The impact of socioeconomic and cultural differences on online trade*

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

Raymond Fisman
Boston University and NBER
rfisman@bu.edu

Brian McManus
University of North Carolina at Chapel Hill
mcmanusb@email.unc.edu

February 2022

Abstract: We use eBay data to investigate how within-U.S. trade is influenced by differences in socioeconomic characteristics. States' similarity in cultural characteristics (ethnicity, religious affiliations, and political behavior) predicts online trade; cultural similarity similarly predicts trade within product categories. The culture-trade relationship 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 in actual transactions.

Keywords: e-commerce, culture, trust, trade frictions
JEL Codes: D12, D91, F14, L15, R12

* 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, Washington University in St. Louis, and the ASSA 2019 meetings for helpful comments.

1. Introduction

Commerce, whether conducted electronically or through traditional channels, requires a match between a seller's production or inventory and a buyer's demands, as well as buyer-seller trust. Electronic platforms use a variety of well-documented mechanisms to help buyers locate products offered by trustworthy sellers capable of providing speedy and error-free delivery (Cabral and Hortascu 2010; Bolton, Greiner, and Ockenfels 2013; Hui, Saeedi, Shen, and Sundaresan 2016; Brynjolfsson, Hui, and Lui 2019). The rise of these platforms has led some observers to posit the emergence of a "flat world" of "frictionless" commerce (Friedman 2005). Others, often drawing on evidence from the international trade literature, have argued that distance – whether geographic, administrative, economic, or cultural – will continue to impact patterns of exchange (Ghemawat 2001; 2007). This debate has prescriptive implications for strategic choices regarding production, in terms of where to allocate investment expenditures, and also regarding sales and marketing, in particular which geographic markets and customer segments warrant the most attention. These strategic implications in turn hinge not just on whether distance (broadly defined) matters, but *why* distance has an impact on costs or revenues.

In this paper, we take up the question of whether and why distance and differences matter, in the context of online trade. We show that – despite the presence of numerous mechanisms to safeguard transactions and the relative opacity of sellers' identities – online buyers in the United States purchase more from sellers in locations that have populations with similar cultural characteristics. Trade between culturally similar populations could arise from more closely overlapping product preferences *or* from greater trust in or affinity for within-group transactions.[1] The latter considerations are motivated by the large body of research on in-group favoritism dating back at least to Tajfel, Billig, Bundy, and Flament (1971), and more recently work which shows that individuals place greater trust in, and are more apt to cooperate with, others

[1] 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.

with whom they share common traits (e.g., Foddy, Platow, and Yamagishi 2009, Buchan, Croson, and Dawes 2002, DeBruine 2002).[2] After documenting the higher level of trade, as captured by transaction revenue and quantity, between buyers and sellers from locations with similar characteristics, we examine the degree to which these patterns are plausibly driven by greater in-group trust. We show that overall trade demonstrates patterns consistent with this interpretation. Finally, to examine whether in-group affinity (or out-group animosity) is a driving force in this online market, we examine buyer satisfaction conditional on trade. We find that, conditional on a transaction taking place, buyer satisfaction does not depend on the seller's location. We interpret this as suggesting that affinity or animus, of the sort considered by Becker (1957), are unlikely to be key factors driving our main results.

Our findings have strategic implications for a broad range of businesses. At its most basic level, our findings indicate that, even for relatively anonymous online merchants for whom one might anticipate cultural distance matters least, it continues to play a substantial role in influencing the ability to sell across geographies. Sellers in our dataset do not choose their locations or exposure to geographically dispersed buyers, which allows us to identify the effect of cultural distance as well as its underlying link to trust. But for commerce (online or otherwise) the link between differences and buyer trust is relevant for seller location decisions as well as where and how to focus marketing efforts. These findings reinforce the continued relevance of Ghemawat's (2007) "The World Isn't Flat" argument, which we show has relevance even in a within-country, online context in which we might expect these effects to be relatively modest. At a more granular level, by examining an underlying friction generated by cultural difference, i.e., mistrust, our research points toward remedies both for platforms that seek to foster exchange regardless of point of origin and for place-based organizations that must overcome disadvantages in selling to certain groups.

---

[2] 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, Laibson, Scheinkman, and Soutter (2000) find no statistically significant relationship between demographic similarity and trust.

Turning to a more detailed overview of our analysis and findings, we utilize a dataset comprising nearly all eBay transactions inside the US during 2015 and 2016. Key features of eBay allow us to isolate and study a subset of potential drivers of trade flows. 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.[3] 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 frequently focused on trade between countries (e.g., Guiso, Sapienza and Zingales 2009).

Specific features of the eBay setting also allow us to examine buyer-seller trust as an underlying driver of trade. In particular, because eBay has several well-established (and well-documented) mechanisms for identifying trustworthy sellers, we can explore how socioeconomic and cultural differences 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 evaluate whether greater cultural or socioeconomic difference leads to perceived advantages from dealing with more trustworthy counterparties, in addition to reflecting overlapping product preferences.

As proxies for cultural difference (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

[3] 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, for the most part, 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 inferences buyers make about socioeconomic and cultural proximity or distance with sellers come primarily from their beliefs about the expected characteristics of sellers at the designated location.

range of socioeconomic characteristics, including income, urban share, home values, and average age, as well as geographic distance. Further, by focusing on within-category trade between states, we may absorb at least some of the broad, unobserved differences across geographies that may be distinct from – but correlated with – cultural or other attributes. All specifications also include fixed effects to absorb source- and destination-specific features. For example, for our category analysis, these include buyer-state×category×year and seller-state×category×year fixed effects, to account for differences in the scale of trade in particular product categories originating in and destined for particular regions.[4]

As a starting point, with our state-pair data aggregated across all categories, we find that, for all three "culture" measures, greater cultural difference between 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 trade.[5] Or, to be more concrete, if the ethnic similarity of Oklahoma and New York (the 75th percentile of ethnic difference) were as similar as that of Massachusetts and Ohio (the 25th percentile difference), trade between these states would be 11 percent greater. In supplementary analysis, we demonstrate that our results also hold at the 3-digit zip-code pair level, which captures 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. We also verify our results in between-state within-category trade, calculating overall effects that are comparable to the aggregate analysis, and document how cultural differences have heterogeneous effects across categories.

We further aim to positively identify the role of trust (as opposed to overlap in tastes) in these more fine-grained analyses, by further disaggregating trade into high- versus low-reputation sellers. Our test is based on the insight that, if social and cultural similarity leads to greater trust, then quality-assurance

---

[4] This amounts to 6732 fixed effects (51 states, 33 categories, 2 years, for seller and buyer locations).

[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).

mechanisms – which serve as an alternative source of trust provided by the platform itself – 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 provides this 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 measures are substantially diminished for eTRS listings. We find a similar pattern when we separate sellers by their level of buyer-provided feedback,[6] which reinforces the notion that cultural similarity is especially important when sellers do not yet have track records to reassure buyers of their reliability.

Finally, we examine the relationship between cultural similarity and buyer satisfaction. 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, and Jolls 2015; Doleac and Stein 2013; Kricheli-Katz and Regev 2016), 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 of cultural difference; however, we find no such association.[7] Rather, we find that, conditional on trade, the

[6] 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 an early survey and Hui, Saeedi, Shen, and Sundaresan (2016) for more recent work. The literature has generally concluded that buyers see greater risk in dealing with sellers with low feedback scores, because their trustworthiness is more uncertain.

[7] Kricheli-Katz and Regev (2016) show that female sellers receive fewer bids than their male counterparts in a large sample of eBay transactions. In an accompanying survey, they show that respondents can often correctly infer the seller's gender based on information available on a listing, suggesting that buyers can and do attend to seller attributes. Ayres, Banaji, and Jolls (2013) show racial discrimination based on an experimental design with a relatively small sample – the design, with black versus white hands holding the item for sale, ensures that buyers can readily infer the seller's race. Our paper is distinct from these studies in our research focus and the type of data we use. While the aforementioned papers study one-way discrimination (e.g., the bias experienced by all black sellers, with no conditioning on buyer race), we examine the effects of differences between populations. And while we do not exploit individual sellers' attributes, our sample's breadth allows us to describe a large share of all trade on eBay within the United States.

relationship between our measures of buyer satisfaction and cultural similarity is not economically significant.[8] We recognize, however, that this does not fully rule out the possibility that buyer selection adjusts trade volume in a precise way across state-pairs, so that consummated trades might capture only the instances in which buyers find sufficiently appealing or trustworthy sellers.

Our work contributes first and foremost to the literature that examines the factors that affect the heterogeneity in performance of e-commerce platforms. 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).[9] 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 factors that impact the ability of e-commerce platforms to generate successful transactions and generally focus on *either* seller attributes (such as charitable contributions, as in Elfenbein, Fisman, and McManus 2012, or programmer satisfaction scores such as Gu and Zhu 2019), buyer attributes (such as experience, as in Perez-Truglia 2017), or platform design choices as the drivers of trust (Elfenbein, Fisman, and McManus 2015; Hui, Saeedi, Shen and Sundaresan 2016). By contrast, we emphasize the importance of the buyer-seller match, which implies a difference in (perceived) opportunism among culturally proximate trading partners, rather than simply the attributes of the seller or attributes or the buyer, and we use features of the platform to examine the mechanisms that define the value of this match. Wang and Overby (2020), which studies online

[8] 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, they are no less likely to betray facially similar partners.

[9] Lendle, Olarreaga, Schropp, and Vezina (2015) also document that geographic distance's impact is relatively small on eBay, focusing on international transactions on the platform versus other trade flows.

lending, find related evidence that lenders use similarities in political ideology as a proxy for borrowers' creditworthiness. Chintagunta and Chu (2021) focus on explaining asymmetries in online trading between locations in China, and find that patterns of trade are correlated with dyadic measures of trust as well as other factors. Also related to our paper, Bailey, Cao, Kuchler, Stroebel, and Wong (2018) examine social media friendships on platforms such as Facebook; they document that trade – online and otherwise – between counties correlates with measures of social connectedness. Our work explores similar cultural determinants of commercial interactions, but allows us to distinguish the role of trust in particular, as distinct from simply an overlap in product interests.[10]

2. Data

We use two 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 level, between pairs of 3-digit zip codes, as well as state-to-state trade disaggregated by eBay's 33 top-level categories. The second is demographic data, drawn from a variety of sources we describe below, which capture cultural and socioeconomic characteristics of individual areas. In the main body of this paper, we present analysis on trade between states; our description and analysis of the zip-code-level data are in the appendix.

2.1 eBay data

[10] 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.

