Presented at the U.S. International Trade Commission
20 October 2011

Ethnic Networks and Price Dispersion

Michael Anderson
The U.S. International Trade Commission
Washington D.C. 20436
michael.anderson2@usitc.gov

Washington and Lee University
Lexington, VA 24550
andersonm@wlu.edu

Martin Davies
Washington and Lee University
Lexington, VA 24550
daviesm@wlu.edu

Stephen L. S. Smith
Gordon College
Wenham, MA 01984
stephen.smith@gordon.edu

Abstract: We offer a simple theoretical model of ethnic networks and price dispersion that suggests that, if networks have the risk-reducing, information-enhancing and contract-enforcing role in trade found by Rauch and Trindade (2002) and others, they should also contribute to decreased price dispersion. Our model predicts both a lower mean price dispersion, and a lower variance in price dispersion, as ethnic network density rises from low to high levels. We further predict that these effects may be diminished or reversed at the highest levels of shared ethnic populations between countries as network discipline breaks down. Using data from Chinese, Indian and Japanese ethnic shares, we find corroborating descriptive evidence on all points: country pairs linked with a large co-ethnic network do indeed have lower mean price dispersion, as well as lower variances, effects that are reversed at the highest levels of ethnic presence. Fixed effects regression results that take advantage of the time dimension of our ethnic shares data confirm the descriptive analysis: for Chinese networks a one-standard deviation increase in network strength is associated with a 44 percent decrease in price dispersion. We find initial evidence that Indian and Japanese ethnic network effects are much smaller than Chinese network effects.

Acknowledgments. We thank our home institutions for generous financial support, and Lauren Gamache (U.S. International Trade Commission), Katherine Boiles and Ian Sturdy (Washington and Lee University), and Julia Mitchell (Gordon College) for able research assistance. The authors acknowledge the helpful comments received from participants at the 2010 Southern Economics Association meetings and the 2011 Mid-West International Economic Group.

Prices vary substantially around their long-run Purchasing Power Parity (PPP) means. This result has been replicated in study after study, leading Obstfeld and Rogoff (2000) to name it as one of their six most important puzzles in international macroeconomics. Asplund and Friberg (2001) discovered that even products sold at the same location at the same time (aboard a cruise line) are found to deviate significantly from PPP. Authors have sought explanations, examining (for example) whether these deviations can be explained by differences in income (Crucini, Telmer and Zachariadis (2005)), the spatial relationship between production and consumption (Anderson and vanWincoop (2004), Inanc and Zachariadis (2006), and Anderson, Schaefer and Smith (2010)), as well as exploring price dispersion via detailed case study (Nakamura and Zerom (2009)).

In this paper we explore a new explanation for price dispersion. Taking our inspiration from the literature that finds strong evidence that information-sharing networks increase trade, we explore a corollary relationship between such networks and price dispersion. Rauch and Trindade (hereafter RT), in an influential 2002 paper, find that trade between countries with ethnic Chinese shares commonly found in Southeast Asia experience an increase in trade of at least 60 percent.[1] Other authors have continued this theme; Kumagai (2007), for example, finds that a country-pair's trade increases as the share of ethnic Japanese in their respective populations increases.[2]

In our theoretical section we offer a simple story for how such networks provide information about arbitrage opportunities, resulting in a movement of prices toward PPP.[3] As we will show in

---

[1] Rauch (2001) documents additional instances of ethnic networks increasing trade, and explores theoretical reasons for this phenomenon.

[2] By contrast, some authors doubt whether the empirical results support the theory. Felbemayr, Jung and Toubal (2009) argue that shared ethnic ties increase trade as a result of common tastes. Moreover, they argue that RT's model suffers from missing variable bias; they find a much more modest 15 percent increase in trade from shared ethnic networks.

[3] Testing the effect of networks on price dispersion is interesting in its own right, and potentially useful in solving the debate noted in footnote 2; if trade effects are as large as RT suggest, then one would expect substantial movement in prices toward PPP. By contrast, trade effects like those

our empirical section, the presence of information-sharing ethnic networks can substantially reduce price dispersion. Moreover, one network – those of Chinese ethnicity – is found to have effects on price dispersion that dwarf those created by Indian and Japanese networks. This latter result has important implications for how networks are studied: not all migration flows have important economic consequences

Our paper proceeds as follows. In the next section we develop a simple model of the role of ethnic trading networks in trade, and examine the implications of that role for international price dispersion. The rest of the paper explores, empirically, the extent to which shared ethnic-Chinese, Indian and Japanese population contributes to lower international price dispersion. Our price data is from the frequently-used Economist Intelligence Unit CityData, including price data on more than 100 products in more than 100 cities covering the years 1990-2009. We find that price dispersion falls between international location pairs as the shared proportion of Chinese population rises, and falls most sharply for goods that are most arbitrageable. Shared populations of Indians and Japanese are also associated with lower price dispersion. Finally, in a preliminary finding, the effects of Japanese ethnic networks in depressing price dispersion seem to be substantially larger than that of Chinese and Indian networks, at least within the OECD and Asia.

I. Ethnic Networks and Price Dispersion: Theory

If ethnic networks really do promote trade, what relationship would we expect to see between the strength of co-ethnic networks and international price dispersion? Following RT and the associated

---

suggested by Felbemayr *et al* would be less likely to show up in strong price movements. This indirect evidence depends upon, among other things, product characteristics, a point we also address.

literature, ethnic networks provide enforcement of sanctions that deter contract violations, they provide market information about price opportunities, and they provide matching services between buyers and sellers.

We begin the analysis by examining a pair of countries, i and j, which each have a continuum population of individuals on $[0,1]$. For illustrative purposes we consider symmetric countries in which the fraction of the population in each country that belongs to an ethnic network is n, and the remaining fraction $(1-n)$ do not.[4]

To engage in arbitrage of price differences between countries, individuals in one population must be matched with individuals in the other. This reflects the matching of buyers and sellers for distribution and supply. We model this as a random matching scheme and while this generates a stochastic system, following Boylan (1992), by appealing to the law of large numbers this is equivalent to a deterministic system in which the fraction of matches between individuals belonging only to the co-ethnic network is n^2. The remaining share, $1-n^2$, consists of mixed matches $(2n(1-n))$ and matches in which neither individual belongs to an ethnic network $(1-n)^2$.

Each match, which is indexed by $k \in [0,1]$, receives a draw $c(k)$ from a distribution of costs, supported on $[0,\infty]$, which represents the resource cost per unit of engaging in trade for good k from a continuum of goods on $[0,1]$.[5] The presence of a co-ethnic network in a population of matches influences the position and dispersion of the distribution of costs because of the lower resource costs of engaging in arbitrage due to lower contracting costs and less opportunistic behavior (cheating), and also the reduced costs of search for arbitrage opportunities for those that belong to the network.

---

[4] If $n_i > n_j$ then the fraction of co-ethnic matches is $n_i.n_j$, and on average each member of the ethnic network in country i engages in $n_i/n_j > 1$ matches.
[5] Implicit in this framework is symmetry of costs, i.e. for good k, $c(k)$ is the same regardless of the direction of arbitrage.

Discipline ensures lower contracting costs and decreased opportunistic behavior amongst members of a network. The sanction of expulsion is the means by which a co-ethnic network creates discipline amongst its members. An individual's choice to exercise discipline stems from the value the individual perceives of belonging to the network, which reflects the opportunity cost of being outside it. The value to a member of belonging to the network is dependent on the discipline of the other members as this both obviates the need for contracting and reduces cheating. A member also gains through access to an information conduit which provides the location of arbitrage opportunities which she would otherwise seek out through costly search. This benefit, which increases with the size of the network, is referred to the network effect.

