	1,571,764				

Table A2: Zip code characteristics

	Mean	Median	SD	Min	Max
Winner vote % 2016	0.55	0.56	0.17	0.04	0.92
Median income (1000)	56.05	53.21	14.28	24.75	116.20
Bachelors share	0.28	0.27	0.10	0.07	0.74
Median age	39.08	38.90	3.67	25.99	52.40
Male share	0.49	0.49	0.01	0.47	0.56
Median home value (1000)	201.62	154.41	134.95	48.79	927.34
Urban share	0.69	0.71	0.24	0.00	1.00
Owner-occupied share	0.68	0.69	0.08	0.20	0.82
Observations	887				

Table A3: Zip code pair characteristics

	Mean	Median	SD	Min	Max
Ethnic difference	0.52	0.52	0.18	0.04	0.95
Religious difference	0.33	0.33	0.07	0.02	0.70
Voting difference, 2016	0.19	0.17	0.14	0.00	0.88
<i>Difference in ...</i>					
Median income (10,000)	1.51	1.14	1.34	0.00	9.14
Bachelors share	0.10	0.08	0.08	0.00	0.67
Median age	4.11	3.45	3.17	0.00	26.41
Male share	0.01	0.01	0.01	0.00	0.09
Median home value (100,000)	1.25	0.71	1.44	0.00	8.79
Urban share	0.28	0.24	0.21	0.00	1.00
Owner-occupied share	0.09	0.07	0.08	0.00	0.62
Dist. between capitals (km)	1633.73	1408.81	1078.88	0.00	8368.32
Observations	785,882				

Table A4: The impact of cultural and socioeconomic differences on trade between zip codes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	All	All	All	Same state	All	All	All
Estimation method	OLS	OLS	OLS	OLS	OLS	PPML	PPML
Dependent variable	log(revenue)	log(revenue)	log(revenue)	log(revenue)	log(quantity)	Revenue	Quantity
Ethnic difference	-0.114*** (0.00877)		-0.112*** (0.00814)	-0.107*** (0.0242)	-0.0611*** (0.00639)	-0.0862*** (0.0129)	-0.0974*** (0.0114)
Religious difference	-0.0595*** (0.00483)		-0.0545*** (0.00476)	-0.0379** (0.0172)	-0.0387*** (0.00370)	-0.0368*** (0.00874)	-0.0433*** (0.00490)
Voting difference	-0.0634*** (0.00438)		-0.0420*** (0.00353)	-0.0404*** (0.00732)	-0.0276*** (0.00276)	-0.0281*** (0.00596)	-0.0166*** (0.00478)
Median income (000)		0.00781* (0.00438)	0.000922 (0.00402)	-0.00750 (0.00779)	0.00302 (0.00289)	0.00463 (0.00571)	0.000154 (0.00447)
Bachelors share		-0.0347*** (0.00505)	-0.0248*** (0.00450)	0.0267*** (0.00790)	-0.0197*** (0.00324)	-0.0238*** (0.00663)	-0.0134*** (0.00439)
Median age		-0.00335 (0.00260)	0.000745 (0.00222)	-0.0163** (0.00660)	0.00196 (0.00167)	0.000357 (0.00303)	-0.000569 (0.00279)
Male share		-0.0179*** (0.00398)	-0.0126*** (0.00372)	-0.00575 (0.00799)	-0.0173*** (0.00274)	-0.0193*** (0.00415)	-0.0128*** (0.00309)
Median home value (000)		-0.0572*** (0.00651)	-0.0281*** (0.00651)	-0.0797*** (0.0124)	-0.0166*** (0.00472)	-0.0104 (0.00777)	-0.00462 (0.00746)
Urban share		-0.0299*** (0.00444)	-0.00893** (0.00413)	-0.0456*** (0.00677)	-0.0368*** (0.00355)	-0.0555*** (0.00633)	-0.0394*** (0.00454)
Owner-occupied share		-0.0269*** (0.00448)	-0.00163 (0.00403)	0.0274*** (0.00806)	-0.00161 (0.00294)	0.000110 (0.00447)	-0.00179 (0.00406)
Log distance between centers	-0.167*** (0.00547)	-0.179*** (0.00569)	-0.161*** (0.00547)	-0.0798*** (0.00999)	-0.133*** (0.00403)	-0.0625*** (0.00686)	-0.0656*** (0.00563)
Observations	1,571,764	1,571,764	1,571,764	43,556	1,571,764	1,571,764	1,571,764
R-squared (within)	0.0699	0.0644	0.0717	0.0900	0.1527	n/a	n/a