Our eBay data come from the firm's U.S. platform, which hosted over $37.5 billion worth of transactions during 2020.[11] 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 some 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. The eTRS badge, if earned, is displayed automatically on qualified sellers' listings. 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) 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 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 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,

[11] eBay US platform revenue from https://www.digitalcommerce360.com/article/ebays-sales/.

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.

Our eBay data comprise a comprehensive record of transactions that occurred between January 2015 and December 2016, aggregated to the year level. To illustrate how the data are constructed, in the description that follows we provide an overview of the state-pair data; a near-identical process is used to generate zip code and category×state 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.[12] We define *High feedback* to denote a seller that has

[12] 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.

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.[13] 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.[14] 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 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's Panel A. 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 of sellers with feedback above 200. The average EPP_{bst} value across state dyad-year combinations is just over one half, and the

[13] 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.

[14] This measure relies on the observation that it is more 'expensive' to leave a negative review than it is to leave a positive one, because sellers may retaliate against or harass the buyer. A basic assumption of this approach, as described by Tadelis (2016: 334), is that "silence is bad news."

mean negative feedback rate is 0.36%.[15] While all state-to-state annual aggregate trade is strictly positive, about 2% of all state-category-pairs have zero trade in a year. When we further divide state-category pairs by eTRS or feedback levels, we find 11% and 3%, respectively, of observations record zero trade. (The trade between zip codes is summarized in Appendix Table A1 Panel A.)

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.

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 in the appendix. 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, which we collected from the Federal Election Commission. We use votes cast in each state for the two major parties to calculate the national winner's fraction of these votes at the state-level; we label this variable *Winner vote share*.[16] In Table 1's Panel B,

---

[15] 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.

[16] 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

we report 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 1 Panel C, 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.

We report several additional measures at the state dyad level in Table 1 Panel C. *Distance* is the "shortest curve" distance (in kilometers) between the states' centers of population, which we obtain from the US Census.[17] The next set of variables capture two further dimensions of cultural similarity between states (in addition to political differences based on 2016 voting behavior). To generate the variable *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. The variable we construct is analogous to a between-group measure of ethnic fractionalization (see, e.g., Alesina, Devleeschauwer, Easterly, Kurlat, and Waczaiarg 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.[18] 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 value 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

---

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.

[17] 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.

[18] 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.

faiths into five main categories, with the residual state population categorized as non-adherents.[19] Appendix Table A1 contains summary statistics at the zip-code level that correspond to the state-level statistics in Table 1.

The summary statistics in Table 1 Panel C are raw difference measures. However, to make the effect sizes more easily comparable across difference measures in our empirical analysis in the next section, we normalize all cultural and demographic difference variables to have a mean of zero and standard deviation of one.

3. Empirical Analysis and Results

3.1 Aggregate state- level analyses

Our aggregate state-level specification takes the form:

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

where C is a vector of cultural difference variables, X a vector of socioeconomic difference variables, $\log(Distance)$ is the logarithm of the distance between the two states' centers of population, and u and v are buyer-state × year and seller-state × year fixed effects, respectively. $Trade_{bst}$ is the revenue value or quantity traded between seller state s and buyer state b during year t.[20] We primarily focus on revenue-based measures of $Trade_{bst}$, and for robustness we also present findings based on quantity-based measures. In our main analysis we estimate (1) using OLS, but we also perform the analysis using Poison pseudo-

[19] 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.

[20] To deal with zeros in revenue and quantity values, we add one to all *Trade* variables before applying the log transformation in (1) and all other linear regression models. As we note above, zeros appear in the trade data in some instances when we deal with zip-code level activity or state-level activity disaggregated by category and/or seller reputation.

maximum likelihood (PPML) specifications, as suggested by Silva and Tenreyro (2006), which use the levels of traded revenue and quantities as the dependent variables, rather than log transformations of them.[21] The variables we use to capture cultural 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 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.

We begin by presenting our baseline state-pair trade findings in Table 2, for both revenue- and quantity-based measures of trade. 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 difference measures are predictive of trade, in the expected directions: the coefficients on the ethnic, political, and religious difference 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

[21] 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. While our dataset has very few zeros, the latter two issues are applicable to the present work, so we include PPML specifications in each of our reported estimates of (1).

of trade in column 3: differences in median age, male share, urban share, and home values. 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).

As expected, all regressions produce coefficients on log(*Distance*) that are negative, economically important, and precisely estimated. Our estimated coefficients are larger, but of a similar order of magnitude, to the estimates produced by Hortacsu, Martinez-Jerez, and Douglas (2009) for eBay trade. (By construction, however, our analysis differs from Hortascu et al. (2009) as we exclude within-state transactions, which Hortascu et al. (2009) show are particularly important for certain categories of transactions, like event tickets.)

The specification in column 4 includes 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 based on total transaction value as the outcome.

In the remaining columns, we use *Summed cultural difference,* an overall measure of cultural similarity which sums our normalized measures of ethnic, religious, and voting differences; the new variable is mean zero with a standard deviation of 1.96. This summary measure facilitates a comparison of the role of cultural difference across product categories, which we turn to in the next subsection. For both revenue (column 5) and quantity (column 6) this overall cultural difference measure is a highly significant predictor of trade.

Finally, in columns 7 and 8 we present PPML specifications for state-pair annual transaction revenue and quantity, respectively. 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 A2 Panel A, we report estimates from PPML specifications which include the three cultural difference variables rather than *Summed cultural difference.*

We next illustrate the culture-trade relationship graphically, using binned scatter plots (Hao et al. 2010) to display the relationship between *Summed cultural difference* and trade. 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 2, column 3), including buyer-state × year and seller-state × 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 difference and interstate trade.[22] The relationship is roughly linear, and it is in line with the estimates presented in Table 2. In Appendix Figures A1 – A3, we present binned scatterplots for each individual cultural difference variable, generated while including all other controls from Table 2, 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.

3.2 Category-level heterogeneity in the role of cultural similarity

A primary candidate explanation for the preceding results is that a pair of geographic areas that differ in cultural and socioeconomic characteristics also suffer from a mismatch in what one area's consumers want and what is available from other area's sellers. For example, California buyers may not purchase clothing items frequently on eBay relative to other products, while Iowa sellers' product mix features a greater share of clothing items than other states' sellers. To address the role of this "inventory-taste overlap," we estimate a version of (1) that examines state-to-state trade within product categories. In terms of the previous example, this analysis investigates whether California clothing buyers favor clothing from sellers in states more demographically similar to themselves. We implement this analysis by replacing the state-year fixed effects u_{bt} and v_{st}, with buyer-state × category × year and seller-state × category × year fixed effects. Our results, which are in Table 3, are quite similar to those in Table 2, suggesting that while category-level

[22] 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.

considerations – tastes or otherwise – may partially explain the relationship between trade patterns and cultural similarity, they do not fully explain it. Appendix Table A2 Panel C contains additional results on PPML estimates within product category.

While Tables 2-3 show the relationships between trade and cultural and socioeconomic differences, averaged across all products in eBay's US market, cross-category differences may exist in the strength of these relationships, and this between-category variation may provide some insight on what drives the cultural similarity effect. In Figure 2, we present 33 separate category-level estimates of the relationship between *Summed Cultural Difference* and the revenue value of state-to-state trade.[23] The biggest effect is for Tickets and Experiences. We note that this is related to – but distinct from – the finding (which we also observe, in unreported results) that physical proximity also plays a particularly prominent role in this category (Hortacsu, Martinez-Jerez, and Douglas 2009). The categories in which cultural difference plays a prominent role also include what one might think of as 'cultural' objects (i.e., those in which tastes plausibly differ across groups) – art, antiques, pottery, memorabilia (again, these patterns are distinct from any role of physical proximity). At the other extreme, cell phones and computer equipment have less obvious taste differences, though the third-smallest effect size is for music, which one might assert has a substantial taste component. Of course, these across-category differences may also be driven by differences in the importance of trust. Plausibly, there is a more substantial leap of faith in buying art or antiques relative to more standardized products like phones and computers.

3.3 Quality assurance, trust, and the impact of cultural differences on trade

To affirmatively assess whether cultural closeness facilitates trust, we now take advantage of formal certification mechanisms that we argue serve as substitutes for (informal) trust. Specifically, we first analyze whether trade volume varies with whether the seller has earned a certification label that may signal trustworthiness, namely the eTRS designation. Intuitively, little trust is required by the buyer to purchase from a seller that has a certified track record for quality and promptness. Similarly, we posit that questions

[23] We find a similar pattern when we replace state-to-state trade revenue with quantity traded.

of trustworthiness are less salient for evaluating sellers with more feedback. To test whether cultural similarity matters less when dealing with more-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 4 contains results from three models based on our category-state-pair data. For the sake of comparison, we first reproduce Table 3 column 3 as column 1 in Table 4. In the next pair of columns (labeled 2a and 2b), we report results from a single model that includes two observations per year and category-state-pair: one for aggregate sales by non-eTRS sellers and the other for aggregate sales from eTRS sellers, with sales measured as the log of sales revenue. Throughout, we include seller-state × category × year × eTRS as well as buyer-state × category × year × eTRS fixed effects. We allow measures of socioeconomic differences 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.[24] For ease of interpretation, we focus on *Summed Cultural Difference* as our measure of the cultural difference between pairs in each case. Table 4, column 2a shows the coefficient estimates for the baseline 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, providing estimates from a single model in which we use *High feedback* rather than *eTRS* to divide sellers by their performance records.

Our primary interest is in the interaction of *eTRS* and *Summed Cultural Difference*. The coefficient estimate we report in column 2a (-0.034) provides the baseline effect of cultural difference on sellers

[24] 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.

without eTRS certification. Column 2b's corresponding estimate (0.014, significant at the 5% level) implies that the role of cultural difference is negative but about 40% lower (0.014/0.034) in magnitude for *eTRS* listings relative to non-*eTRS* ones. The eTRS interactions with the socioeconomic variables are of inconsistent sign and an F-test of their sum has a *p*-value of 0.73. 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 are broadly consistent with the message from our eTRS-based analyses in 2a and 2b. In particular, the joint effect of cultural differences is smaller for *High feedback* sellers, though the difference is significant only at the 10% level. Curiously, the estimates for socioeconomic differences for *eTRS* versus non-*eTRS* sellers are generally negative (i.e., the "wrong" sign); an F-test indicates that their sum is significantly different from zero at the 5% level.