The presence of a co-ethnic network influences the distribution of costs directly, as this fraction of total matches has lower resource costs, and also indirectly through the threat of entry in goods traded through non-network matches. Although the threat of entry is imperfect it may influence discipline exerted by individuals engaged in non-network matches. Thus a large and highly disciplined co-ethnic network will have a significant effect on the distribution of costs. We argue that the influence of the co-ethnic network on the cost distribution is determined by the level of discipline of members and the size of the network.[6] Below we define the *effective discipline* of the network as the product of these two terms. The effective discipline of the co-ethnic network in countries i and j influences the cost distribution for the entire population of matches, and hence the dispersion of prices between countries i and j.

We present a simple model to show how discipline varies with the fraction of co-ethnic matches in the population, which we refer to as the network density. The value to an individual of

[6] Note that n is both the fraction of the total population in each country that belong to the ethnic group, and also the number of individuals in this group.

belonging to the co-ethnic network, v, depends on the discipline exerted by each individual within the network, d, and also the network effect which is proportional to n^2. Thus,

(1) $$v = d + n^2$$

The level of discipline, d, that each member of the network chooses depends on the value of the network to the individual, v, and has the following properties:[7]

i. *for given v, the individual's discipline is decreasing in the number of co-ethnic matches.*

There are two reasons for declining discipline: firstly, as n increases enforcing exclusion becomes more costly as an expelled member who re-enters the network under a different guise will face a lower probability of detection; secondly, one could imagine that as n increases distribution and supply become more complex and the number of parties involved in the arbitrage of a good increases leading to increased difficulty in the detection of opportunistic behavior. Both effects alter the cost-benefit calculus of cheating leading to lower discipline.[8]

ii. *for a given number of matches, the individual's discipline is increasing in v.*

With the size of the co-ethnic network fixed, from (1), when all other members of the network increase their discipline then the value of the network to the individual rises. This increases the opportunity cost of being excluded, and the individual responds by increasing her level of discipline. Following from *i.* and *ii.*, the individual's discipline is represented as

---

[7] We simplify the analysis by assuming individuals are homogenous, and thus under the same circumstances will choose the same d.

[8] Although not modeled explicitly here, one could imagine a supergame in which at every point in time a member of a match chooses whether to collude (choose a high level of discipline) or defect (choose a lower level of discipline).

(2) $d = v(\alpha - \beta n)$

where $\alpha, \beta \in [0,1]$ *and* $\alpha \geq \beta$.[9] The constants α and β represents the innate characteristics of an ethnic group in forming a disciplined network, with α scaling the level of discipline, and β representing the sensitivity of discipline to network size at constant *v*. When *n* is small, the network effect is small and despite the fact that discipline is higher at low *n* (for given *v*) because the value of the network is low the level of discipline is also low. As *n* increases the network effect increases, the value of the network rises and so does discipline. However at high *n* discipline may eventually fall despite a higher network effect depending on the sensitivity of discipline to *n*, as parameterized by β. Note that while the network effect is increasing at an increasing rate in *n* reflecting the rate at which network connections increase, discipline (at constant *v*) decreases linearly with *n*. Substituting (2) into (1) leads to an expression for the value of the network in terms of *n*,

(3) $v(n) = n^2 / (1-(\alpha-\beta n))$

The numerator represents the influence of the network effect, and the denominator the influence of discipline, on the value of the network to the individual. As *n* increases the value increases through a higher network effect (higher numerator), and decreases through lower discipline (higher denominator). Setting $\alpha=1$ for simplicity, then $v = n/\beta$ and substituting for *v* in (2) leads to

(4) $d(n) = n(1- \beta n)/\beta$

---

[9] This ensures that $d \geq 0$.

Discipline increases for low values of n however beyond the threshold $n_d = 1/2\beta$ discipline falls as n increases. When discipline is less sensitive to n, *and* $\beta \leq 1/2$, then $n_d = 1$ and discipline is non-decreasing in n.

The influence of the co-ethnic network on the cost distribution for the population of matches depends on the level of discipline of individuals within the network, $d(n)$, and the number of co-ethnic matches in the population, n^2. We define effective discipline as the product of these two influences (discipline is being weighted by network density) and

(5) $$E(n) = n^2.d(n) = n^3(1- \beta n)/\beta$$

A small co-ethnic network with a high level of discipline may have a smaller effect on the cost distribution than a larger network with a lower level of individual discipline. We model the influence of effective discipline on the cost distribution by imposing that the mean, $\mu(E)$, and variance, $\sigma^2(E)$, of the cost distribution for the population of matches are decreasing in $E(n)$.[10] How does the distribution of costs influence price dispersion? Defining $p_i(k)$ as the common-currency price of good k in location i, and suppressing time subscripts for brevity, whenever $c(k) < |p_i(k)-p_j(k)|$ a match will engage in arbitrage to ensure that prices are driven so that $c(k)=| p_i(k)-p_j(k)|$. When $c(k) \geq| p_i(k)-p_j(k)|$ there is no entry. If international price differences are initially large then arbitrage will ensure that $c(k) =| p_i(k)-p_j(k)|$ for all k, and the distribution of price differences will fully reflect the distribution of costs. Further, the expected moments (first and second) of any random sample drawn from the distribution of price differences will be given by the distribution of

[10] $\mu'(E) < 0,\ \sigma^{2\prime}(E) < 0$

costs. More generally, the effect of a cost distribution with lower mean and variance will be to reduce the mean and variance of the distribution of price differences.[11]

From (5) effective discipline is increasing in n until the maximal point $n^* = 3/4\beta$ beyond which it decreases, however for $\beta \leq 3/4$ then $n^* = 1$ and E is increasing in n.[12] It follows that price dispersion will be decreasing in network density. When $\beta > 3/4$ price dispersion has a 'u-shaped' relationship with number of co-ethnic matches (network density), initially decreasing until n^* and then increasing above it. Variations in the rate of change of effective discipline, E, are also determined by β.[13] For example when $\beta = 3/4$ then E is increasing at an increasing rate until the inflexion point $n^{**} = 2/3$ after which it is increasing at a decreasing rate.

II. Data and Descriptive Analysis

Our empirical work uses data from the Economist Intelligence Unit (EIU). The full EIU data set comprises annual data on the retail prices, including taxes, of (close to) 200 goods and services across 142 cities internationally drawn from all continents except Antarctica. Our study includes as many of these cities as possible (lack of data on covariates means that not all cities are included in all of the analyses that follow). Appendix Table 2 lists the EIU cities and countries.

---

[11] For the purposes of illustration suppose that the distribution of costs is Poisson with mean and the variance λ, and that the co-ethnic network influences the moments of cost distribution according to $\lambda = 1/(1+E)$. Then as network effectiveness, E, increases from zero λ falls below one.
[12] the maximum point n^* is determined by $dE/dn = (n^2/\beta)(3 - 4\beta n) = 0$.
[13] the inflexion point n^{**} is determined by $d^2E/dn^2 = (n/\beta)(6 - 12\beta n) = 0$. When $\beta = 1$ then $n^{**} = 1/2$, when $3/4 < \beta < 1$ then $1/2 < n^{**} < 2/3$, when $\beta = 3/4$ then $n^{**} = 2/3$, when $1/2 < \beta < 3/4$ then $2/3 < n^{**} < 1$, when $\beta = 1/2$ then $n^{**} = 1$, and when $\beta < 1/2$ then $d^2E/dn^2 > 0$ at $n=1$.

As our interest is in the effect of information-sharing networks on price dispersion, we retain only those EIU goods that are tradable. This division is somewhat *ad hoc* but relies on distinctions typically made in economic analysis, so that, for instance, haircuts are considered nontradable while shirts are tradable. Thus we use 120 of the available EIU goods price data series. In what follows we present results for "all goods," namely, all goods that we deem tradable. In addition we present results to two prominent categories in the EIU database, food and apparel. (We also at times employ a subset of tradables we subjectivity consider to be particularly "arbitrageable"—non-branded generic goods, for instance, as opposed to branded products such as Coca-Cola.) Appendix 1 lists the goods used in this study.