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, all models include buyer-zip code × year and seller-zip code × year fixed effects. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer zip code and seller zip code.

Table A5: Quantity-focused models of taste and trade between states

Estimation method	(1)	(2)	(3)	(4)	(5)
Dependent variable	OLS	OLS	OLS	OLS	PPML
	log(quantity)	log(quantity)	log(quantity)	log(quantity)	Quantity
Ethnic difference	-0.0635*** (0.0123)	-0.0247*** (0.00642)	-0.0388*** (0.00924)	-0.0246*** (0.00646)	-0.0207*** (0.00738)
Religious difference	-0.0410*** (0.0101)	-0.0317*** (0.00831)	-0.0353*** (0.00862)	-0.0317*** (0.00840)	-0.0260*** (0.00739)
Voting difference	-0.0120** (0.00491)	-0.0104** (0.00424)	-0.0103** (0.00456)	-0.0104** (0.00433)	-0.0144*** (0.00283)
Median income (000)	0.00140 (0.00540)	0.000335 (0.00458)	0.000215 (0.00485)	0.000359 (0.00453)	-0.00101 (0.00389)
Bachelors share	-0.00886 (0.00642)	-0.0103* (0.00538)	-0.0103* (0.00538)	-0.0102* (0.00539)	0.00111 (0.00463)
Median age	0.00741 (0.00521)	0.00575 (0.00432)	0.00585 (0.00432)	0.00578 (0.00435)	0.00217 (0.00272)
Male share	-0.0108** (0.00450)	-0.00942** (0.00441)	-0.0101** (0.00454)	-0.00940** (0.00440)	-0.00818* (0.00485)
Median home value (000)	-0.00789 (0.00826)	-0.00761 (0.00645)	-0.00698 (0.00681)	-0.00765 (0.00635)	-0.00720** (0.00365)
Urban share	-0.00866* (0.00462)	-0.00868** (0.00360)	-0.00852** (0.00395)	-0.00869** (0.00358)	-0.00663* (0.00393)
Owner-occupied share	-0.0141*** (0.00437)	-0.0141*** (0.00364)	-0.0153*** (0.00362)	-0.0140*** (0.00357)	-0.00625*** (0.00242)
Euclidean category distance (revenue)		-0.139*** (0.0211)		-0.144*** (0.0323)	-0.121*** (0.0374)
Cross-product category difference (revenue)			-0.103*** (0.0382)	0.00573 (0.0399)	-6.85e-05 (0.0540)
Log distance between capitals	-0.115*** (0.00754)	-0.115*** (0.00723)	-0.115*** (0.00735)	-0.115*** (0.00723)	-0.0868*** (0.00452)
Observations	5,100	5,100	5,100	5,100	5,100
R-squared (within)	.6308	.6880	.6709	.6881	n/a

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, all models include buyer-state \times year and seller-state \times year fixed effects. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

