**Presented at the Mid-West International Economics Group meetings**
**Notre Dame University**
**April 29, 2011**

# Ethnic Networks and Price Dispersion

Michael Anderson
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 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 24 percent decrease in price dispersion. We find initial evidence that Indian and Japanese ethnic network effects are much smaller than Chinese network effects, at least in the OECD and Asian countries we use in this study.

**Acknowledgments.** We thank our home institutions for generous financial support; Katherine Boiles and Ian Sturdy (Washington and Lee University), and Julia Mitchell (Gordon College) for able research assistance; and seminar participants at the 2010 Southern Economics Association meetings for comments on an earlier version of this paper.

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, across three ethnic identities that have been found to be associated with increases in bilateral trade, is associated with large decreases in price dispersion.

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 now 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 ethnic networks and international price dispersion? Following RT and the associated literature, ethnic networks provide enforcement of sanctions that deter contract violations, they

---

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.

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]*. In each country the fraction of the population that belongs to an ethnic network is $n$, and the remaining fraction $(1-n)$ do not.

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 system is equivalent to a deterministic system in which the fraction of matches between individuals belonging only to an ethnic network (the co-ethnic network) is $n^2$. The remaining share, $1-n^2$, consists of matches where neither individual belongs to an ethnic network $(1-n)^2$ and mixed matches $(2n(1-n))$, which are equivalent for our purposes.

Each match, which is indexed by $k \in [0,1]$, receives a draw $c(k)$ from a distribution of costs, which represents the resource cost per unit of engaging in trade for good $k$ from a continuum of goods on *[0,1]*.[4] 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 effects of discipline on contracting costs and opportunistic behavior, and also the reduced costs of searching for arbitrage opportunities for those that belong to the network.

The sanction of expulsion is the means by which a co-ethnic network creates discipline amongst its members. This discipline is the key mechanism which ensures lower contracting costs,

[4] Implicit in this framework is symmetry of costs, i.e. for good k, $c(k)$ is the same regardless of the direction of arbitrage.

and decreases opportunistic behavior amongst members. An individual's choice to exercise discipline stems from the value they perceive of belonging to the network, which is reflective of the opportunity cost of being outside it. The value to a member of belonging to the network is thus dependent on the discipline of the other members. As noted this discipline both obviates the need for contracting and reduces cheating. The member also gains through having access to a conduit which provides information about arbitrage opportunities that he or she would otherwise seek out through costly search, which we call network effects.

We present a simple model to show how discipline varies with network size, and given this how the effectiveness of the network influences the cost distribution for the population of matches. The value to an individual of belonging to the co-ethnic network, $v$, depends on the discipline exerted by individuals within the network, $d$, and also the network effects which are proportional to $n^2$. Thus,

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

The level of discipline, $d$, that each member of the network chooses is dependent on the value of the network to the individual, $v$, and has the following properties:[5]

i) for a given $v$, the individual's discipline is decreasing in $n$. There are two reasons for declining discipline: firstly, as $n$ increases enforcing exclusion becomes more costly[6]; secondly as $n$ increases detection of opportunistic behavior becomes more difficult. Both effects alter the cost-benefit calculus of cheating, which is not explicitly modeled here, leading to lower discipline.

[5] We simplify the analysis by assuming all individuals are homogenous, and will thus make the same choices when faced with the same circumstances.

[6] when networks are large, an expelled member who re-enters the network under a different guise will face a lower probability of detection.

ii) for a given $n$, an increase in the value of the network to the individual will lead to an increase in discipline. This increases the opportunity cost of exclusion for the individual. Discipline is modeled as

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

where $\alpha, \beta \in [0,1]$, and $\alpha \geq \beta$ which ensures that $d$ is non-negative. The coefficient $\beta$ represents the sensitivity of discipline to network size (at constant $v$). When $n$ is small, network effects are small and despite the fact that discipline is high at low $n$, because the value of the network is small the level of discipline is low. As $n$ increases and network effects rise, the value of the network rises and so does discipline. Despite rising network effects at higher $n$ and thus higher value, discipline may eventually fall at high $n$, depending on the magnitude of $\beta$ relative to $\alpha$.