After converting all prices to U.S. dollar equivalents at market exchange rates (from the same EIU source), we calculate all possible relative price permutations across city pairs for each good and each time period. Letting i, j, k, and t subscript city i, city j, product k, and time t, respectively, each U.S. dollar price of a product in location i or j can be defined as P_{ikt} or P_{jkt}. We (arbitrarily) choose to calculate price differentials (for product k in period t) as P_{ikt}/P_{jkt}. The log of this ratio provides a good approximation of the percentage price differential. Therefore our main dependent variable, the city pair price differential for product k in time period t, suppressing time and location subscripts, is

(6) $$\text{Price differential} = | \ln P_i - \ln P_j |$$

We will refer to this price differential as a price "wedge."

Overall price dispersion in this data set of tradable goods is large. The mean wedge, over all goods and location pairs, is 0.54, with a standard deviation of 0.46 (Table 1). The overall size of price dispersion, and its variance, changes little over the 1990-2009 time period. In descriptive results not reported here, restricting the sample to only countries in the OECD substantially reduces mean price dispersion to 0.44. International price dispersion is much larger than intra-national. When measured using only city pairs within the same country, mean price dispersion is about 0.30; measured between countries the value is near the overall mean, at 0.55.[14]

Next, consider our measure of network density for location pairs (i, j), based on ethnic-population shares in the two locations. As noted above, we use the phrase "network density" to refer to the number of co-ethnic matches across a country pair or, equivalently in our theoretical model, the product of the pair's co-ethnic shares, n^2. This differs from network effectiveness. For empirical work we define density as:

(7) $$\text{Network Density} = \ln(\text{zPopShare}_i * \text{zPopShare}_j)\ ,$$

where zPopShare_i is the percent of country i's population made up of ethnicity z. We measure the depth of the potential trading network created by this ethnic group in product form. This variable is the same as that used by Rauch and Trindade to measure Chinese ethnic network strength.

We have a rich data set on bilateral ethnic density, as described in Tables 2, 3 and 4. To our knowledge, this is the most extensive data set yet assembled to examine the effect of ethnic networks on market outcomes. Data sources are detailed in Appendix 3.

Table 2 presents a summary of the ethnic-share data (zPopShare), and Table 3 presents a summary of the network density variable ($\text{zPopShare}_i * \text{zPopShare}_j$) from equation (9). Note that both Table 2 and 3 present unlogged variable summaries.

Several important points are revealed in these two tables. The mean ethnic share variable by nation (Table 2) is a small number for each ethnicity. Said differently, the Chinese, Indians and

[14] Source: authors' calculations using EIU data. Price dispersion break-outs by country groups, and price dispersion statistics calculated within and between countries are available from the authors.

(especially) the Japanese are minorities nearly everywhere outside their home countries. Next, the Chinese network is the densest of the three networks (Table 3); the mean of its shared ethnic density variable is more than ten times the value for the Indian network, and two times the mean for the Japanese network. Finally, the standard deviation of network density (Table 3) for each ethnicity is typically 10 times larger than the mean. These huge variances (which give each distribution of network density a right-skewed shape) will be important later when we consider our regression results of price dispersion on network density.

Table 4 tabulates, by year, the number of shared ethnic density observations that exist in the dataset for each ethnicity. Recall that our data measure price dispersion by product by city pair by year. Ethnic density data is by country, and is imputed to each city pair. Thus it is not difficult to have more than 700,000 observations for a single ethnicity in a single year if, for example, for Indian ethnicity in 2001 there are data on 120 city and 100 goods. Though rich in size of years and number of networks, a weakness of our data is the small number of overlap years.[15]

In Tables 5-7 we present our measure of price dispersion measure in a cross tab with deciles of the logged product of ethnic population shares (Chinese, Indian and Japanese shares are in Table 5, 6, and 7, respectively). In Table 5 we see the mean and variance of price dispersion declining with our measure of network density up to the 9th decile. Both increase slightly in the 10th decile. This descriptive result suggests that network effectiveness ("E" in equation (5) in our theory section) may be rising for the first nine deciles in network density and declining thereafter. Interestingly, in the subsequent Tables 6 and 7 we also see declines in the mean and variance of price dispersion, though less dramatic and less smoothly, which stops at the 9th decile. For each of the three networks,

[15] The overlap is less than what appears to exist in the table; countries for which we have data on one ethnic network may be absent in the data for another network.

these general results hold across the board for the food, apparel and "arbitrageable" breakouts as well as for the all-goods data.

Our takeaway from the descriptive results is that there is, absent any controls, initial evidence for the idea that network density matters for price dispersion—the mean and the variance both respond to the size of the network. Interestingly, both moments fall through 9 deciles, but increase for the largest decile.

III. Empirical Model and Approaches to Estimation

The price dispersion measures discussed above are unconditioned. To fully account for ethnic networks' effect on price dispersion, we need a model of retail price dispersion that allows us to identify and control for sources of price dispersion that cannot be arbitraged away. Networks reduce the cost of international arbitrage, but do not necessarily affect the costs of retailing particular items in particular locations—costs which are reflected in retail prices. Networks may help clothing sellers in New York access inexpensive suppliers of clothing from abroad, but shirts sold in New York are sold from shops that pay New York wages, rents and sales taxes.

We adopt a simple arbitrage model based on the model used by Engel and Rogers (1996). All final goods are produced by a local monopolist and are nontraded. But production of final goods for local sale requires combining a traded intermediate input with non-traded factors. Thus, for instance, the final good may be Cornflakes sold in Hong Kong in 2002; the traded input is the box of Cornflakes while the nontraded inputs are the factors necessary for retailing Cornflakes there at that time. Suppressing k (good) and t (time) subscripts, define I_i to be the price of the traded intermediate input and W_i to be the price of the nontraded factors used in producing good k.

Assuming Cobb-Douglas production technology, the price of good k in location i can then be written as

$$(8)\qquad P_i = \mu_i (I_i)^{\alpha} (W_i)^{(1-\alpha)}$$

where μ_i is the markup and α is the share of the traded input in production cost. Several nontraded factors may need to be combined to produce the final product. We assume a nested Cobb-Douglas formulation in which

$$(9)\qquad W_i = \phi_i (\mathrm{w}_{i1})^{\gamma} (\mathrm{w}_{i2})^{(1-\gamma)},$$

where w_{i1} and w_{i2} are the prices of non-traded factors 1 and 2, respectively. Taking the log of relative prices in locations i and j, we obtain

$$(10)\qquad \ln P_i - \ln P_j = (1-\alpha)\ln(\phi_i/\phi_j) + \ln(\mu_i / \mu_j) + \alpha \ln(I_i / I_j)$$
$$+ (1-\alpha)\gamma \ln(\mathrm{w}_{i1} / \mathrm{w}_{j1}) + (1-\alpha)(1-\gamma)\ln(\mathrm{w}_{i2} / \mathrm{w}_{j2}).$$

Consider the terms on the right hand side, from left to right. The first term reflects the relative productivity of nontraded inputs in producing goods i and j. The second term, the relative markup, can be expected to vary with firms' ability to price-to-market (itself related to the elasticity of demand) and (since P_i and P_j are final sales prices) with relative local sales taxes.

The price of the traded input in location i relative to j (I_i/I_j) would be unity in a world of frictionless trade with perfect information. However, in practice, we expect this differential to be related to differentials in transport costs, to national trade taxes (or tax equivalents), and to all the non-pecuniary costs of transacting across international borders, including exchange-rate risk. We measure these directly where possible—with dummy variables for common currency, for borders, for free trade areas, for common language—and capture the effects of networks with the ethnic population share variable.