Table A6: The impact of taste differences on trade between zip codes

Estimation method	(1)	(2)	(3)	(4)	(5)
Dependent variable	OLS	OLS	OLS	OLS	PPML
	log(revenue)	log(revenue)	log(revenue)	log(revenue)	Revenue
Ethnic difference	-0.112*** (0.00814)	-0.0826*** (0.00726)	-0.0831*** (0.00689)	-0.0812*** (0.00695)	-0.0582*** (0.00920)
Religious difference	-0.0545*** (0.00476)	-0.0556*** (0.00451)	-0.0561*** (0.00448)	-0.0560*** (0.00447)	-0.0400*** (0.00825)
Voting difference	-0.0420*** (0.00353)	-0.0324*** (0.00323)	-0.0323*** (0.00319)	-0.0318*** (0.00320)	-0.0123*** (0.00458)
Median income (000)	0.000922 (0.00402)	-0.00209 (0.00377)	-0.00168 (0.00375)	-0.00192 (0.00375)	-0.00113 (0.00458)
Bachelors share	-0.0248*** (0.00450)	-0.0191*** (0.00419)	-0.0210*** (0.00414)	-0.0204*** (0.00417)	-0.0175*** (0.00481)
Median age	0.000745 (0.00222)	-0.000751 (0.00203)	-0.000606 (0.00197)	-0.000718 (0.00197)	-0.00532** (0.00267)
Male share	-0.0126*** (0.00372)	-0.00605* (0.00342)	-0.00640* (0.00336)	-0.00596* (0.00337)	-0.00875*** (0.00328)
Median home value (000)	-0.0281*** (0.00651)	-0.0215*** (0.00592)	-0.0208*** (0.00588)	-0.0205*** (0.00586)	-0.00415 (0.00520)
Urban share	-0.00893** (0.00413)	0.000223 (0.00390)	-0.000170 (0.00386)	0.000435 (0.00388)	-0.0401*** (0.00604)
Owner-occupied share	-0.00163 (0.00403)	0.00403 (0.00358)	0.00210 (0.00341)	0.00275 (0.00346)	0.00458 (0.00324)
Euclidean category distance (revenue)		-0.372*** (0.0166)		-0.0840*** (0.0291)	-0.198*** (0.0253)
Cross-product category difference (revenue)			-0.127*** (0.00670)	-0.106*** (0.0114)	-0.0733*** (0.0101)
Log distance between capitals	-0.161*** (0.00547)	-0.158*** (0.00513)	-0.159*** (0.00513)	-0.159*** (0.00511)	-0.0629*** (0.00598)
Observations	1,571,764	1,571,764	1,571,764	1,571,764	1,571,764
R-squared (within)	0.0717	0.0825	0.0853	0.0854	n/a

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, all models include buyer-zip code × year and seller- zip code × year fixed effects. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

Table A7: Interstate trade within product categories

	(1)	(2)	(3)	(4)
Estimation method	OLS	OLS	PPML	PPML
Dependent variable	log(revenue)	log(quantity)	Revenue	Quantity
Ethnic difference	-0.0860*** (0.0177)	-0.0443*** (0.0117)	-0.0637*** (0.0110)	-0.0724*** (0.0111)
Religious difference	-0.0548*** (0.0121)	-0.0385*** (0.00737)	-0.00848 (0.0155)	-0.0280*** (0.0104)
Voting difference	-0.0143** (0.00653)	-0.0113** (0.00446)	-0.0225*** (0.00726)	-0.00889** (0.00377)
Median income (000)	0.00750* (0.00425)	0.00148 (0.00308)	0.00766 (0.00776)	0.00614 (0.00721)
Bachelors share	-0.0136 (0.0103)	-0.000223 (0.00397)	0.000738 (0.00980)	-0.00348 (0.00666)
Median age	0.00546 (0.00629)	0.00165 (0.00416)	0.00815*** (0.00232)	0.00526** (0.00224)
Male share	-0.00794 (0.00685)	-0.0167** (0.00703)	-0.00964 (0.00908)	-0.00427 (0.00835)
Median home value (000)	-0.0263** (0.0103)	-0.0152** (0.00621)	-0.0139* (0.00800)	-0.00843 (0.00551)
Urban share	0.00981 (0.00586)	-0.0121*** (0.00389)	-0.0168 (0.0115)	-0.0141 (0.0105)
Owner-occupied share	-0.00889 (0.00627)	-0.00427 (0.00340)	-0.0190*** (0.00712)	-0.0111** (0.00488)
Log distance between capitals	-0.158*** (0.00879)	-0.125*** (0.00678)	-0.0987*** (0.0101)	-0.0847*** (0.00643)
Observations	168,300	168,300	168,300	168,300
R-squared (within)	0.0500	0.1157	n/a	n/a