Collectively, we interpret the estimates in Table 4 as suggesting that cultural similarity supports trust, so that cultural difference 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 and with PPML estimation. These sets of analyses show that cultural variables' impact on trade is consistently attenuated for eTRS-badged and high feedback sellers. The magnitudes are quite comparable to those reported in Table 4, though the PPML estimation provides much greater precision in our point estimates. These results are in Appendix Tables A4-A6.

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 difference 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. We investigate the satisfaction of buyers by analyzing the volume and content of feedback they provide for sellers. Conditional on a 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. This raises a concern about selection: a buyer must be willing to purchase from a seller in order to have the opportunity to leave feedback. We cannot capture the sentiments of buyers from state *b* who never purchase from sellers of state *s*, but we observe that trade often does occur between dissimilar locations, so if there is a tendency to give negative feedback, it may appear (although attenuated) in the feedback results. Some trade between dissimilar locations is likely to occur because of 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.

We provide results on buyer satisfaction, represented by effective percent positive (*EPP*) and fraction negative feedback (*Negative feedback*), derived from models that are variants of Table 3's specifications which include state-category-year fixed effects. One strength of *EPP* is that it removes concerns about an additional form of selection in buyer satisfaction: the decision of whether to leave feedback. *EPP* measures buyer satisfaction through the feedback choice itself, relying on the regularity that virtually all eBay feedback is positive (Nosko and Tadelis, 2015). *Negative feedback* is only observed when a buyer chooses to leave feedback, but it provides a sharper view of buyer satisfaction with individual transactions. We view the two feedback measures as providing complementary evidence on buyer satisfaction. In Table 5, the odd-numbered specifications include the three separate cultural difference measures, and the even specifications include *Summed Cultural Difference*. We find that 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 no case is any measure of cultural difference (individual or in aggregate) a significant predictor of either proxy for customer satisfaction. Overall, 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, Coffman, Gennaioli, and Schleifer 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 difference – 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. 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 the negative association between cultural difference and trade varies significantly across product categories – for example, the weakest association was found for cell phones and video games, and the strongest for tickets, art, and antiques. We conjecture that, for the latter set of products, quality and/or authenticity are less easily verified. Consistent with this pattern, our third contribution is to use the differences in patterns of trade with certified versus non-certified sellers to show that cultural similarity matters more when there is more uncertainty about the *quality of the seller*, suggesting that buyers associate cultural similarity with greater trustworthiness. Nonetheless, our concluding examination of customer satisfaction, raises questions about whether these perceptions of trustworthiness are accurate or driven by mistaken beliefs about the impact of buyer-seller differences. Collectively, these findings indicate a positive link between cultural similarity and trade that is at least partly driven by trust concerns. These findings contribute to a small but growing body of work that examines the relationship between cultural similarity and trade, such as Guiso, Sapienza, and Zingales (2009), Bailey et al. (2018), Chinagunta and Chu (2021), and Wang and Overby (2021).

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 difficult 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.

4.1. *Implications for Business Strategy*

Our findings suggest that the salience of location – beyond the choice of location itself – may itself be a strategic choice. In our setting, sellers may wish to make location more salient to buyers (e.g., in product descriptions) depending on desired target markets, or indeed may wish to highlight location if they are endowed with a high-trust location. In considering the role of place in firm strategy more broadly, we see evidence of such efforts at tethering brands to location both in businesses that use place names to differentiate themselves (e.g., Bayerische Motoren Werke, OshKosh B'Gosh, Tom's of Maine) and in those that use location in advertising (e.g., "Miller Brewing Company – Milwaukee, Wisconsin" or "Ben & Jerry's – Vermont's Finest"). The empirical patterns we uncover also emphasize the importance to new or small firms of understanding that they may be judged by prospective consumers, at least in part, based on their location.

For early-stage entrepreneurs, our results point toward important considerations that influence product-market fit (Eisenmann, Ries, and Dillard 2011), by highlighting product categories where seller location influences buyer decision-making and suggesting mechanisms, like quality assurance, that moderate the influence of inferences based on location. These insights can offer guidance to entrepreneurs engaged in experiment-based approaches to developing business models (Koning, Hasan, and Chatterji 2019, Carmuffo, Cordova, Gambardella, and Spina 2020), by providing informative priors (e.g., Chen, Croson, Elfenbein, and Posen 2018) or shaping the development of high-precision tests (Agarwal, Gans, and Stern 2021). A startup in the Bay Area may need to work harder to earn the trust of customers in Bakersfield than Brooklyn, despite the fact that Bakersfield is considerably closer.

For platform operators, our results 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 cultural distance in economic transactions. In the absence of complete quality assurance or reputational assurances, however, our results suggest that location choice may significantly shape the market available to a firm, especially in their early stages. More generally, gains from trade are likely to be greater and more widespread if firms are aware of these potential frictions and can work to overcome them.

REFERENCES

Alesina A, Devleeschauwer A, Easterly W, Kurlat, S and Wacziarg R. 2003. Fractionalization. *Journal of Economic Growth*, 8(2): 155-194.

Agrawal A, Gans JS and Stern, S. 2021. Enabling entrepreneurial choice. *Management Science*.

Anderson JE and Van Wincoop E. 2004. Trade costs. *Journal of Economic Literature*, 42(3): 691-751.

Ayres I, Banaji M and Jolls C. 2015. Race effects on eBay. *The RAND Journal of Economics*, *46*(4), pp.891-917.

Bailey M, Cao RR, 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.

Bolton G, Greiner B and Ockenfels A. 2013. Engineering trust: reciprocity in the production of reputation information. *Management Science*, *59*(2), pp.265-285.

Bordalo P, Coffman K, Gennaioli N and Shleifer A. 2016. Stereotypes. *The Quarterly Journal of Economics*, 131(4): 1753-1794.

Brynjolfsson E, Hui X and Liu M. 2019. Does machine translation affect international trade? Evidence from a large digital platform. *Management Science*. 65(12): 5449-5460.

Buchan NR, Croson RT and Dawes RM. 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 NR, Johnson, EJ, and Croson RT. 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.

Camuffo A, Cordova A, Gambardella A, and Spina, C. 2020. A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. *Management Science*, *66*(2), pp.564-586.

Chen JS, Croson DC, Elfenbein DW, and Posen HE. 2018. The impact of learning and overconfidence on entrepreneurial entry and exit. *Organization Science*, *29*(6), pp.989-1009.

Cui R, Li J, and Zhang DJ. 2016. Discrimination with incomplete information in the sharing economy: Field evidence from Airbnb. Washington University in St. Louis working paper.

Chintagunta PK and Chu J. 2021. Geography as branding: Descriptive evidence from Taobao. *Quantitative Marketing and Economics*. 19: 53-92.

DeBruine LM. 2002. Facial resemblance enhances trust. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1498): 1307-1312.

Doleac JL and Stein LC. 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.

Eisemann T, Ries E, and Dillard S. 2011. *Hypothesis-driven entrepreneurship: the lean startup*. Harvard Business School.

Elfenbein DW, 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 DW, 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.

Fisman RJ. 2003. Ethnic ties and the provision of credit: Relationship-level evidence from African firms. *Advances in Economic Analysis & Policy*, *3*(1).

Foddy M, Platow MJ and Yamagishi T. 2009. Group-based trust in strangers: The role of stereotypes and expectations. *Psychological Science*, 20(4): 419-422.

Friedman TL. *The World Is Flat: A Brief History of the Twenty-First Century*. New York: Farrar, Straus, and Giroux, 2005.

Ghemawat P. 2001. Distance still matters. The hard reality of global expansion. *Harvard Business Review*, 79(8): pp.137-147.

Ghemawat P. 2007. Why the world isn't flat. *Foreign Policy* 159 (March/April): 54-60.

Glaeser EL, Laibson DI, Scheinkman JA and Soutter CL. 2000. Measuring trust. *The Quarterly Journal of Economics*, 115(3): 811-846.

Gu G and Zhu F. 2019. Trust and disintermediation: Evidence from an online freelance marketplace. *Management Science.* 67(2): 794-807.

Guiso L, Sapienza P and Zingales L. 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 MC, Dayal U, Sharma RK, Keim DA 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.

Kricheli-Katz T and Regev T. 2016. How many cents on the dollar? Women and men in product markets. *Science Advances*, 2(2): DOI: 10.1126/sciadv.1500599.

Koning R, Hasan S and Chatterji A. 2019. *Experimentation and startup performance: Evidence from A/B testing* (No. w26278). National Bureau of Economic Research.

Lendle A, Olarreaga M, Schropp S and Vezina P. 2015. There Goes Gravity: eBay and the Death of Distance. *Economic Journal* 126: 406-441.

Nosko C and Tadelis S. 2015. The limits of reputation in platform markets: An empirical analysis and field experiment. National Bureau of Economic Research working paper w20830.

Perez-Truglia, R. 2018. Markets, trust and cultural biases evidence from eBay. *Journal of Behavioral and Experimental Economics* 72: 17-27.

Silva JMCS and Tenreyro S. 2006. The log of gravity. *Review of Economics and Statistics* 88(4): 641-658.

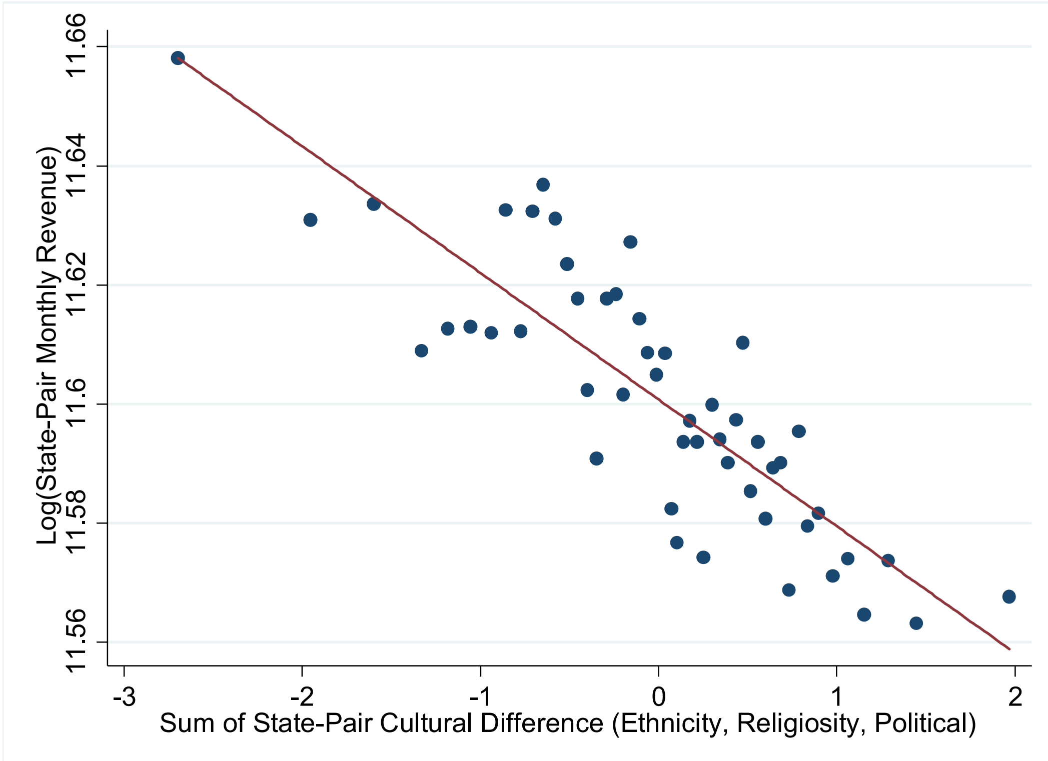
Tadelis S. 2016. Reputation and feedback systems in online platform markets. Annual Review of Economics, 8, pp.321-340.