Relative non-traded factor prices can be measured directly and are expected to show considerable variation across city pairs precisely because they are nontraded and reflect all local supply and demand conditions. We expect that large positive factor price differentials between locations should generate, all else equal, large final goods price differentials. Low land prices in Lexington should lead lemons to cost less there than in London. In our formulation above we are assuming that each good faces the same non-traded factor prices at any one location. To estimate (10) we use the standard proxy for transport costs, distance. We take labor wages and land rents as the factor prices, and calculate relative factor price wedges between locations i and j as the natural log of the factor price ratios.[16]

To take advantage of the multiple years of network density data available for Chinese, Indian and Japanese networks, we employ the fixed effects ("within") estimator. The panel variable is city-pairs by good (for instance, Manila-Boston-apples). Thus for locations i and j, good k, and period t, the estimating equation is:

$$(11) \quad |\ln P_{ikt} - \ln P_{jkt}| = \beta_1(\mathrm{zEthnicDensity}_{ijkt}) + \beta_2\ln(\mathrm{wagediff}_{ijt}) + \beta_3\ln(\mathrm{rentdiff}_{ijt}) + \beta_4\ln(\mathrm{VATdiff}_{ijkt}) + \beta_5\ln(\mathrm{Tariffdiff}_{ijkt}) + \Sigma\theta_{ijk} + \varepsilon_{ijkt}.$$

[16] As described earlier, we take the dependent variable in absolute value as its sign is arbitrarily dependent on which city is "i" and which is "j," and adjust the factor price wedges accordingly.

The θ_{ijk} are the fixed effects, which capture the impact of time-invariant variables, leaving on the right hand side only the time-varying variables of wage and rent differentials (which vary by city-pair, not by good), VAT/sales tax and tariff differentials (which do vary by city-pair and good) and the ethnic density variable(s). Most of the results that follow are derived from the fixed effects estimator.

Secondarily, we report some results estimated with pooled OLS. Because relatively few observations for all three ethnic networks overlap in multiple years, it is not possible to use the within estimator for joint estimation of networks' effects. After grouping the data in reasonable time periods pooled OLS allows us to offer a handful of such joint estimates. The estimating equation is:

$$(12) \quad |\ln P_{ikt} - \ln P_{jkt}| = \beta_0 + \sum\beta_z (\text{zEthnicDensity}_{ijt}) + \beta_2 \ln \text{Dist}_{ij} + \beta_3 \text{Border}_{ij} + \beta_4 \ln(\text{wagediff}_{ijt}) + \beta_5 \ln(\text{rentdiff}_{ijt}) + \beta_6 \ln(\text{VATdiff}_{ijkt}) + \beta_7 \ln(\text{Tariffdiff}_{ijkt}) + \beta_8 \ln(\text{Common Currency}_{ijt}) + \beta_9 \ln(\text{FTA}_{ijt}) + \beta_{10}\,\text{time effects} + \beta_{11}\,\text{city effects} + \varepsilon_{ijkt}.$$

"Dist_{ij}" measures the great-circle distance, in miles, between locations i and j. The dummy variable "Border_{ij}" equals 1 if the two locations span an international border. Its expected sign is positive – all else equal there should be larger goods price wedges between countries than inside countries. We also included dummies for the presence of FTAs and common currencies between locations.

The wage differential variable, $\ln(\text{wagediff})_{ij}$, is the natural log of hourly labor costs, in U.S. dollar terms, in city i relative to city j (using EIU data), while the rent differential variable

$\ln(\text{rentdiff})_{ij}$ proxies the overall land price differential between cities as the natural log of the city i over city j ratio of monthly rent on a two-bedroom unfurnished apartment (again with EIU data).[17]

With both estimators, we explore potential non-linearity in the effectiveness of networks in two ways. We include a polynomial (squared) term on network density on the right hand side or, alternatively, estimate the entire relationship on a subset of the data in the top decile or quartile of the ethnic density measure.

IV. Estimation Results

In this discussion we focus on the fixed-effects estimations. First consider the results for Chinese networks, reported in Table 8. The columns show the results for all goods and then for food and apparel; within each column we consider first the regression on the left which does not include a polynomial term. All coefficient values are statistically significant, owing in part to large sample sizes, and have expected signs.

The all-goods regression returns a coefficient value of -0.023 for the elasticity of price dispersion with respect to Chinese ethnic density. While absolutely small, this implies a large effect on price dispersion, which can be seen as follows. A one-standard deviation increase in the Chinese network density variable (relative to the mean) is an increase of 1910 percent. Multiplying this 1910 by -0.023 yields a predicted decrease in price dispersion of 43.5 percent. (This and similar calculations for the economic significance of our estimated ethnic coefficients are tabulated in the bottom rows of Tables 8-10.) In effect, the one standard deviation increase in network density

[17] The EIU supplies various labor cost and housing rental series that can be used to proxy labor and land costs. We found that our results were invariant to using other series than the ones selected here.

reduces price dispersion by about one-half of a standard deviation (measured in sample), a not inconsiderable effect. Results for food and apparel are similarly substantial; the price dispersion coefficients translate into predicted reductions in price dispersion of 44.9 and 70.2 percent, respectively, when Chinese ethnic density rises by one standard deviation.

Results on Indian networks are reported in Table 9. The coefficient on the Indian network (-0.0002) is not statistically significant, despite a sample size of over 4 million observations. Taking the coefficient as our best, if noisy, estimate, then a one standard deviation of Indian network density would decrease price dispersion by one third of one percent. The coefficient on the network in the food-only regression is -.003, statistically significant, and suggests a decline in price dispersion of about 5 percent. Finally, the coefficient on apparel has an unanticipated positive sign, and is statistically significant. If accepted, it would suggest that price dispersion would increase by 16 percent for a one-standard-deviation increase in Indian network density.

Consider finally the results for the Japanese network. Following the discussion above, a one standard deviation change in the size of the Japanese networks is associated with a 23 percent decrease in price dispersion in the all-goods regression, and 103 percent with food. Price dispersion in food is, overall, quite large; mean price dispersion is nearly 200 percent (with a standard deviation nearly as large as the mean); hence the Japanese network can reduce the mean by about half. In the apparel case the coefficient has the wrong sign. Taken together, the results for the Japanese network are inconsistent, and difficult to reconcile.

As discussed earlier, it is possible that not all increases in ethnic-population shares will add to network density. Instead, beyond some level adding additional members may diminish network effectiveness if the policing of trust violations becomes more costly with a larger network. To gain some initial insight into this question we add to the model discussed above an additional term, the square of the log-product of ethnic shares, to allow the coefficient (the elasticity of price dispersion to the size of the network) to be non-constant, to vary with the size of the network. These results are in the right hand side of every column in Tables 8-10, for all-goods, food only, and apparel only, for each ethnic network. The implied elasticity of price dispersion with respect to ethnic density, evaluated at the mean, is indicated at the bottom of each column along with the implied percentage change in price dispersion. [18]

For the Chinese network, the all-goods regression, a one-standard deviation increase in ethnic network density (starting at the mean) will decrease price dispersion by 31 percent. Notice that this is smaller than that found in the non-polynomial model, where the result was 43 percent. The corresponding result for the food-only regression is a 26 percent decrease in price dispersion (smaller than the 45 percent in the non-polynomial model). For apparel the result is a decrease in price dispersion of of 75 percent, comparable to the result of 70 percent for the non-polynomial model.

Next, consider the results for the Indian network. For the all-goods regression, the effect of a one-standard deviation increase in network size is a negligible increase in price dispersion of about 2 percent. For the food-only regression the effect is again unexpectedly positive, with a similarly sized increase in network size associated with a 14 percent increase in price dispersion. Finally, in the apparel-only regression, we predict an increase in price dispersion of about 17 percent. Overall, these results suggest that the diaspora we have been calling the Indian "network" appears to have little effect on market integration.

---

[18] In our polynomial specification, where $y = \beta_1 \ln x + \beta_2 (\ln x)^2$, the elasticity of y with respect to x is $\beta_1 + 2\beta_2 \ln x$. Thus, for China, the elasticity at the mean is calculated as -0.009 + 2* 0.0005 *(-6.56), where -6.56 is the log of the mean of the ethnic network variable, and equals -0.016.