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, all models include buyer-state \times product category \times year and seller-state \times product category \times year fixed effects. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

Table A8: Quantity-focused trust and trade models

	(1)	(2a)	(2b)	(3a)	(3b)
	<i>base effect</i>	<i>base effect, eTRS = 0</i>	<i>add'l effect for eTRS = 1</i>	<i>base effect, High fdbk = 0</i>	<i>add'l effect for High fdbk = 1</i>
[1] Ethnic difference	-0.0246*** (0.00646)	-0.0453*** (0.0116)	0.0301 (0.0189)	-0.0711*** (0.0125)	0.0628*** (0.0132)
[2] Religious difference	-0.0317*** (0.00840)	-0.0350*** (0.00818)	0.00498 (0.00527)	-0.0362*** (0.00808)	0.00673 (0.00777)
[3] Voting difference	-0.0104** (0.00433)	-0.0125*** (0.00411)	0.00570 (0.00575)	-0.0151*** (0.00296)	0.00636 (0.00479)
Test of sum of [1]-[3] = 0	F = 58.69 (p = 0.000)	F = 46.00 (p = 0.000)	F = 4.02 (p = 0.050)	F = 66.12 (p = 0.000)	F = 30.70 (p = 0.000)
[4] Median income (000)	0.000359 (0.00453)	0.00145 (0.00404)	0.000215 (0.00557)	0.000755 (0.00322)	-3.83e-05 (0.00471)
[5] Bachelors share	-0.0102* (0.00539)	-0.0104** (0.00508)	-0.00222 (0.00478)	-0.0114** (0.00492)	0.00114 (0.00409)
[6] Median age	0.00578 (0.00435)	0.00700 (0.00516)	-0.00335 (0.00456)	0.00825*** (0.00187)	-0.00316 (0.00543)
[7] Male share	-0.00940** (0.00440)	-0.00751 (0.00534)	-0.00556 (0.00534)	-0.0206*** (0.00399)	0.0155*** (0.00468)
[8] Median home value (000)	-0.00765 (0.00635)	-0.0119** (0.00525)	0.00737 (0.00902)	-0.0203*** (0.00421)	0.0172** (0.00713)
[9] Urban share	-0.00869** (0.00358)	-0.0108*** (0.00370)	0.00501 (0.00423)	-0.0133*** (0.00259)	0.00496 (0.00348)
[10] Owner-occup. share	-0.0140*** (0.00357)	-0.0154*** (0.00440)	0.00505 (0.00546)	-0.0151*** (0.00522)	0.00111 (0.00438)
Test of sum of [4]-[10] = 0	F = 24.20 (p = 0.000)	F = 25.58 (p = 0.000)	F = 0.27 (p = 0.604)	F = 41.97 (p = 0.000)	F = 8.56 (p = 0.005)
Euclidean categ dist (quantity)	-0.144*** (0.0323)	-0.119*** (0.0258)	-0.0438 (0.0486)	-0.0531** (0.0207)	-0.119*** (0.0350)
Cross-product categ diff (quantity)	0.00573 (0.0399)	0.0315 (0.0399)	-0.0419 (0.0367)	0.0157 (0.0185)	-0.00903 (0.0342)
Log distance between capitals	-0.115*** (0.00723)	-0.125*** (0.00797)	0.0231*** (0.00459)	-0.134*** (0.00593)	0.0244*** (0.00576)
Observations	5,100	10,200		10,200	
R-squared (within)	.6881	.5922		.7090	