Tajfel H, Billig MG, Bundy RP and Flament C. 1971. Social categorization and intergroup behaviour. *European journal of social psychology*, 1(2): 149-178.

Wang H and Overby EM. 2021. Do political differences inhibit market transactions? An investigation in the context of online lending. Working Paper, Georgia Institute of Technology.

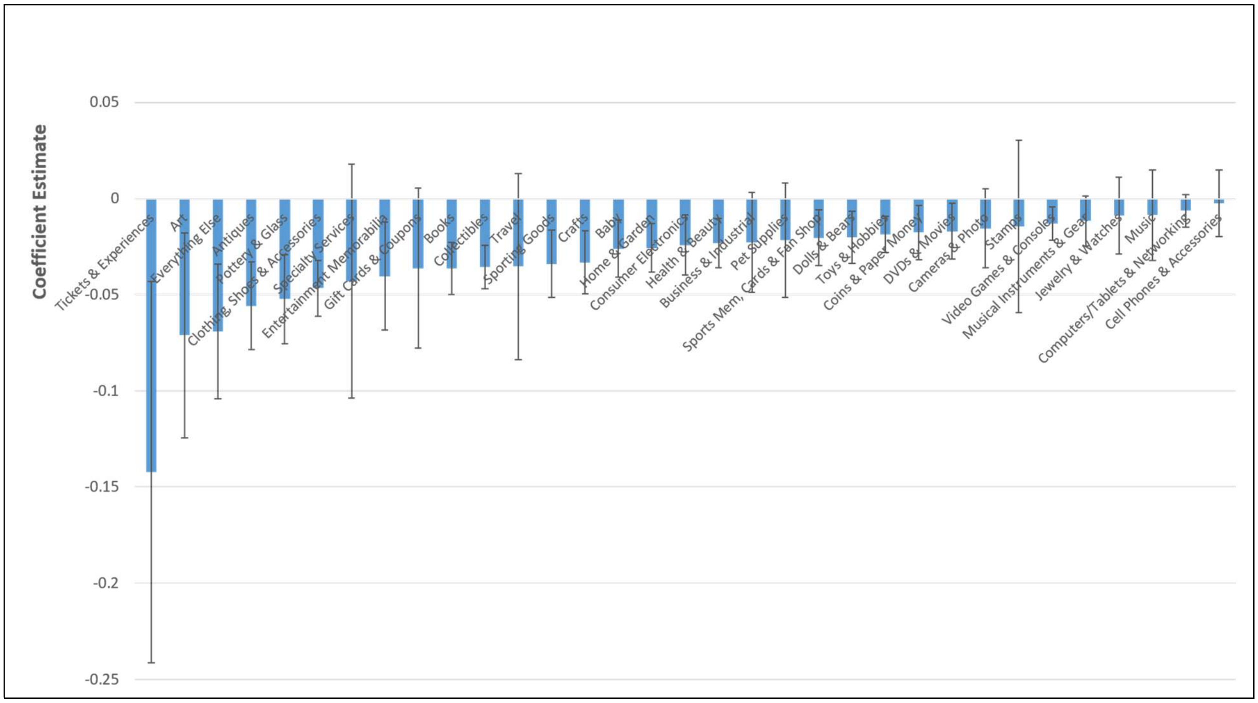
Zussman A. 2013. Ethnic discrimination: Lessons from the Israeli online market for used cars. *The Economic Journal*, 123(572): F433-468.

Figure 1: Relationship between cultural difference 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 difference: ethnicity, religiosity, and voting. We use 50 bins, and residualize the data to account for all control variables included in Table 2, column 3.

Figure 2. Coefficient estimates on sum of difference measures, by product category, with 95% confidence intervals, state-to-state trade data, log (value) specification



Note: Y-axis is the estimated coefficient on *Summed Cultural Difference* from regressions at the category level including all controls in Table 2 column (2) and buyer-state × year and seller-state × year fixed effects. The dependent variable is log of the total value of trade between states. 95% confidence intervals are shown using standard errors that are two-way clustered by buyer state and seller state.

Table 1: Annual sales and feedback data summary

	Mean	Median	SD	Min	Max
Panel A: Annual sales and feedback data summary, state-to-state transactions (N = 5,100)					
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
Panel B: State characteristics (N = 51)					
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
Panel C: Differences in state characteristics (N = 2550)					
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
Difference in …					
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 states (km)	1964.91	1600.78	1468.42	31.73	8226.99

Note: State data includes all 50 states and the District of Columbia. For variable definitions, see paper text.

Table 2: The impact of cultural and socioeconomic differences on interstate trade

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	OLS	OLS	OLS	OLS	OLS	OLS	PPML	PPML
Dependent variable	log(revenue)	log(revenue)	log(revenue)	log(quantity)	log(revenue)	log(quantity)	Revenue	Quantity
Ethnic difference	-0.110***		-0.0798***	-0.0635***				
	(0.0206)		(0.0167)	(0.0123)				
Religious difference	-0.0286**		-0.0327**	-0.0410***				
	(0.0129)		(0.0136)	(0.0101)				
Voting difference	-0.0403***		-0.0262***	-0.0120**				
	(0.00801)		(0.00631)	(0.00491)				
Summed cultural difference					-0.0332***	-0.0245***	-0.0223***	-0.0212***
					(0.00461)	(0.00333)	(0.00633)	(0.00441)
Median income difference		0.0122*	0.00980	0.00140	0.00831	-0.00161	0.00743	0.00308
		(0.00715)	(0.00713)	(0.00540)	(0.00659)	(0.00532)	(0.00760)	(0.00742)
Bachelors share difference		-0.0210**	-0.0133	-0.00886	-0.00820	-0.00214	0.00154	0.00242
		(0.00991)	(0.00992)	(0.00642)	(0.00987)	(0.00638)	(0.00887)	(0.00659)
Median age difference		0.0104*	0.0119**	0.00741	0.0131**	0.00806	0.00612***	0.00250
		(0.00576)	(0.00532)	(0.00521)	(0.00512)	(0.00564)	(0.00232)	(0.00349)
Male share difference		-0.0257***	-0.0237***	-0.0108**	-0.0258***	-0.0107**	-0.0139	-0.00506
		(0.00859)	(0.00797)	(0.00450)	(0.00786)	(0.00494)	(0.00942)	(0.00781)
Median home value difference		-0.0340**	-0.0197	-0.00789	-0.0168	-0.00346	-0.0111	0.000547
		(0.0160)	(0.0156)	(0.00826)	(0.0159)	(0.00874)	(0.00826)	(0.00744)
Urban share difference		-0.0261***	-0.0117*	-0.00866*	-0.0143**	-0.0105**	-0.0224**	-0.0213**
		(0.00715)	(0.00653)	(0.00462)	(0.00668)	(0.00491)	(0.0107)	(0.0109)
Owner-occupied share difference		-0.0270***	-0.0152**	-0.0141***	-0.0231***	-0.0200***	-0.0245***	-0.0187***
		(0.00614)	(0.00613)	(0.00437)	(0.00573)	(0.00438)	(0.00749)	(0.00469)
Log distance between states	-0.146***	-0.145***	-0.134***	-0.115***	-0.136***	-0.120***	-0.0970***	-0.0891***
	(0.00955)	(0.0103)	(0.00919)	(0.00754)	(0.00947)	(0.00812)	(0.0102)	(0.00673)
Observations	5,100	5,100	5,100	5,100	5,100	5,100	5,100	5,100
R-squared (within)	.5702	.5605	.5873	.6308	.583	.623	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 × year and seller-state × 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 3: Interstate trade within product categories

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	OLS	OLS	OLS	OLS	PPML	PPML
Dependent variable	log(revenue)	log(quantity)	log(revenue)	log(quantity)	Revenue	Quantity
Ethnic difference	-0.0860***	-0.0443***				
	(0.0177)	(0.0117)				
Religious difference	-0.0548***	-0.0385***				
	(0.0121)	(0.00737)				
Voting difference	-0.0143**	-0.0113**				
	(0.00653)	(0.00446)				
Summed cultural difference			-0.0317***	-0.0215***	-0.0189***	-0.0193***
			(0.00500)	(0.00347)	(0.00484)	(0.00322)
Median income difference	0.00750*	0.00148	0.00329	-0.00106	-0.000513	-0.000586
	(0.00425)	(0.00308)	(0.00430)	(0.00364)	(0.00620)	(0.00514)
Bachelors share difference	-0.0136	-0.000223	-0.00427	0.00478	0.00538	0.00276
	(0.0103)	(0.00397)	(0.0102)	(0.00435)	(0.00663)	(0.00495)
Median age difference	0.00546	0.00165	0.00637	0.00188	0.00482***	0.00268*
	(0.00629)	(0.00416)	(0.00655)	(0.00481)	(0.00180)	(0.00144)
Male share difference	-0.00794	-0.0167**	-0.00775	-0.0158**	-0.00721	-0.00471
	(0.00685)	(0.00703)	(0.00691)	(0.00681)	(0.00620)	(0.00511)
Median home value difference	-0.0263**	-0.0152**	-0.0202*	-0.0117*	-0.00499	-0.00138
	(0.0103)	(0.00621)	(0.0106)	(0.00673)	(0.00595)	(0.00527)
Urban share difference	0.00981	-0.0121***	0.00721	-0.0131***	-0.0160**	-0.0120*
	(0.00586)	(0.00389)	(0.00610)	(0.00401)	(0.00727)	(0.00651)
Owner-occupied share difference	-0.00889	-0.00427	-0.0171**	-0.00746**	-0.0181***	-0.0129***
	(0.00627)	(0.00340)	(0.00689)	(0.00323)	(0.00447)	(0.00316)
Log distance between states	-0.158***	-0.125***	-0.165***	-0.129***	-0.0978***	-0.0895***
	(0.00879)	(0.00678)	(0.00929)	(0.00719)	(0.00693)	(0.00550)
Observations	168,300	168,300	168,300	168,300	168,300	168,300
R-squared (within)	0.0500	0.1157	0.0493	0.115	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 × product category × year and seller-state × product category × 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 4: The moderating role of seller quality in the impact of cultural difference on trade. Dependent variable: log(revenue)