Finally, consider the results for the Japanese network. A one standard deviation increase in network size (from the mean) is predicted to decrease price dispersion of all goods by about 2 percent, decrease price dispersion in food by 106 percent, and decrease price dispersion for apparel items by about 28 percent. The results for Japan are therefore quite strong, larger in some cases than what was found for the Chinese network. What is surprising here is the small effect for all-goods, contrasted with the large effect for the sub-categories, given that the two sub-categories constitute approximately 75 percent of the all-goods category.

V. Conclusion

Our study offers several contributions. First, we offer a new explanation for price dispersion. Our theory predicts that strong ethnic networks diminish the costs of arbitrage and lead to price convergence. Both descriptive data analysis and our empirical model support that prediction. In our baseline specification a one standard deviation increase in our measure of ethnic Chinese network strength is associated with a reduction in price dispersion of one half of a standard deviation. Said differently, we predict mean price dispersion will decrease by about 44 percent when the Chinese ethnic-network density is one standard deviation above its mean.

An additional finding and our second contribution is that not all ethnic networks are the same. The Chinese ethnicity has the most reliably large effect in our study, and it is also the group found to rigorously punish cheating with exclusion. The literature on trade-and-networks, with some notable exceptions such as Bandyopadhyay *et al* (2008), treats population groups as equivalent in their effect on markets; our results here suggest a re-thinking of that assumption. If our results are robust across further studies, we should begin to regard different ethnicities as particular and distinct in their operation and effect.

Third, we find support for the notion that the marginal effect of a network depends upon its density. That is, our theory's second prediction that network effectiveness may decline as ethnic population shares increase beyond some threshold finds support, and our investigation on this point continues.

Table 1. Global Price Dispersion

Price dispersion between locations i and j measured by $|\ln P_i - \ln P_j|$.
All tradable goods, all regions.

Year	Mean	Standard Deviation	Number of Observations
1990	0.64	0.55	522,957
1991	0.62	0.54	555,059
1992	0.61	0.53	565,853
1993	0.58	0.50	601,255
1994	0.56	0.49	624,289
1995	0.55	0.47	643,984
1996	0.53	0.46	659,315
1997	0.50	0.43	633,974
1998	0.52	0.45	689,914
1999	0.53	0.46	735,994
2000	0.54	0.47	892,593
2001	0.55	0.46	912,430
2002	0.57	0.49	938,471
2003	0.56	0.48	931,708
2004	0.56	0.48	961,455
2005	0.56	0.47	980,822
2006	0.56	0.47	991,596
2007	0.57	0.47	1,005,373
2008	0.56	0.46	1,008,639
2009	0.55	0.45	842,385
OVERALL	0.56	0.48	15,698,066

Source: Authors' calculations from EIU CityData, described in Appendices 1 and 2.

Table 2. Ethnic Shares: Summary of Country Data (Percent)

	Chinese	Indian	Japanese
Number of countries	128	134	129
Minimum	0.00008	0.00006	0.00008
Maximum	97.37	98.00	100.00
Mean	8.76	3.00	1.60
Standard deviation	25.4	12.44	12.40

Sources: authors' calculations from sources detailed in Appendix 3.

Table 3. Network Density Measure: Summary Statistics

Network Density = product of ethnic group's proportion of population in nations i, j.

	CHINA	**INDIA**	**JAPAN**
Levels:			
Mean	0.001468	0.000906	0.000641
SD	0.027864	0.016987	0.000002
Minimum	6.05e-13	5.30e-14	1.96e-12
Maximum	0.961908	0.960401	1.00
Logs:			
Mean	-12.46742	-14.76008	-16.06871
SD	3.635763	5.179547	3.021516
Minimum	-28.13433	-30.56821	-26.95694
Maximum	-0.038836	-0.040405	0.00
Obs.	3,335,382	7,014,695	2,783,693

Sources: Authors' calculations based on ethnic share data described in Appendix 3.

Table 4. Availability of Network Density Measure by Ethnic Group and Year

Number of paired observations of network density measure

Year	Chinese	Indian	Japanese
1990	362,475	10,910	252,400
1991	336,454	266	0
1992	355,201	87	0
1993	175,822	86	0
1994	671	85	0
1995	3,941	260	315,878
1996	294,294	1,528	109
1997	503,453	13,238	110
1998	517,621	87	3,712
1999	537,925	91	1,557
2000	2,406	284	412,275
2001	115	781,688	109
2002	0	740,212	108
2003	245,004	734,438	106
2004	0	761,049	104
2005	0	777,881	407,738
2006	0	786,930	411,119
2007	0	799,292	478,618
2008	0	802,650	467,481
2009	0	803,633	32,269
Total	3,335,382	7,014695	2,783,693

Table 5. Mean and Standard Deviation of Price Dispersion by Chinese Ethnic Network Strength (Deciles)

	All Goods	Arbitrageable	Food Products	Apparel
Decile	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
1	0.64 (0.53)	0.70 (0.57)	0.69 (0.56)	0.62 (0.51
2	0.62 (0.53)	0.68 (0.57)	0.68 (0.56)	0.63 (0.51)
3	0.57 (0.49)	0.63 (0.54)	0.63 (0.53)	0.54 (0.45)
4	0.56 (0.49)	0.62 (0.53)	0.62 (0.52)	0.54 (0.44)
5	0.52 (0.45)	0.58 (0.49)	0.58 (0.48)	0.50 (0.41)
6	0.51 (0.43)	0.56 (0.46)	0.56 (0.46)	0.47 (0.39)
7	0.50 (0.42)	0.55 (0.45)	0.55 (0.44)	0.44 (0.37)
8	0.49 (0.42)	0.55 (0.46)	0.54 (0.46)	0.45 (0.38)
9	0.44 (0.39)	0.48 (0.41)	0.48 (0.41)	0.43 (0.38)
10	0.48 (0.40)	0.48 (0.40)	0.48 (0.40)	0.49 (0.40)
N per decile	333,487	167, 140	157,108	44,365

Pooled observations, various years 1990-2003.

Table 6. Mean and Standard Deviation of Price Dispersion by Indian Ethnic Network Strength (Deciles)

	All Goods	Arbitrageable	Food Products	Apparel
Decile	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
1	0.58 (0.47)	0.55 (0.46)	0.62 (0.50)	0.62 (0.49)
2	0.60 (0.49)	0.59 (0.49)	0.65 (0.52)	0.63 (0.49)
3	0.60 (0.49)	0.59 (0.48)	0.63 (0.50)	0.62 (0.48)
4	0.59 (0.49)	0.59 (0.48)	0.64 (0.50)	0.60 (0.48)
5	0.56 (0.48)	0.58 (0.48)	0.60 (0.49)	0.58 (0.48)
6	0.54 (0.45)	0.58 (0.48)	0.56 (0.45)	0.57 (0.46)
7	0.50 (0.42)	0.58 (0.48)	0.50 (0.42)	0.56 (0.46)
8	0.48 (0.41)	0.59 (0.48)	0.49 (0.40)	0.53 (0.45)
9	0.48 (0.40)	0.58 (0.48)	0.48 (0.39)	0.52 (0.43)
10	0.52 (0.46)	0.58 (0.48)	0.51 (0.43)	0.61 (0.51)
N per decile	701,470	343,976	324,824	90,068

Pooled observations, 2001-2009.

Table 7. Mean and Standard Deviation of Price Dispersion by Japanese Ethnic Network Strength (Deciles)

	All Goods	Arbitrageable	Food Products	Apparel
Decile	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
1	0.67 (0.56)	0.71 (0.60)	0.68 (0.60)	0.81 (0.59)
2	0.62 (0.52)	0.66 (0.55)	0.63 (0.58)	0.69 (0.53)
3	0.58 (0.48)	0.61 (0.49)	0.59 (0.52)	0.60 (0.46)
4	0.54 (0.45)	0.58 (0.47)	0.57 (0.43)	0.54 (0.44)
5	0.51 (0.43)	0.54 (0.44)	0.53 (0.45)	0.52 (0.42)
6	0.52 (0.42)	0.54 (0.43)	0.53 (0.44)	0.53 (0.42)
7	0.51 (0.41)	0.53 (0.42)	0.53 (0.43)	0.50 (0.38)
8	0.48 (0.40)	0.50 (0.42)	0.50 (0.42)	0.46 (0.36)
9	0.40 (0.35)	0.42 (0.36)	0.42 (0.45)	0.43 (0.35)
10	0.57 (0.50)	0.65 (0.56)	0.65 (0.41)	0.58 (0.48)
N per decile	278,361	137,463	129,614	36,915

Pooled observations, various years 1990-2009.