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, the model in columns 2a and 2b includes buyer-state \times *eTRS* \times year and seller-state \times *eTRS* \times year fixed effects; the model in 3a and 3b replaces *eTRS* with *High feedback*. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

Table A9: Revenue-focused trust and trade models, PPML estimation

	(1)	(2a)	(2b)	(3a)	(3b)
	<i>base effect</i>	<i>base effect, eTRS = 0</i>	<i>add'l effect for eTRS = 1</i>	<i>base effect, High fdbk = 0</i>	<i>add'l effect for High fdbk = 1</i>
[1] Ethnic difference	-0.0371*** (0.00726)	-0.0449*** (0.00848)	0.0227* (0.0116)	-0.0572*** (0.00789)	0.0320*** (0.0101)
[2] Religious difference	-0.0214* (0.0124)	-0.0262** (0.0113)	0.00945 (0.0111)	-0.0384*** (0.00910)	0.0223* (0.0116)
[3] Voting difference	-0.00754 (0.00628)	-0.0147** (0.00617)	0.0178** (0.00718)	-0.0215*** (0.00462)	0.0211*** (0.00617)
Test of sum of [1]-[3] = 0	$\chi^2 = 15.38$ (p = 0.001)	$\chi^2 = 27.03$ (p = 0.000)	$\chi^2 = 6.96$ (p = 0.008)	$\chi^2 = 84.08$ (p = 0.000)	$\chi^2 = 15.09$ (p = 0.000)
[4] Median income (000)	-0.00292 (0.00663)	-0.00254 (0.00616)	0.000599 (0.00656)	-0.00169 (0.00367)	-0.00170 (0.00743)
[5] Bachelors share	0.00431 (0.00680)	0.000109 (0.00522)	0.00949 (0.00802)	0.000259 (0.00412)	0.00441 (0.00829)
[6] Median age	0.00293 (0.00236)	0.00572** (0.00263)	-0.00751 (0.00587)	0.00644*** (0.00213)	-0.00553 (0.00348)
[7] Male share	-0.00681 (0.00591)	-0.00653 (0.00609)	-0.000514 (0.0102)	-0.0150*** (0.00579)	0.0123* (0.00714)
[8] Median home value (000)	-0.00627 (0.00738)	-0.00392 (0.00638)	-0.00648 (0.00705)	-0.0135*** (0.00448)	0.00992 (0.00669)
[9] Urban share	-0.0103 (0.00629)	-0.0126** (0.00598)	0.00712 (0.00836)	-0.0184*** (0.00505)	0.0109 (0.00729)
[10] Owner-occup. share	-0.00810** (0.00372)	-0.0124*** (0.00455)	0.0104* (0.00623)	-0.0196*** (0.00505)	0.0161*** (0.00551)
Test of sum of [4]-[10] = 0	$\chi^2 = 6.26$ (p = 0.012)	$\chi^2 = 8.18$ (p = 0.004)	$\chi^2 = 0.55$ (p = 0.459)	$\chi^2 = 24.34$ (p = 0.000)	$\chi^2 = 9.15$ (p = 0.003)
Euclidean categ dist (revenue)	-0.197*** (0.0139)	-0.165*** (0.0193)	-0.121*** (0.0404)	-0.0812*** (0.0253)	-0.189*** (0.0231)
Cross-product categ diff (revenue)	-0.0410*** (0.00914)	-0.0334*** (0.0101)	-0.00287 (0.00828)	-0.00689 (0.0114)	-0.0370* (0.0200)
Log distance between capitals	-0.0937*** (0.00661)	-0.0994*** (0.00645)	0.0153** (0.00606)	-0.112*** (0.00702)	0.0265*** (0.00652)
Observations	5,100		10,200		10,200

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, the model in columns 2a and 2b includes buyer-state \times *eTRS* \times year and seller-state \times *eTRS* \times year fixed effects; the model in 3a and 3b replaces *eTRS* with *High feedback*. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