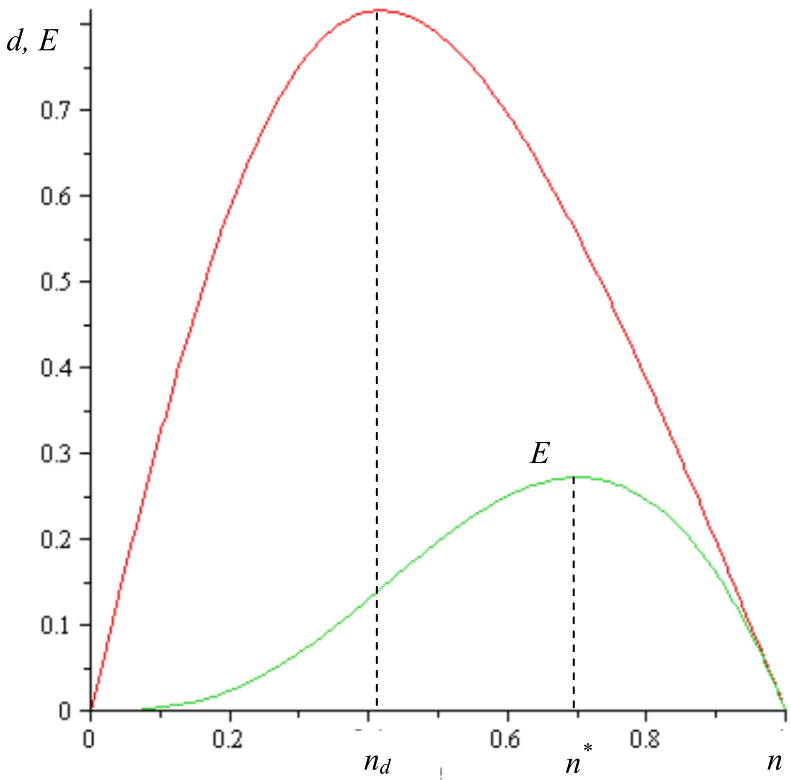
Substituting (2) into (1) leads to an expression for the value of the network as determined by $n$.

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

The numerator represents the network effect, and the denominator the influence of discipline. An increase in $n$ increases the value of the network through larger network effects, and decreases it as discipline falls. Substituting for $v$ in (2) leads to

(4) $d = n^2(\alpha-\beta n) / (1-(\alpha-\beta n))$

Discipline increases for low values of $n$ however beyond the threshold $n_d = (1/\beta)(\alpha\text{-}3/4+(1/4)(9\text{-}8\alpha)^{1/2}$ discipline falls as $n$ increases.[7] The diagram below illustrates the level of discipline as a function of $n$.




The influence of the co-ethnic network on the cost distribution for the population of matches depends on the effectiveness of the network, $E$. The effectiveness of the network depends on the level of discipline of individuals within the network, $d(n)$, and also the number of co-ethnic matches in the population, $n^2$. We model effectiveness as the product of $d(n)$ and $n^2$,

(5) $E = n^2.d(n)$

---

[7] If $\beta \leq (1/\beta)(\alpha\text{-}3/4+(1/4)(9\text{-}8\alpha)^{1/2}$ then $n^* = 1$, and discipline is non-decreasing in $n$.

For example, a small network with a high level of discipline will have a smaller influence on the cost distribution than a large network with slightly lower level of individual discipline. The mean, $\mu(E)$, and variance, $\sigma^2(E)$, of the cost distribution for the population of matches are decreasing in $E$.[8] The effectiveness of the co-ethnic network is at a maximum when $E$ is at its largest. This occurs at[9]

(6) $\quad n^* = -\,2d(n^*)/d'(n^*)$

A match from a population of matches in which the fraction of co-ethnic matches is $(n^*)^2$ draws from a distribution with lower mean and variance than a match from a pair of populations with a different $n$. If $n^* = 1$ then price dispersion is decreasing in $n$.[10] If, on the other hand, $n^* < 1$ then price dispersion is decreasing for $n < n^*$ and increasing for $n > n^*$, as in the diagram above.

For the purposes of illustration we assume the distribution of costs to be Poisson, which has the property that the mean and the variance are equal. Defining the mean and variance as λ, suppose then that the co-ethnic network influences the distribution of cost draws through the following relationship

(7) $\quad \lambda = 1 + 1/(1+E)$

When $E=0$ the distribution of costs is Poisson with $\lambda=2$. If $E > 0$ then $\lambda < 2$.