Table 8. Chinese Network Effects—Fixed Effects ("Within") Estimator

Dependent variable: $|\ln P_i - \ln P_j|$
Group variable: city-pair by good
Cluster standard errors, by group variable, in parentheses
Data range: 1990-2003

Ind. Var.	**All Goods**		**Food**		**Apparel**	
ln(Chinese Network)	-0.023	-0.009	-0.023	-0.004	-0.038	-0.043
	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)	(0.004)
$[\ln(\text{Chinese Network})]^2$	x	0.001	x	0.001	x	-0.0002
		(0.000)		(0.000)		(0.0002)
ln(Wage Ratio)	0.102	0.103	0.115	0.115	0.082	0.082
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)
ln(Rent Ratio)	0.107	0.107	0.110	0.110	0.113	0.113
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
ln(VAT Ratio)	0.279	0.284	0.326	0.332	0.903	0.899
	(0.041)	(0.041)	(0.053)	(0.053)	(0.126)	(0.126)
ln(Tariff Ratio)	-0.055)	-0.055	-0.020)	-0.021	-0.229	-0.229
	(0.004	(0.009)	(0.005)	(0.012)	(0.011)	(0.011)
Observations	2,036,355		1,363,285		302,693	
Groups	590,001		402,060		85,250	
R-square: within	0.0493	0.0494	0.0517	0.0519	0.0618	0.0618
(between)	0.1173	0.1184	0.1159	0.1165	0.1365	0.1359
(overall)	0.1161	0.1164	0.1156	0.1155	0.1319	0.1315
Elasticity of PD w.r.t ethnic density, evaluated at mean PD	-0.023	-0.016	-0.023	-0.013	-0.038	-0.041
Implied percent change in PD of increase in ethnic density by σ	-43.4	-30.5	-44.9	-26.3	-70.2	-74.9

Table 9. Indian Network Effects—Fixed Effects ("Within") Estimator

Dependent variable: $|\ln P_i - \ln P_j|$
Group variable: city-pair by good
Cluster standard errors, by group variable, in parentheses
Data range: 2001-2009

Ind. Var.	All Goods		Food		Apparel	
ln(Indian Network)	-0.0002	0.013	-0.003	0.014	0.007	0.010
	(0.0002)	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)
$[\ln(\text{Indian Network})]^2$	x	0.0003	x	0.0005	x	0.0001
		(0.0000)		(0.0000)		(0.0001)
ln(Wage Ratio)	0.250	0.250	0.245	0.245	0.318	0.318
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)
ln(Rent Ratio)	0.083	0.083	0.070	0.070	0.130	0.130
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
ln(VAT Ratio)	1.279	1.278	2.016	2.017	0.359	0.358
	(0.040)	(0.040)	(0.053)	(0.053)	(0.066)	(0.066)
ln(Tariff Ratio)	0.325	0.324	0.318	0.317	0.087	0.087
	(0.006)	(0.006)	(0.007)	(0.007)	(0.0116)	(0.0116)
Observations	4,190,146		2,877,130		585,239	
Groups	821,957		569,410		115,419	
R-square: within	0.0904	0.0905	0.0853	0.0854	0.1182	0.1182
(between)	0.0860	0.0842	0.0795	0.0765	0.1036	0.1036
(overall)	0.0832	0.0814	0.0763	0.0733	0.1027	0.1027
Elasticity of PD w.r.t ethnic density, evaluated at mean PD	-0.0002	0.010	-0.003	0.008	0.007	0.009
Implied percent change in PD of increase in ethnic density by σ	-0.3	1.9	-5.0	14.9	13.3	16.8

Table 10. Japanese Network Effects—Fixed Effects ("Within") Estimator

Dependent variable: $|\ln P_i - \ln P_j)|$
Group variable: city-pair by good
Cluster standard errors, by group variable, in parentheses
Data range: 1995-2009

Ind. Var.	**All Goods**		**Food**		**Apparel**	
ln(Japanese Network)	-0.006	0.005	-0.027	-0.029	0.019	-0.032
	(0.001)	(0.006)	(0.001)	(0.007)	(0.003)	(0.017)
$[\ln(\text{Japanese Network})]^2$	x	0.0004	x	-0.00005	x	-0.002
		(0.0002)		(0.0002)		(0.0005)
ln(Wage Ratio)	0.217	0.217	0.235	0.235	0.203	0.203
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
ln(Rent Ratio)	0.122	0.122	0.119	0.119	0.118	0.118
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)
ln(VAT Ratio)	0.801	0.802	0.818	0.818	0.989	0.984
	(0.041)	(0.041)	(0.054)	(0.054)	(0.107)	(0.107)
ln(Tariff Ratio)	0.372	0.370	0.406	0.407	0.113	0.112
	(0.008)	(0.008)	(0.011)	(0.011)	(0.023)	(0.023)
Observations	1,873,489		1,194,390		283,947	
Groups	562,028		366,590		84,415	
R-square: within	0.0960	0.0960	0.1091	0.1091	0.0714	0.0715
(between)	0.0962	0.0977	0.0999	0.0997	0.0943	0.0871
(overall)	0.1018	0.1032	0.1093	0.1011	0.1003	0.0936
Elasticity of PD w.r.t ethnic density, evaluated at mean PD	-0.006	-0.001	-0.027	-0.028	0.019	-0.009
Implied percent change in PD of increase in ethnic density by σ	-22.6	-1.8	-103.1	-106.0	60.3	-28.2

Table 11. Indian and Japanese Network Effects Combined—Fixed Effects (“Within”) Estimator

Dependent variable: $|\ln P_i - \ln P_j|$
Group variable: city-pair by good
Cluster standard errors, by group variable, in parentheses

	All Goods	Food	Apparel
ln(Indian Network)	0.001 (0.001)	0.015 (0.002)	-0.032 (0.004)
ln(Japanese Network)	-0.003 (0.003)	-0.004 (0.003)	-0.061 (0.008)
ln(Wage Ratio)	0.264 (0.002)	0.285 (0.002)	0.291 (0.006)
ln(Rent Ratio)	0.078 (0.001)	0.081 (0.002)	0.086 (0.004)
ln(VAT Ratio)	1.500 (0.050)	2.166 (0.072)	1.500 (0.119)
ln(Tariff Ratio)	0.360 (0.013)	0.389 (0.016)	0.024 (0.039)
Number of observations	1,121,765	714,302	168,476
Number of groups	415,924	273,518	61,243
R-square: within	0.0600	0.0694	0.0619
(between)	0.0899	0.0732	0.0923
(overall)	0.0911	0.0758	0.0892

Appendix Table 1. Price Dispersion and Joint Network Strength, Pooled OLS Estimation

Dependent variable: $|\ln P_i - \ln P_j|$
Robust standard errors in parentheses
Data grouped into two periods: "2000" and "2005"