	Model 1: All sellers	Model 2: Sellers separated by eTRS status		Model 3: Sellers separated by feedback level	
	(1)	(2a)	(2b)	(3a)	(3b)
		base effect, eTRS = 0	*add'l effect for eTRS = 1*	*base effect, High fdbk = 0*	*add'l effect for High fdbk = 1*
Summed cultural difference	-0.0317***	-0.0337***	0.0138**	-0.0346***	0.0109*
	(0.00500)	-0.00507	-0.00554	-0.00495	-0.00578
Difference in:					
[1] Median income	0.00329	0.00248	-0.00212	0.0043	-0.0072
	(0.00430)	-0.00476	-0.00554	-0.00723	-0.00572
[2] Bachelors share	-0.00427	-0.00289	-0.000704	-0.0105	0.0183***
	(0.0102)	-0.0104	-0.00807	-0.0116	-0.0055
[3] Median age	0.00637	0.00605	-0.00394	0.00184	0.0064
	(0.00655)	-0.00716	-0.0077	-0.00667	-0.00504
[4] Male share	-0.00775	-0.00491	-0.00379	-0.00574	0.00567
	(0.00691)	-0.00748	-0.00967	-0.00677	-0.00551
[5] Med. home value	-0.0202*	-0.0235**	0.0128	-0.0290***	0.0055
	(0.0106)	-0.00985	-0.00847	-0.00907	-0.00815
[6] Urban share	0.00721	0.00781	-0.0138*	0.00627	-0.00282
	(0.00610)	-0.00697	-0.00762	-0.00812	-0.00539
[7] Owner-occup. share	-0.0171**	-0.0190**	0.00597	-0.0206**	0.00668
	(0.00689)	-0.00834	-0.00776	-0.0101	-0.00804
Test of sum of [1]-[7] = 0	F = 4.63	F = 3.41	F = 0.12	F = 9.05	F = 4.46
	(p = 0.0362)	(p = 0.0707)	(p = 0.7345)	(p = 0.0041)	(p = 0.0396)
Log distance between states	-0.165***	-0.172***	0.00959	-0.196***	0.0334***
	(0.00929)	-0.00916	-0.00607	-0.00905	-0.00677
Observations	168,300	336,600		336,600	
R-squared (within)	0.0493	0.0312		0.0364	

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

Note: This table reports results from 3 models. The model with results in column 1 aggregates the impact of difference on all sellers. The second model's results are in columns 2s and 2b; this model provides baseline (2a) and eTRS-interacted (2b) effects of difference measures on trade. Model 3 has the same structure as model 2, but for seller feedback. In addition to the listed variables, the model in columns 2a and 2b includes buyer-state × *eTRS* × year and seller-state × *eTRS* × 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 5: The impact of cultural difference on customer satisfaction

VARIABLES	(1) EPP	(2) EPP	(3) Share neg. fdbk	(4) Share neg. fdbk
Ethnic difference	-0.0411		0.0340	
	(0.170)		(0.0236)	
Religious difference	-0.0861		0.0219	
	(0.0801)		(0.0150)	
Voting Difference	0.0750		-0.00412	
	(0.0609)		(0.00768)	
Summed cultural difference		0.0220		0.00637
		(0.0352)		(0.00686)
Median income difference	-0.0247	-0.0382	0.00234	0.00485
	(0.0436)	(0.0442)	(0.00865)	(0.00826)
Bachelors share difference	-0.0616*	-0.0392	-0.00577	-0.0110
	(0.0344)	(0.0407)	(0.0112)	(0.0121)
Median age difference	-0.0912**	-0.0917**	-0.00796	-0.00835
	(0.0423)	(0.0414)	(0.00879)	(0.00851)
Male share difference	0.0192	0.0283	-0.0104	-0.0109
	(0.0669)	(0.0639)	(0.00789)	(0.00789)
Median home value difference	0.0834	0.100	-0.0225*	-0.0262**
	(0.0758)	(0.0719)	(0.0128)	(0.0125)
Urban share difference	0.0725	0.0709	0.00927	0.0105
	(0.0521)	(0.0519)	(0.00826)	(0.00841)
Owner-occupied share difference	-0.0469	-0.0535	-0.0112	-0.00723
	(0.0429)	(0.0466)	(0.00932)	(0.00808)
Log distance between states	-0.422***	-0.445***	0.0288***	0.0329***
	(0.0698)	(0.0670)	(0.00609)	(0.00657)
Observations	166,606	166,606	165,008	165,008
R-squared (within)	0.000880	0.000863	0.000142	0.000121

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 × year and seller-state × 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

A.1: Supplementary analysis of trade between zip codes

A.1.1 Zip-code data

In supplementary analysis, we examine trade between 3-digit zip-code areas (e.g., the area "275" includes zip codes 27514, 27516, and others). There are 887 of these areas in the U.S. with civilian populations and complete demographic data. For these zip code areas, we analyze trade and demographic data that is similar in structure to our data on state-to-state trade. In Table A1 Panel A, we include summary statistics for the zip-to-zip annual transaction data, which has similar patterns to the state-to-state data in its trade and feedback values.

In constructing demographic data at the zip-code level, we use county-level analogues to the state-level data described in Section 3 together with a weighting procedure. 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.[25] We construct our zip code demographic variables by using these weights with the county-level versions of the demographic variables described above.[26] In Table A1 Panels B and C, we present summary statistics for zip-code-level and zip-code-pair attributes respectively.

A.1.2 Zip-code level analysis

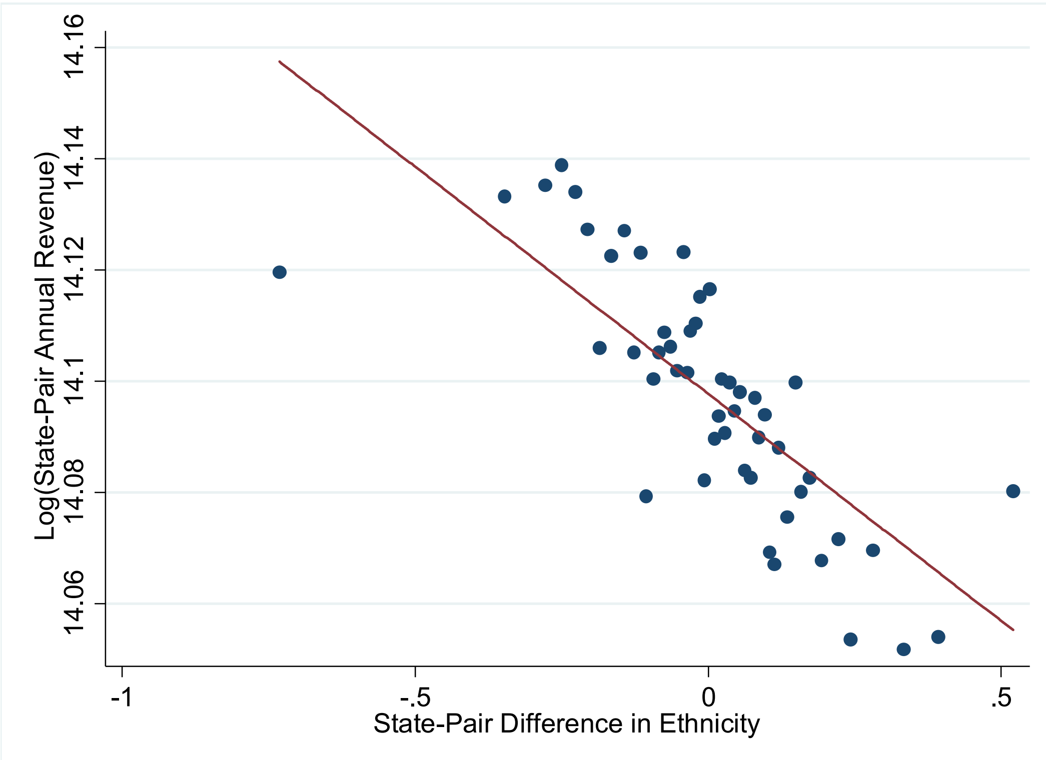
For our zip-code-level analyses, the specification is virtually identical to (1), with 3-digit zip codes defining b and s throughout. In Table A2, we present analogous results for zip-to-zip annual transaction data, which

[25] 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+0.2×100,000+0.6×200,000), and the counties' weights in z are their population shares of z's total population, e.g. 0.36 for county A.

[26] 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.