How does the distribution of costs influence price dispersion? Define $p_i(k)$ as the common-currency price of good $k$ in location $i$; we suppress time subscripts for brevity. Whenever $c(k) <$

---

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

[9] It is straightford to show that $n^* > n_d$ ,

[10] $n^* < 1$ requires that $d'(n^*) < 1$. If $d'(n) \geq 0$ for all $n$ then $n^*=1$. See the diagram above which shows how $E$ varies with $n$.

$|p_i(k)-p_j(k)|$ then 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 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.

## II. Data and Descriptive Analysis

Our empirical work uses data from the Economist Intelligence Unit (EIU). The full data set comprises annual data on the retail prices, including taxes, of (close to) 200 goods and services across 142 cities internationally. From this data we select for this present study 105 available cities from North America, Europe (both EU and non-EU members), Australia and New Zealand, all of Asia (including East Asia, Southeast Asia, South Asia, and Central Asia) and the Middle East.

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 119 of the available EIU goods and services 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. As part of our analysis we present results for two sets of commodities, food and apparel. 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

(8) 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 is large. The mean wedge, over all goods and location pairs, is 0.53 , with a standard deviation of 0.45 (Table 2). In descriptive results not reported here restricting the sample to only countries in the OECD does not substantially reduce mean price dispersion; average price dispersion for this group is about 0.44. Price dispersion is larger between countries than within countries 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. When measured between countries the value is near the overall mean, at 0.55.[11]

Next, consider our measure of network strength for location pairs (i, j), based on ethnic-population shares in the two locations. (Note that we will use the phrases "network strength" and "network density" interchangeable to refer to the total size of the ethnic network.) This variable is defined as:

[11] 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.

(9) Network strength = $\ln(zPopShare_i * zPopShare_j)$ ,

where $kPopShare_i$ 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 RT to measure Chinese ethnic network strength.

We have two years of Chinese population-share data (1990 and 2000). two years of Indian data (2001 and 2009), and one year (2000) of Japanese ethnic-share data. Data sources are detailed in Appendix 3. Table 1 presents a summary of the ethnic-share data (zPopShare), and Table 3 presents a summary of the network strength variable ($zPopShare_i * zPopShare_j$ ) from equation (8). Note that both Table 1 and 3 present unlogged variable summaries.

Several important points are revealed in these two tables. The mean ethnic share variable (Table 1) is a small number for each ethnicity. Said differently, the Chinese, Indians and (especially) the Japanese are minorities nearly everywhere outside their home countries. Next, the Chinese network is the largest of the three networks (Table 3); the mean here is three times the value for the Indian network, and three-hundred times the mean for the Japanese network. Finally, the standard deviation of network strength (Table 3) for each ethnicity is typically 10 times larger than the mean. These huge variances (which give each distribution of network strength a right-skewed shape) will be important later when we consider our regression results of price dispersion on network strength.

In Tables 4-6 we present our measure of price dispersion measure in a cross tab with deciles of the logged product of ethnic population shares (Chinese shares in Table 4, Indian shares in Table 5, Japanese shares in Table 6). In Table 4 we see the mean and variance of price dispersion declining with our measure of network strength, up to the $8^{th}$ decile, where both again begin to increase. This descriptive result suggests that network effectiveness ("E" in equation 5 in our theory section) may be rising for the first seven deciles network size, and declining thereafter. Interestingly, in the subsequent Tables 5 and 6 we also see declines in price dispersion for the first deciles. In Table 5 price dispersion declines for the first seven deciles of Indian network strength, plateauing for the last deciles. The Japanese price-dispersion by network strength comparison, Table 6, follows the pattern for the Chinese data. Mean price dispersion and the variance of price dispersion fall for the first seven deciles of Japanese network strength; both measures show increases for the final three deciles, These results are supported for the food and apparel 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 strength matters for price dispersion. Mean price dispersion decreases, and the variance of price dispersion decreases, as ethnic-network strength increases – though only up to a particular point. Beyond a particular threshold (for two of the networks) both moments of the price-dispersion distribution increase for subsequent increases in network strength.