	All Goods	**Food**	**Apparel**
ln(Chinese Network)	-0.018 (0.0004)	-0.019 (0.000)	-0.024 (0.001)
ln(Indian Network)	-0.006 (0.0003)	-0.008 (0.000)	-0.006 (0.001)
ln(Japanese Network)	0.010 (0.0004)	0.019 (0.000)	-0.003 (0.001)
ln(Distance)	0.014 (0.001)	0.008 (0.001)	0.026 (0.003)
ln(Wage Ratio)	0.078 (0.001)	0.081 (0.001)	0.070 (0.002)
ln(Rent Ratio)	0.090 (0.001)	0.088 (0.001)	0.114 (0.003)
ln(VAT Ratio)	0.240 (0.011)	0.049 (0.015)	0.132 (0.028)
ln(Tariff Ratio)	0.281 (0.012)	0.351 (0.015)	-0.001 (0.030)
Border	0.050 (0.005)	0.070 (0.007)	-0.009 (0.013)
Common language	-0.029 (0.002)	-0.060 (0.003)	0.019 (0.005)
Common currency	-0.019 (0.005)	-0.008 (0.006)	-0.050 (0.011)
Internal EU	-0.129 (0.004)	-0.145 (0.005)	-0.103 (0.011)
Internal US-CAN	-0.047 (0.004)	-0.044 (0.005)	-0.034 (0.009)
Number of observations	258,864	164,933	39,361
R-squared	0.1461	0.1733	0.1641
RMSE	0.3852	0.389	0.379

Appendix Table 2. Price Dispersion and Joint Network Strength, Fixed Effects ("Within") Estimation

Dependent variable: $|\ln P_i - \ln P_j|$
Group variable: city-pair by good
Cluster standard errors, by group variable, in parentheses
Data grouped into two periods: "2000" and "2005"

	All Goods	**Food**	**Apparel**
ln(Chinese Network)	-0.016 (0.002)	-0.013 (0.002)	-0.034 (0.006)
ln(Indian Network)	-0.373 (0.037)	-0.018 (0.046)	-1.472 (0.113)
ln(Japanese Network)	-0.010 (0.007)	-0.019 (0.008)	-0.054 (0.020)
ln(Wage Ratio)	0.390 (0.004)	0.391 (0.005)	0.308 (0.012)
ln(Rent Ratio)	0.154 (0.004)	0.114 (0.005)	0.316 (0.012)
ln(VAT Ratio)	-0.742 (0.771)	x	-5.607 (1.468)
ln(Tariff Ratio)	0.479 (0.025)	0.428 (0.030)	0.699 (0.069)
Observations	258,864	164,933	39,361
Groups	199,115	125,770	31,246
R-square: within	0.2569	0.2416	0.2923
(between)	0.0368	0.0951	0.0050
(overall)	0.0396	0.0895	0.0075

Appendix Table 3. Elasticities of Price Dispersion w.r.t. Ethnic Density—Estimated from the Top Quartile of Ethnic Density

Each entry is an elasticity (with cluster SE) estimated from fixed effects ("within") estimation of price dispersion on the top quartile of ethnic density observations; full results not reported.

	Ethnic Network		
	Chinese	**Indian**	**Japanese**
All goods	-0.017 (0.001)	0.007 (0.001)	0.009 (0.002)
Food	-0.017 (0.001)	0.012 (0.001)	-0.018 (0.002)
Apparel	-0.037 (0.002)	0.0006 (0.002)	0.038 (0.005)

Appendix 1. Tradable Goods

Source: EIU CityData, authors' characterization as "tradable" and, a subset of tradables, "arbitrageable" (italicized).
Where applicable, all items are chosen from the "supermarket" or "chain store" category.
Category designations are from the EIU.

Food	**Food (cont.)**	**Household Goods**
Apples (1 kg)	*Spaghetti (1 kg)*	Batteries, D-LR20 (2)
Bacon (1 kg)	*Sugar, white (1 kg)*	Dishwashing soap (750 ml)
Bananas (1 kg)	*Tea bags (25)*	Frying pan, Teflon
Beef, entrecote (1 kg)	*Tomatoes (1 kg)*	Insecticide spray (330 g)
Beef, filet mignon (1 kg)	*Tomatoes, canned (250 g)*	Laundry detergent (3 l)
Beef, ground (1 kg)	*Veal chops (1 kg)*	*Light bulb, 60 watt (2)*
Beef, roast (1 kg)	*Veal fillet (1 kg)*	*Soap (100 g)*
Beef, shoulder (1 kg)	*Veal roast (1 kg)*	Toaster, electric
Bread, white (1 kg)	*Water, mineral (1 l)*	Toilet tissue (2 rolls)
Butter (500 g)	*Water, tonic (1 l)*	
Carrots (1 kg)	*Yogurt, natural (150 g)*	**Personal Care Goods**
Cheese, imported (500 g)		*Aspirin (100 tablets)*
Chicken, fresh (1 kg)	**Alcohol**	Lipstick, deluxe
Chicken, frozen (1 kg)	Beer, local (1 l)	Lotion, hand (125 ml)
Coca-Cola (1 l)	Beer, quality (1 l)	*Razor blades (5)*
Cocoa (250 g)	Cognac, French (700 ml)	*Shampoo (400 ml)*
Cocoa, beverage (500 g)	Gin (700 ml)	*Tissues, facial (100)*
Coffee, ground (500 g)	Liqueur (700 ml)	*Toothpaste (120 g)*
Coffee, instant (125 g)	Scotch whisky (700 ml)	
Cornflakes (375 g)	Vermouth (1 l)	**Recreation**
Eggs (12)	Wine, common (750 ml)	Compact disc album
Fish, fresh (1 kg)	Wine, fine (750 ml)	Film, Kodak color (36 exp.)
Fish, frozen (1 kg)	Wine, superior (750 ml)	News magazine, Time
Flour, white (1 kg)		Newspaper, daily local
Ham, whole (1 kg)	**Clothing**	Newspaper, international
Lamb, chops (1 kg)	Boys dress trousers	Novel, paperback (bookstore)
Lamb, leg (1 kg)	Boys jacket	Personal computer (64 MB)
Lamb, stewing (1 kg)	Childs jeans	Television, color (66 cm)
Lemons (1 kg)	Childs shoes, dress	
Lettuce (one)	Childs shoes, sport	**Tobacco**
Margarine (500 g)	Girls dress	Cigarettes, local (20)
Milk, pasteurized (1 l)	Mens raincoat	Cigarettes, Marlboro (20)
Mushrooms (1 kg)	Mens shirt	Pipe tobacco (50 g)
Olive oil (1 l)	Mens shoes	
Onions (1 kg)	Mens suit	**Automobiles**
Orange juice (1 l)	Socks, wool	Car, compact (high)
Oranges (1 kg)	Women's dress	Car, compact (low)
Peaches, canned (500 g)	Womens panty hose	Car, deluxe (high)
Peanut or corn oil (1 l)	Womens raincoat	Car, deluxe (low)
Peas, canned (250 g)	Womens shoes	Car, family (high)
Pineapples, sliced (500 g)	Womens sweater	Car, family (low)
Pork chops (1 kg)		Car, low-priced (high)
Pork loin (1 kg)	**Social**	Car, low-priced (low)
Potatoes (2 kg)	Tennis balls (6)	Gas, regular unleaded (1 l)
Rice, white (1 kg)		

Appendix 2. Cities and Countries

As described in text, these are all cities for which EIU price data exists, in countries for which Chinese population share data exists.