Table A10: Revenue-focused trust and trade models, OLS and all zip codes

	(1)	(2a)	(2b)
	<i>base effect</i>	<i>base effect, eTRS = 0</i>	<i>add'l effect for eTRS = 1</i>
[1] Ethnic difference	-0.0812*** (0.00695)	-0.0919*** (0.00757)	0.0236*** (0.00785)
[2] Religious difference	-0.0560*** (0.00447)	-0.0633*** (0.00483)	0.00374 (0.00434)
[3] Voting difference	-0.0318*** (0.00320)	-0.0332*** (0.00348)	0.00686 (0.00465)
Test of sum of [1]-[3] = 0	F = 382.59 (p = 0.000)	F = 411.48 (p = 0.000)	F = 10.35 (p = 0.001)
[4] Median income (000)	-0.00192 (0.00375)	0.000967 (0.00398)	0.00150 (0.00429)
[5] Bachelors share	-0.0204*** (0.00417)	-0.0214*** (0.00454)	0.00362 (0.00468)
[6] Median age	-0.000718 (0.00197)	-0.000677 (0.00219)	0.00216 (0.00263)
[7] Male share	-0.00596* (0.00337)	-0.00811** (0.00360)	0.00338 (0.00353)
[8] Median home value (000)	-0.0205*** (0.00586)	-0.0243*** (0.00619)	0.0124** (0.00537)
[9] Urban share	0.000435 (0.00388)	0.00404 (0.00434)	0.000207 (0.00472)
[10] Owner-occup. share	0.00275 (0.00346)	0.00179 (0.00377)	0.00323 (0.00376)
Test of sum of [4]-[10] = 0	F = 19.47 (p = 0.000)	F = 16.66 (p = 0.000)	F = 7.01 (p = 0.008)
Euclidean categ dist (revenue)	-0.0840*** (0.0291)	-0.138*** (0.0366)	-0.0124 (0.0371)
Cross-product categ diff (revenue)	-0.106*** (0.0114)	-0.0570*** (0.0140)	-0.0514*** (0.0152)
Log distance between capitals	-0.159*** (0.00511)	-0.176*** (0.00549)	0.0308*** (0.00281)
Observations	1,571,764	3,143,528	
R-squared (within)	0.0854	0.0534	

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1 (two-sided test)

Note: In addition to the listed variables, the model in columns 2a and 2b includes buyer-state \times eTRS \times year and seller-state \times eTRS \times year fixed effects. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.

Table A11: The impact of cultural distance on customer satisfaction, OLS and all zip codes

VARIABLES	(1) EPP	(2) Share neg. feedback
Ethnic difference	0.335*** (0.0673)	-0.0231*** (0.00709)
Religious difference	-0.00526 (0.0363)	0.0130*** (0.00414)
Voting Difference	0.00983 (0.0285)	0.000125 (0.00300)
Median income (000)	-0.0533* (0.0315)	-0.00367 (0.00326)
Bachelors share	0.0674** (0.0303)	0.000590 (0.00304)
Median age	0.0221 (0.0235)	0.00181 (0.00207)
Male share	0.0411* (0.0247)	0.00111 (0.00327)
Median home value (000)	0.0539 (0.0407)	0.00137 (0.00399)
Urban share	0.0803*** (0.0270)	-0.00420 (0.00274)
Owner-occupied share	-0.00431 (0.0305)	-0.00298 (0.00340)
Log distance between capitals	-0.471*** (0.0219)	0.0200*** (0.00217)
Observations	1,552,549	1,552,549
R-squared (within)	0.0008	0.0001

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: In addition to the listed variables, all models include buyer-state \times year and seller-state \times year fixed effects. The variables median income, bachelors share, median age, male share, median home value, urban share and owner-occupied share all refer to the absolute value of state-level differences. Standard errors are two-way clustered by buyer state and seller state.