## 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

$$(10) \qquad P_i = \mu_i (I_i)^{\alpha} (W_i)^{(1-\alpha)}$$

where $\mu_i$ is the markup and $\alpha$ 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

$$(11) \qquad W_i = \phi_i (\mathrm{w}_{i1})^{\gamma} (\mathrm{w}_{i2})^{(1-\gamma)},$$

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

$$(12) \qquad \ln P_i - \ln P_j = (1-\alpha)\ln(\phi_i/\phi_j) + \ln(\mu_i/\mu_j) + \alpha \ln(I_i/I_j) + (1-\alpha)\gamma \ln(\mathrm{w}_{i1}/\mathrm{w}_{j1}) + (1-\alpha)(1-\gamma)\ln(\mathrm{w}_{i2}/\mathrm{w}_{j2}).$$

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 let lemons 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 (9) 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.[12]

[12] 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.

We employ two approaches to estimating this model. First, to take advantage of the multiple years of ethnic network strength data that are available for Chines and Indian networks, we employ the "within" (fixed effects) 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:

$$(13) \quad |\ln P_{ikt} - \ln P_{jkt}| = \beta_1(\text{zEthnicStrength}_{ijkt}) + \beta_2\ln(\text{wagediff}_{ijt}) + \beta_3\ln(\text{rentdiff}_{ijt}) + \beta_4\ln(\text{VATdiff}_{ijkt}) + \beta_5\ln(\text{Tariffdiff}_{ijkt}) + \Sigma\theta_{ijk} + \varepsilon_{ijkt}.$$

The $\theta_{ijk}$ 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 strength variable(s).

Our second estimation technique takes advantage of the fact that we have more years of price dispersion data than of network strength data. Though the network strength variable is measured only in particular years, we impute its value to adjacent years and thus use OLS to make estimates across the entire twenty-year period. Because the ethnic strength variable changes only slowly this seems reasonable; for the Japanese ethnic network, observed only for one year, it is in fact the only technique we can use. The estimating equation:

$$(14) \quad |\ln P_{ikt} - \ln P_{jkt}| = \beta_0 + \beta_1(\text{zEthnicStrength}_{ijt}) + \beta_2\ln\text{Dist}_{ij} + \beta_3\text{Border}_{ij} + \beta_4\ln(\text{wagediff}_{ijt}) + \beta_5\ln(\text{rentdiff}_{ijt}) + \beta_6\ln(\text{VATdiff}_{ijkt}) + \beta_7\ln(\text{Tariffdiff}_{ijkt}) + \beta_8\ln(\text{Common Currency}_{ijt}) + \beta_9\ln(\text{FTA}_{ijt}) + \beta_{10}\,\text{time effects} + \beta_{11}\,\text{city effects} + \varepsilon_{ijkt}.$$

"$Dist_{ij}$" 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).[13]

One final consideration: inclusion of the ethnic networks' home countries risks highly-leveraged regression results. To side-step this potential problem, in all the regressions reported below we exclude any country pairs where one of the locations is the ethnic networks' home country. So, for instance, when the Chinese ethnic network strength variable is included on the right hand side, China, Hong Kong and Taiwan are excluded. The excluded countries all have a close to a 100 percent value for the population share of Chinese. In effect, our regressions look for network effects across "third party" countries, where neither country is the ethnic network's original home nation; as such, our estimates might be considered lower bounds on networks' effects on price dispersion.

[13] 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.

## IV. Estimation Results

In this discussion we focus on the fixed effects estimations. First consider the results for Chinese networks, reported in Table 7. All coefficient values are statistically significant, owing in part to large sample sizes. The columns show the results for all goods and then for food and apparel , two large categories in the EIU dataset. All coefficients have expected signs.

For all goods the regression returns a coefficient value of -0.020 for the elasticity of price dispersion with respect to Chinese ethnic strength. 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 strength variable (relative to the mean) is a 1217 percent change in network strength. Multiplying this 1217 by -0.020 yields a predicted decrease in price dispersion of 24.3 percent. (This and similar calculations for the economic significance of our estimated ethnic coefficients are tabulated in Table 14.) In effect, the one standard deviation increase in network strength reduces price dispersion by more than half of a standard deviation (globally-measured), a not inconsiderable effect. Results for food and apparel are similarly substantial; the price dispersion coefficients translate into predicted reductions in price dispersion of 17 and 46 percent, respectively, when Chinese ethnic strength rises by one standard deviation.