City	Country	City	Country	City	Country
Buenos Aires	Argentina	Athens	Greece	Barcelona	Spain
Adelaide	Australia	Budapest	Hungary	Madrid	Spain
Brisbane	Australia	Mumbai	India	Stockholm	Sweden
Melbourne	Australia	New Delhi	India	Geneva	Switzerland
Perth	Australia	Jakarta	Indonesia	Zurich	Switzerland
Sydney	Australia	Tehran	Iran	Bangkok	Thailand
Vienna	Austria	Dublin	Ireland	Istanbul	Turkey
Brussels	Belgium	Tel Aviv	Israel	London	UK
Sao Paulo	Brazil	Milan	Italy	Manchester	UK
Calgary	Canada	Rome	Italy	Atlanta	US
Montreal	Canada	Osaka	Japan	Boston	US
Toronto	Canada	Tokyo	Japan	Chicago	US
Vancouver	Canada	Nairobi	Kenya	Cleveland	US
Santiago	Chile	Seoul	Korea	Detroit	US
Beijing	China	Kuwait	Kuwait	Honolulu	US
Dalian	China	Tripoli	Libya	Houston	US
Guangzhou	China	Kuala Lumpur	Malaysia	Lexington	US
Hong Kong	China	Mexico City	Mexico	Los Angeles	US
Qingdao	China	Casablanca	Morocco	Miami	US
Shanghai	China	Amsterdam	Netherlands	Minneapolis	US
Shenzhen	China	Auckland	New Zealand	New York	US
Suzhou	China	Wellington	New Zealand	Pittsburgh	US
Tianjin	China	Lagos	Nigeria	San Francisco	US
Bogota	Colombia	Oslo	Norway	San Juan	US
Copenhagen	Denmark	Karachi	Pakistan	Seattle	US
Quito	Ecuador	Asuncion	Paraguay	Washington D.C.	US
Cairo	Egypt	Lima	Peru	Montevideo	Uruguay
Helsinki	Finland	Manila	Philippines	Caracas	Venezuela
Lyon	France	Warsaw	Poland	Singapore	Singapore
Paris	France	Lisbon	Portugal	Taipei	Taiwan
Berlin	Germany	Al Khobar	Saudi Arabia		
Dusseldorf	Germany	Jeddah	Saudi Arabia		
Frankfurt	Germany	Riyadh	Saudi Arabia		
Hamburg	Germany	Johannesburg	South Africa		
Munich	Germany	Pretoria	South Africa		

Appendix 3. Data Sources and Description

Price data comes from EIU CityData as described in the text, 1990-2009. Some cities and some years dropped for obvious data errors. Common USD prices calculated as (local currency prices) x (USD per local currency unit), using the EIU market exchange rate values at time of the price survey. 2009 data are from the April survey,1990-2008 from the November survey.

Chinese population data is taken largely from various editions of the *Overseas Chinese Economy Yearbook*, published by the Overseas Chinese Affairs Commission, R.O.C., and supplemented by additional data from the website of the same organization (http://www.ocac.gov.tw/stat/chinese/cstat.htm). This was the main source of data reported by Poston *et al* (1994), Kumugai (2007), and Rauch and Trindade (2002). Selected additional data was obtained from the *CIAWorld Fact Book* (for Taiwan); from the Hong Kong Census and Statistics Department (http://www.censtatd.gov.hk/hong_kong_statistics/statistical_tables/index.jsp?htmlTableID=139&excelID=&chartID=&tableID=139&ID=&subjectID=1), for Hong Kong; and from the *China Statistical Yearbook 2010* report of China's National Population Census, for China.

Japanese population data is from tabulations made by the Ministry of Foreign Affairs and reported by that ministry and by the Japanese Statistical Bureau. JSB data is found at http://www.stat.go.jp/english/data/nenkan/1431-02.htm and http://www.stat.go.jp/english/data/chouki/02.htm. MFA data is available at http://www.mofa.go.jp/region/index.html

Indian population data for 2001 and 2009 is from http://www.moia.gov.in/. Values for 2002-2008 were imputed using geometric growth rates between the 2001 and 2009 endpoints.

Distance is in miles using a great circle distance formula and geographic coordinates of each city.

Factor prices. Wages used in the relative wage variable are taken to be the hourly rate for maid service as reported in EIU CityData. Rents are those reported by EIU CityData for unfurnished two-bedroom apartments.

VAT, GST and local sales tax data are assembled from a wide variety of sources, including

(a) European Commission - Taxation and Customs Union "VAT Rates Applied in the Member States of the European Union" (2010);
(b) USCIB - ATA Carnet Value Added Taxes http://www.uscib.org/index.asp?documentID=1676;
(c) The TMF Group - http://www.tmf-vat.com/international-vat-rates-2010.html;
(d) TaxRates http://www.taxrates.cc;
(e) Meridian Global Services http://www.meridianglobalservices.com.

Expressing VATs/sales taxes as proportions and not as percentage rates—that is, as "0.10" rather than "10 percent"—relative taxes for (i,j) location pairs are calculated as ln[(1+iVAT)/(1+jVAT)].

Tariff data is from the World Bank's *World Development Indicators*' annual average national tariffs for agricultural and manufactured goods, which we attributed to food and all other products, respectively. This is a crude measure of actual tariffs; our preferred interpretation is that this variable is an index of relative trade protection across individual (i,j) location pairs. As with VATs, this variable is constructed as ln[(1+iTariff)/(1+jTariff)] where tariffs enter as proportions, not percentage rates.

Language is a binary variable coded to be unity when location pairs share an official language; designations are based on the official language information provided by the *CIA World Factbook*.

References

Aker, Jenny C., Michael W. Klein, Stephen A. O'Connell and Muzhe Yang. 2010. "Are Borders Barriers? The Impact of International and Internal Ethnic Borders on Agricultural Markets in West Africa." (April). Available at SSRN: http://ssrn.com/abstract= 1646622.

Anderson, James and Eric Van Wincoop. 2004. "Trade Costs." *Journal of Economic Literature*, Vol. 42, No. 3. (September), pp. 691-751.

Anderson, Michael A., Kurt C. Schaefer and Stephen L.S. Smith, 2010, "A Spatial Model of Price Dispersion: Theoretical and Empirical Findings," processed.

Asplund, Marcus and Richard Friberg. 2001. "The Law of One Price in Scandinavian Duty-Free Stores." *American Economic Review*. Vol. 91, No. 4 (September) pp. 1072-1083.

Bandyopadhyay, Subhayu, Cletus Coughlin and Howard J. Wall. 2008. "Ethnic Networks and State Exports." *Review of International Economics*. Vol. 16, No. 1 pp. 199-213.

Boylan, Richard T. 1992. "Laws of large numbers for dynamical systems with randomly matched individuals." Journal of Economic Theory. Vol. 57, No. 2 (August), pp. 473-504.

Crucini, Mario J., Chris I. Telmer, and Marios Zachariadis. 2005 "Understanding European Real Exchange Rates," *American Economic Review* Vol. 95, No. 3, pp. 724-738.

Engel, Charles, and John H. Rogers. 1996. "How Wide is the Border?" *American Economic Review* Vol. 86, No. 5, pp. 1112-25.

Feenstra, Robert and Gordon H. Hanson. 2004. "Intermediaries in Entrepôt Trade: Hong Kong Re-Exports of Chinese Goods," *Journal of Economics & Management Strategy*, Vol. 13, No. 1 (spring), pp. 3-35.

Felbemayr, Gabriel, Benjamin Jung and Farid Toubal. 2009. "Ethnic Networks, Information, and International Trade: Revisiting the Evidence." University of Hohenheim Discussion Paper Number 306.

Kumagai, Satoru. 2007, "Comparing the Networks of Ethnic Japanese and Ethnic Chinese in International Trade," IDE Discussion Paper. No. 113.

Nakamura, Emi and Dawit Zerom, 2009. "Accounting for Incomplete Pass-Through," NBER Working Papers 15255, National Bureau of Economic Research.

Ozlem Inanc and Marios Zachariadis, 2006. "International Price Dispersion and the Direction of Trade," University of Cyprus Working Papers in Economics.

Poston, Dudley L., Michael Xinxiang Mao and Mei-Yu Yu. 1994. "The Global Distribution of the Overseas Chinese Around 1990." *Population and Development Review*, Vol. 20, No. 3 (September), pp. 631-645.

Rauch, James E. 2001. "Business and Social Networks in International Trade." *Journal of Economic Literature*, Vol. 39 (December), pp. 1177-1203.

Rauch, James E. and Vitor Trindade. 2002. Ethnic Chinese Networks and International Trade." *The Review of Economics and Statistics*, Vol. 84, No. 1 (February).

Rogoff, Kenneth. 1996. "The Purchasing Power Parity Puzzle." *Journal of Economic Literature*, Vol. 34, No. 2 (June).