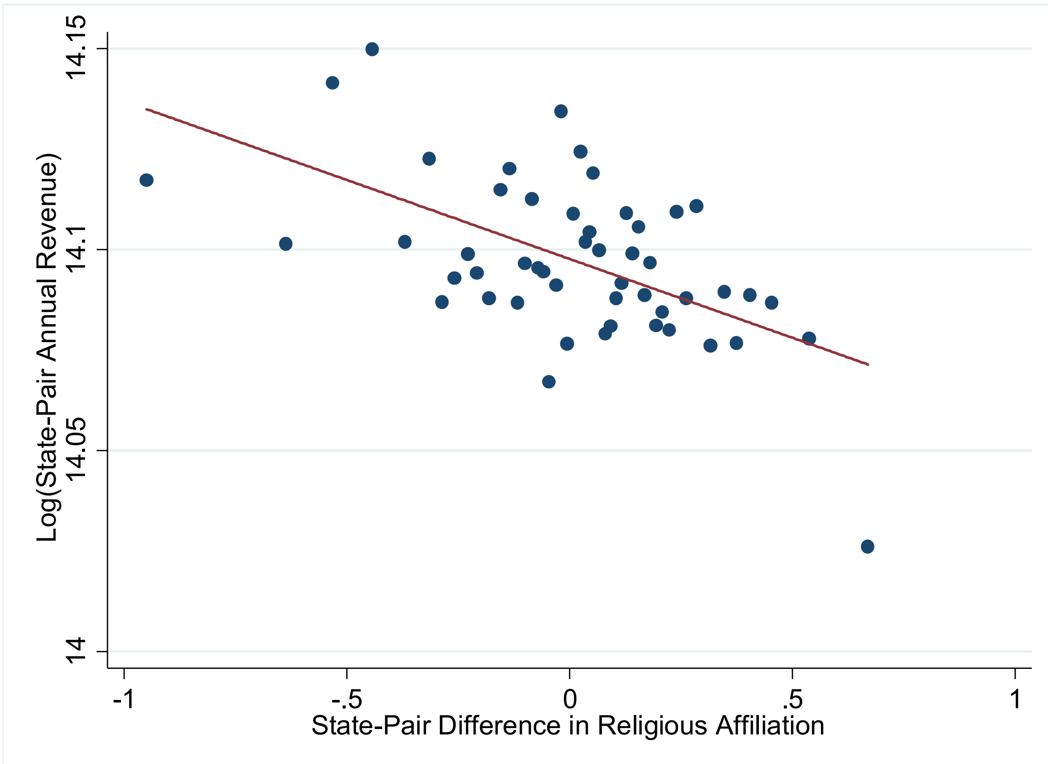
allows us to account for the very different perceptions that buyers may have of, say, Manhattan versus upstate New York, or Austin versus rural Texas. Column 1 shows a consistent and robust relationship between culture proxies and trade at the zip-code-pair level. In general, the magnitudes of the coefficients on *Religious difference* and *Voting Difference* are larger, and all variables are (unsurprisingly) more precisely estimated than in the state-to-state trade analyses. This specification exploits both cross- and within-state variation in estimating the culture-trade relationship. In column 2, 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. Turning to our quantity-based measure of trade in column 3 the results are again largely unchanged. We use our overall measure of cultural similarity (*Summed cultural difference*) in the remaining columns and observe a strong relationship between cultural similarity and trade whether we employ OLS (columns 4 and 5) or PPML (columns 6 and 7) specifications. See Appendix Table A3 Panel B for PPML results with the three separate cultural difference variables. In Appendix Figures A4-A6 we present binned scatter plots for each individual cultural difference variable's relationship with total trade based on our zip code-level data.

Figure A1: Relationship between ethnic differences and interstate trade



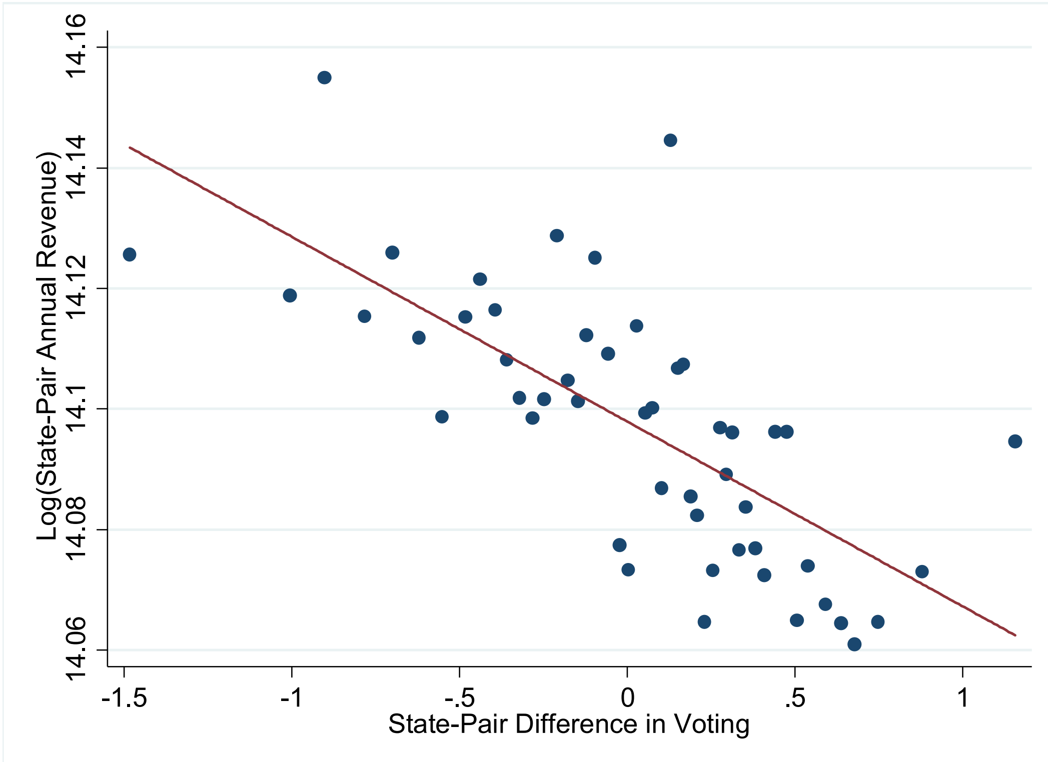
Note: The figure shows a binned scatter plot to illustrate the relationship between state-pair differences in ethnicity and interstate trade. We use 50 bins, and residualize the data to account for all other variables included in Table 2, column 3.

Figure A2: Relationship between differences in religiosity interstate trade



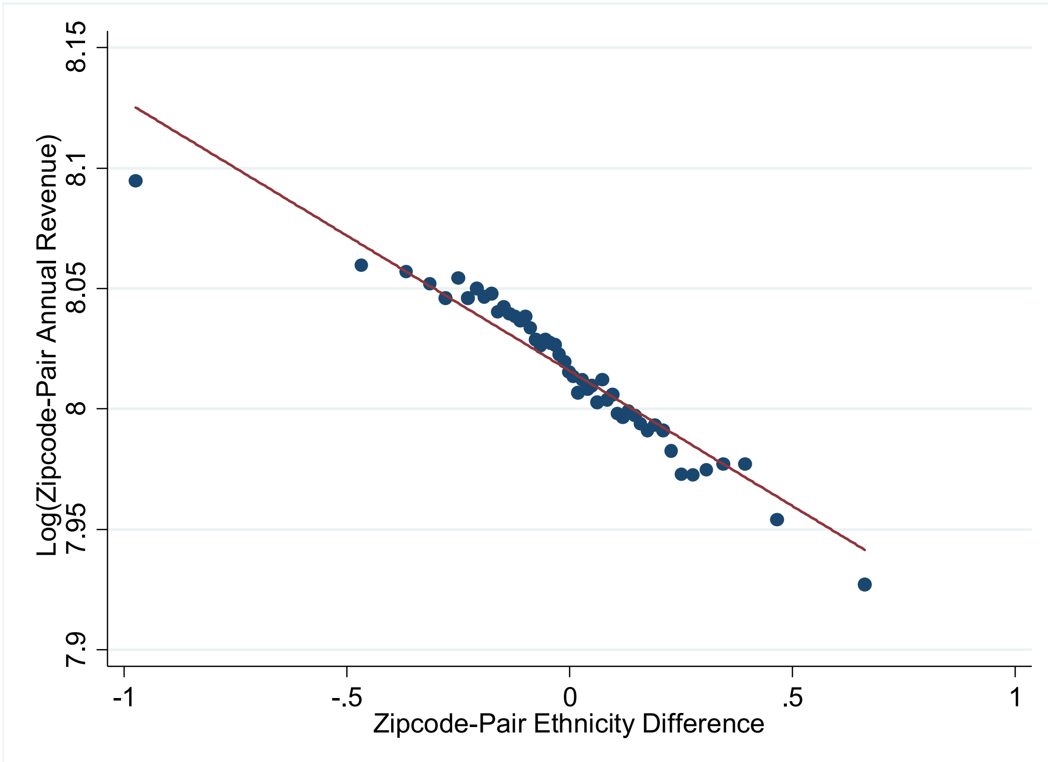
Note: The figure shows a binned scatter plot to illustrate the relationship between state-pair differences in religiosity and interstate trade. We use 50 bins, and residualize the data to account for all other variables included in Table 2, column 3.

Figure A3: Relationship between differences in voting and interstate trade



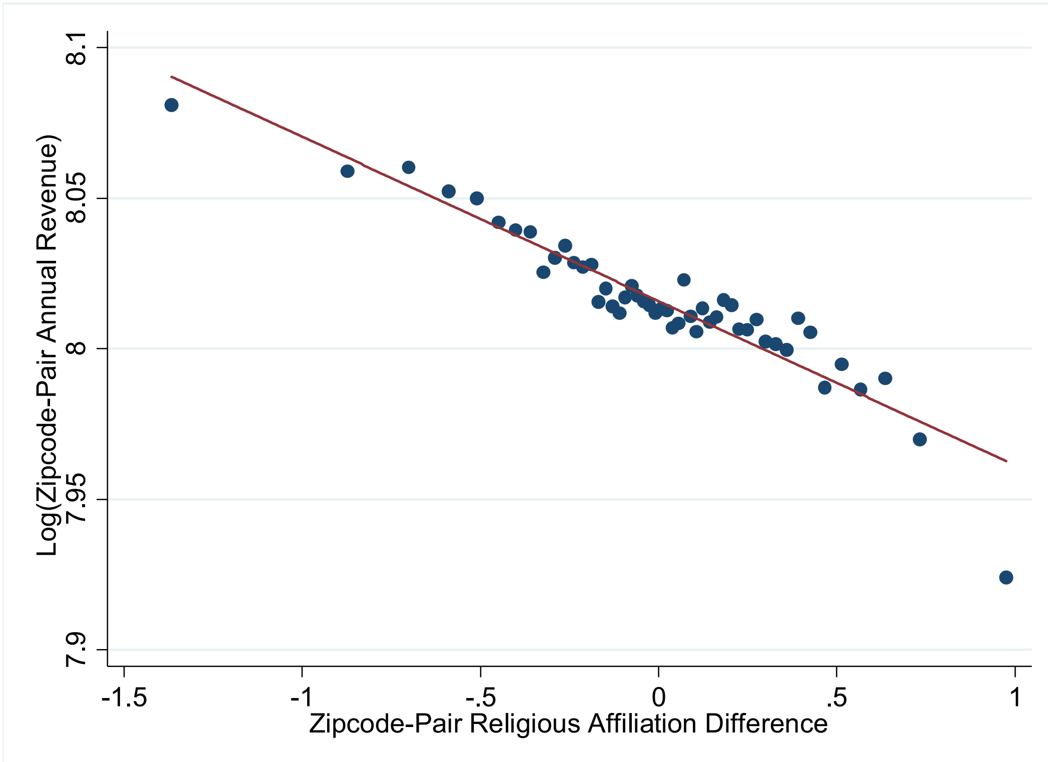
Note: The figure shows a binned scatter plot to illustrate the relationship between state-pair differences in voting in the 2016 presidential election and interstate trade. We use 50 bins, and residualize the data to account for all other variables included in Table 2, column 3.

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



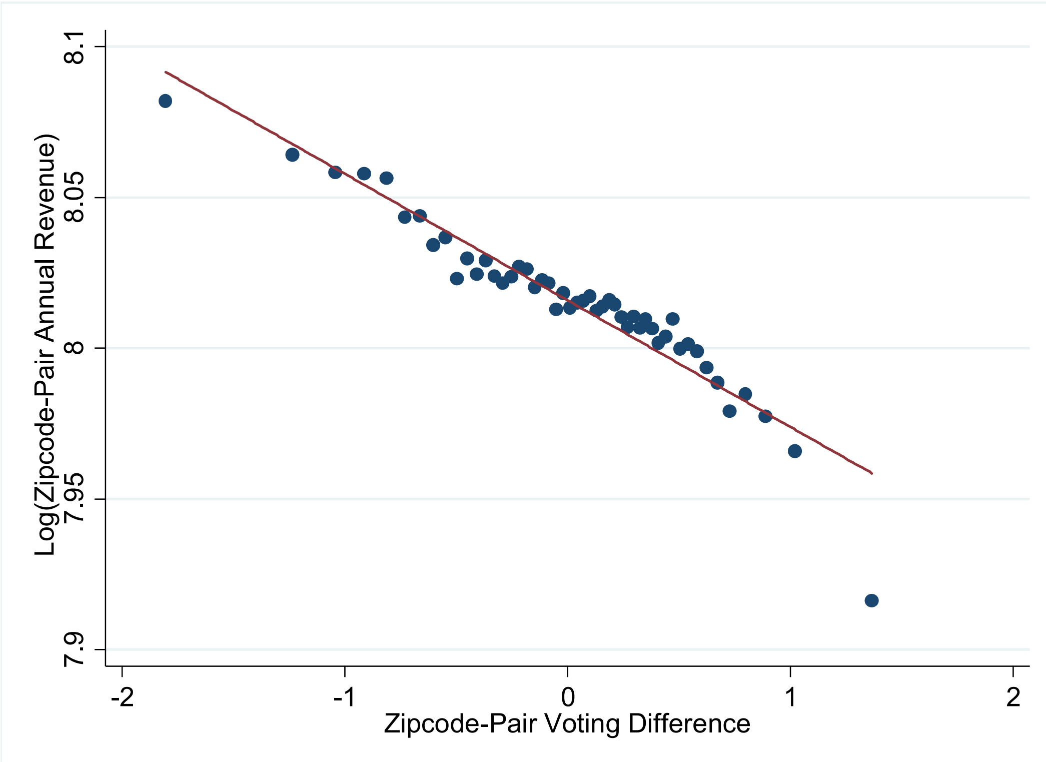
Note: The figure shows a binned scatter plot to illustrate the relationship between zip-code-pair differences in ethnicity and trade. We use 50 bins, and residualize the data to account for all other variables included in Table A4, column 3.

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



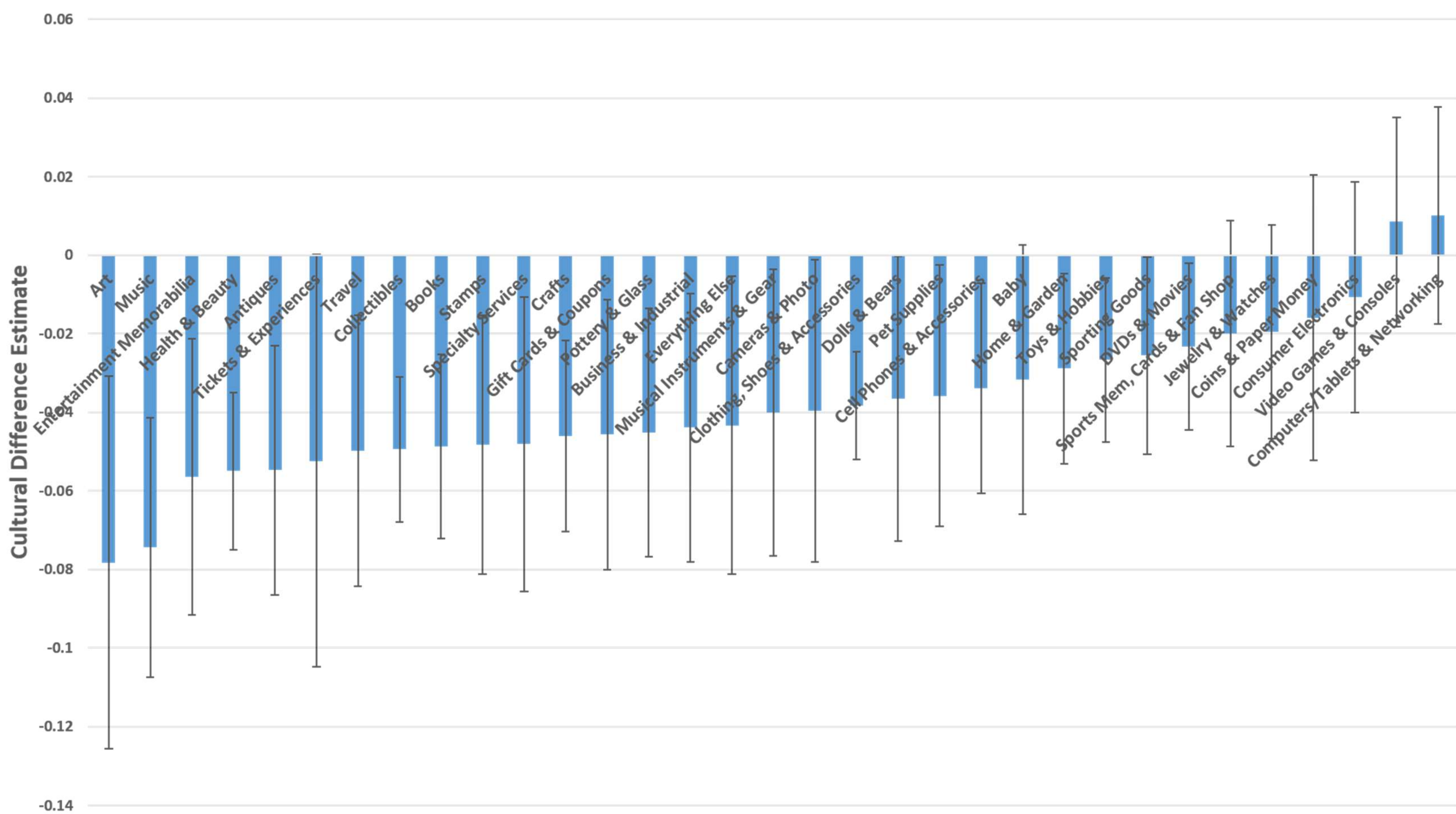
Note: The figure shows a binned scatter plot to illustrate the relationship between zip-code-pair differences in religiosity and trade. We use 50 bins, and residualize the data to account for all other variables included in Table A4, column 3.

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.

Figure A7. Correlation between coefficient estimates on sum of difference measures, across product categories, 3-digit zip to 3-digit zip data, log (1+total trade) specification



Appendix Table A1: Summary statistics, zip-code data

	Mean	Median	SD	Min	Max
Panel A: Annual sales and feedback data summary, zip-to-zip transactions (N = 1,571,764)					
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
Panel B: Zip code characteristics (N=887)					
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
Panel C: Differences in zip-code area characteristics (785,882)					
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
Difference in …					
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 zip codes (km)	1633.73	1408.81	1078.88	0.00	8368.32

Note: See text for variable definitions

Appendix Table A2: The impact of cultural and socioeconomic differences on trade between zip codes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	All	Same state	All	All	All	All	All
Estimation method	OLS	OLS	OLS	OLS	OLS	PPML	PPML
Dependent variable	log(revenue)	log(revenue)	log(quantity)	log(revenue)	log(quantity)	Revenue	Quantity
Ethnic difference	-0.112***	-0.107***	-0.0611***				
	(0.00814)	(0.0242)	(0.00639)				
Religious difference	-0.0545***	-0.0379**	-0.0387***				
	(0.00476)	(0.0172)	(0.00370)				
Voting difference	-0.0420***	-0.0404***	-0.0276***				
	(0.00353)	(0.00732)	(0.00276)				
Summed cultural difference				-0.0553***	-0.0353***	-0.0403***	-0.0372***
				(0.00263)	(0.00210)	(0.00560)	(0.00365)
Median income difference	0.000922	-0.00750	0.00302	-0.000510	0.00182	0.00308	-0.00279
	(0.00402)	(0.00779)	(0.00289)	(0.00413)	(0.00292)	(0.00581)	(0.00448)
Bachelors share difference	-0.0248***	0.0267***	-0.0197***	-0.0172***	-0.0156***	-0.0193***	-0.00595
	(0.00450)	(0.00790)	(0.00324)	(0.00471)	(0.00332)	(0.00695)	(0.00443)
Median age difference	0.000745	-0.0163**	0.00196	-0.00310	0.000131	-0.00155	-0.00347
	(0.00222)	(0.00660)	(0.00167)	(0.00229)	(0.00164)	(0.00338)	(0.00313)
Male share difference	-0.0126***	-0.00575	-0.0173***	-0.0130***	-0.0172***	-0.0192***	-0.0127***
	(0.00372)	(0.00799)	(0.00274)	(0.00377)	(0.00279)	(0.00402)	(0.00281)
Median home value difference	-0.0281***	-0.0797***	-0.0166***	-0.0274***	-0.0160***	-0.00817	-0.00173
	(0.00651)	(0.0124)	(0.00472)	(0.00654)	(0.00477)	(0.00768)	(0.00811)
Urban share difference	-0.00893**	-0.0456***	-0.0368***	-0.0103**	-0.0373***	-0.0577***	-0.0412***
	(0.00413)	(0.00677)	(0.00355)	(0.00415)	(0.00354)	(0.00669)	(0.00464)
Owner-occupied share difference	-0.00163	0.0274***	-0.00161	-0.00863**	-0.00426	-0.00479	-0.00778*
	(0.00403)	(0.00806)	(0.00294)	(0.00415)	(0.00316)	(0.00526)	(0.00462)
Log distance between zip codes	-0.161***	-0.0798***	-0.133***	-0.165***	-0.135***	-0.0644***	-0.0692***
	(0.00547)	(0.00999)	(0.00403)	(0.00539)	(0.00395)	(0.00673)	(0.00544)
Observations	1,571,764	43,556	1,571,764	1,571,764	1,571,764	1,571,764	1,571,764
R-squared (within)	0.0717	0.0900	0.1527	0.0710	0.152	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.

Appendix Table A3: PPML estimates of state-to-state and zip-to-zip trade with separate cultural difference measures

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A:		Panel B:		Panel C:	
Sample	State-to-State		Zip-to-Zip		StateXCat.-to-StateXCat.	
Dependent variable	Revenue	Quantity	Revenue	Quantity	Revenue	Quantity
Ethnic difference	-0.0647***	-0.0721***	-0.0862***	-0.0974***	-0.0637***	-0.0724***
	(0.0116)	(0.0111)	(0.0129)	(0.0114)	(0.0110)	(0.0111)
Religious difference	-0.00986	-0.0277***	-0.0368***	-0.0433***	-0.00848	-0.0280***
	(0.0155)	(0.0103)	(0.00874)	(0.00490)	(0.0155)	(0.0104)
Voting difference	-0.0215***	-0.00901**	-0.0281***	-0.0166***	-0.0225***	-0.00889**
	(0.00658)	(0.00362)	(0.00596)	(0.00478)	(0.00726)	(0.00377)
Median income difference	0.00737	0.00560	0.00463	0.000154	0.00766	0.00614
	(0.00751)	(0.00714)	(0.00571)	(0.00447)	(0.00776)	(0.00721)
Bachelors share difference	0.000356	-0.00327	-0.0238***	-0.0134***	0.000738	-0.00348
	(0.00945)	(0.00652)	(0.00663)	(0.00439)	(0.00980)	(0.00666)
Median age difference	0.00860***	0.00506**	0.000357	-0.000569	0.00815***	0.00526**
	(0.00239)	(0.00216)	(0.00303)	(0.00279)	(0.00232)	(0.00224)
Male share difference	-0.0107	-0.00425	-0.0193***	-0.0128***	-0.00964	-0.00427
	(0.00937)	(0.00835)	(0.00415)	(0.00309)	(0.00908)	(0.00835)
Median home value difference	-0.0147*	-0.00778	-0.0104	-0.00462	-0.0139*	-0.00843
	(0.00857)	(0.00543)	(0.00777)	(0.00746)	(0.00800)	(0.00551)
Urban share difference	-0.0171	-0.0140	-0.0555***	-0.0394***	-0.0168	-0.0141
	(0.0116)	(0.0105)	(0.00633)	(0.00454)	(0.0115)	(0.0105)
Owner-occupied share difference	-0.0178**	-0.0113**	0.000110	-0.00179	-0.0190***	-0.0111**
	(0.00714)	(0.00476)	(0.00447)	(0.00406)	(0.00712)	(0.00488)
Log distance between areas	-0.0973***	-0.0849***	-0.0625***	-0.0656***	-0.0987***	-0.0847***
	(0.0102)	(0.00633)	(0.00686)	(0.00563)	(0.0101)	(0.00643)
Observations	5,100	5,100	1,571,764	1,571,764	168,300	168,300

Note: In addition to the listed variables, all models include buyer-state × product category × year and seller-state × product category × 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 Table A4: The moderating role of seller quality in the impact of cultural difference on trade. Dependent variable: log(quantity)

	(1)	(2a)	(2b)	(3a)	(3b)
	base effect	*base effect, eTRS = 0*	*add'l effect for eTRS = 1*	*base effect, High fdbk = 0*	*add'l effect for High fdbk = 1*
Summed cultural difference	-0.0215***	-0.0231***	0.00750**	-0.0228***	0.00390*
	(0.00347)	(0.00355)	(0.00320)	(0.00346)	(0.00230)
Difference in:					
[1] Median income	-0.00106	-0.00138	0.00263	-0.00106	-0.000353
	(0.00364)	(0.00354)	(0.00277)	(0.00363)	(0.00282)
[2] Bachelors share	0.00478	0.00464	0.000708	0.000425	0.00971***
	(0.00435)	(0.00377)	(0.00501)	(0.00390)	(0.00301)
[3] Median age	0.00188	0.00169	-0.00280	-0.000392	0.00285
	(0.00481)	(0.00533)	(0.00495)	(0.00732)	(0.00447)
[4] Male share	-0.0158**	-0.0152**	-0.00953*	-0.0220**	0.00346
	(0.00681)	(0.00754)	(0.00556)	(0.00863)	(0.00319)
[5] Med. home value	-0.0117*	-0.0121*	0.00514	-0.0128**	0.000230
	(0.00673)	(0.00674)	(0.00514)	(0.00628)	(0.00514)
[6] Urban share	-0.0131***	-0.0154***	-0.00618	-0.0215***	0.00392
	(0.00401)	(0.00432)	(0.00431)	(0.00518)	(0.00261)
[7] Owner-occup. Share	-0.00746**	-0.00893***	0.00836	-0.00980***	0.00633*
	(0.00323)	(0.00320)	(0.00544)	(0.00347)	(0.00349)
Test of sum of [1]-[7] = 0	F = 22.80	F = 23.50	F = 0.03	F = 35.98	F = 10.94
	(p = 0.0000)	(p = 0.0000)	(p = 0.8688)	(p = 0.000)	(p = 0.0017)
Log distance between states	-0.129***	-0.132***	0.0269***	-0.141***	0.0239***
	(0.00719)	(0.00729)	(0.00384)	(0.00743)	(0.00402)
Observations	168,300	336,600		336,600	
R-squared (within)	0.115	0.0797		0.0994	

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 × *eTRS* × year and seller-state × *eTRS* × 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.

Appendix Table A5: The moderating role of seller quality in the impact of cultural difference on trade. Dependent variable: Revenue (PPML)

	(1)	(2a)	(2b)	(3a)	(3b)
	base effect	*base effect, eTRS = 0*	*add'l effect for eTRS = 1*	*base effect, High fdbk = 0*	*add'l effect for High fdbk = 1*
Summed cultural difference	-0.0189***	-0.0243***	0.0107**	-0.0290***	0.0140***
	(0.00484)	(0.00479)	(0.00449)	(0.00364)	(0.00455)
Difference in:					
[1] Median income	-0.000513	0.000252	-0.000841	-0.00166	0.000224
	(0.00620)	(0.00571)	(0.00563)	(0.00357)	(0.00573)
[2] Bachelors share	0.00538	0.00118	0.00963	0.00364	0.00209
	(0.00663)	(0.00596)	(0.00667)	(0.00453)	(0.00707)
[3] Median age	0.00482***	0.00611***	-0.00355	0.00340**	0.00200
	(0.00180)	(0.00166)	(0.00315)	(0.00142)	(0.00251)
[4] Male share	-0.00721	-0.00595	-0.00165	-0.0102**	0.00588
	(0.00620)	(0.00580)	(0.00747)	(0.00503)	(0.00488)
[5] Med. home value	-0.00499	-0.00241	-0.00743	-0.00913***	0.00529
	(0.00595)	(0.00573)	(0.00941)	(0.00352)	(0.00625)
[6] Urban share	-0.0160**	-0.0171**	0.00530	-0.0225***	0.00970*
	(0.00727)	(0.00695)	(0.00733)	(0.00558)	(0.00564)
[7] Owner-occup. share	-0.0181***	-0.0232***	0.0113*	-0.0217***	0.00522
	(0.00447)	(0.00561)	(0.00616)	(0.00452)	(0.00460)
Test of sum of [1]-[7] = 0	Chi sq = 7.11	Chi sq = 9.36	Chi sq = 0.73	Chi sq = 24.81	Chi sq = 3.98
	(p = 0.0076)	(p = 0.0022)	(p = 0.3915)	(p = 0.0000)	(p = 0.0461)
Log distance between states	-0.0978***	-0.103***	0.0151***	-0.115***	0.0250***
	(0.00693)	(0.00663)	(0.00434)	(0.00687)	(0.00517)
Observations	168,300	336,600		336,600	

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 × *eTRS* × year and seller-state × *eTRS* × 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.

Appendix Table A6: The moderating role of seller quality in the impact of cultural difference on trade. Dependent variable: Quantity (PPML)

	(1)	(2a)	(2b)	(3a)	(3b)
	base effect	*base effect, eTRS = 0*	*add'l effect for eTRS = 1*	*base effect, High fdbk = 0*	*add'l effect for High fdbk = 1*
Summed cultural difference	-0.0193***	-0.0232***	0.00770***	-0.0284***	0.0117***
	(0.00322)	(0.00329)	(0.00260)	(0.00306)	(0.00264)
Difference in:					
[1] Median income	-0.000586	0.000724	-0.00263	-0.000628	-0.000640
	(0.00514)	(0.00367)	(0.00607)	(0.00290)	(0.00479)
[2] Bachelors share	0.00276	-0.000180	0.00706	0.00313	0.000134
	(0.00495)	(0.00377)	(0.00506)	(0.00367)	(0.00412)
[3] Median age	0.00268*	0.00323**	-0.000767	0.00335**	-0.000789
	(0.00144)	(0.00161)	(0.00255)	(0.00148)	(0.00145)
[4] Male share	-0.00471	-0.00515	0.000951	-0.00985**	0.00695*
	(0.00511)	(0.00516)	(0.00376)	(0.00420)	(0.00363)
[5] Med. home value	-0.00138	-0.00159	3.95e-05	-0.00743**	0.00705
	(0.00527)	(0.00450)	(0.00449)	(0.00314)	(0.00476)
[6] Urban share	-0.0120*	-0.0155***	0.00820	-0.0193***	0.00922
	(0.00651)	(0.00543)	(0.00657)	(0.00465)	(0.00568)
[7] Owner-occup. share	-0.0129***	-0.0155***	0.00489*	-0.0192***	0.00800***
	(0.00316)	(0.00320)	(0.00274)	(0.00340)	(0.00258)
Test of sum of [1]-[7] = 0	Chi sq = 5.76	Chi sq = 13.20	Chi sq = 1.54	Chi sq = 34.18	Chi sq = 7.53
	(p = 0.0164)	(p = 0.0003)	(p = 0.2145)	(p = 0.0000)	(p = 0.0061)
Log distance between states	-0.0895***	-0.0952***	0.0133***	-0.113***	0.0300***
	(0.00550)	(0.00544)	(0.00347)	(0.00530)	(0.00315)
Observations	168,300	336,600		336,600	

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 × *eTRS* × year and seller-state × *eTRS* × 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.