Results on Indian networks are reported in Table 8. The fixed effects estimator applied to 2001 and 2009 returns a mixed outcome. Food products display the expected negative elasticity between price dispersion and ethnic network strength (-0.008, highly significant), which implies that a one standard deviation increase in network strength reduces price dispersion by 6.9 percent (Table 14). The elasticity is unexpectedly positive in apparel and overall. This may be evidence that all ethnic networks are not alike; perhaps there are aspects of the Chinese diaspora that allow Chinese networks to have a larger effect on price dispersion than Indian networks. It is also possible that this finding may be reversed when we include data from Africa, where Indian networks may be particularly strong in East Africa.

Table 9 reports fixed effect results where both Chinese and Indian networks are use as controls. The results should be interpreted with caution, as both Chinese and Indian home countries have been dropped for this estimation, which may penalize the estimation of the effects of the other (non-home) network. Note further that to make this estimation we had to impute Chinese and Indian network strength to years outside the years in which it is measured (for instance, imputing Chinese ethnic network strength in 1990 to all of the years in 1990-1999). Qualitatively the results are broadly similar to what is found in the individual regressions: Chinese network effects are negative across all goods, food and apparel, but the effects are smaller. Indian networks have the expected negative sign only in food.

With a single year of data on the Japanese ethnic network's strength, we estimate its effect on price dispersion with OLS; results are reported in Table 12. The network elasticity is negative for all goods, food and apparel, though it translates into relatively small effects on price dispersion: -4 percent, -4 percent and -3 percent, respectively (Table 14). Comparable OLS estimates of the Chinese and Indian network effects (Tables 10, 11) all have the expected negative sign and imply price dispersion reductions of between 11 and 12 percent—that is, substantially larger effects than for Japanese networks.

Finally, we tested a number of specifications to assess whether network strength's effectiveness in reducing price dispersion erodes at high levels. In contrast to our descriptive results we found no evidence of this, though we continue to explore it.

## V. Conclusion

Our study offers a new explanation for price dispersion. Our theory predicts that strong ethnic networks diminish the costs of arbitrage and lead to price convergence. Conversely, pairs of locations with weak ethnic networks will see larger price dispersion. Both descriptive data analysis and our empirical model support that prediction, though with the surprising finding that not all ethnic networks are the same. Our main finding is that 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. Looked at differently, we predict mean price dispersion will decrease by about 50 percent when ethnic-network strength is one standard deviation above its mean. Interestingly, neither Japanese nor Indian ethnic networks appear to reduce price dispersion to the same extent, though we continue to investigate this point.

Our theory's second prediction, that network effectiveness may decline as ethnic population shares increase beyond some threshold, finds only limited support, though our investigation on this point continues.

**Table 1. Ethnic Shares: Summary of Country Data, Percent**

| | Chinese 1990 | Chinese 2000 | Indian 2001 | Indian 2010 | Japanese 2000 |
|---|---|---|---|---|---|
| Number of countries | 55 | 83 | 79 | 86 | 83 |
| Minimum | 0.000084 | 0 | 0.000024 | 0.000171 | 0.000100 |
| Maximum | 98.00 | 98.38 | 31.92 | 44.22 | 100.00 |
| Mean | 7.68 | 4.90 | 1.76 | 2.76 | 1.24 |
| Standard deviation | 24.09 | 19.14 | 5.25 | 8.17 | 10.97 |

Sources: author's calculations from sources detailed in Appendix 3.

**Table 2. Global Price Dispersion**

Price dispersion between locations i and j measured by |ln(Pi/Pj)|
All tradable goods, all regions (OECD, all Asia, and Middle East)

| Year | Mean | Standard Deviation | Number of Observations |
|---|---|---|---|
| 1990 | 0.59 | 0.51 | 242,115 |
| 1991 | 0.57 | 0.49 | 244,791 |
| 1992 | 0.56 | 0.48 | 251,993 |
| 1993 | 0.54 | 0.47 | 272,696 |
| 1994 | 0.53 | 0.46 | 281,567 |
| 1995 | 0.52 | 0.44 | 283,140 |
| 1996 | 0.49 | 0.42 | 291,911 |
| 1997 | 0.48 | 0.41 | 294,281 |
| 1998 | 0.49 | 0.43 | 319,285 |
| 1999 | 0.49 | 0.43 | 341,436 |
| 2000 | 0.51 | 0.45 | 379,690 |
| 2001 | 0.52 | 0.44 | 389,513 |
| 2002 | 0.52 | 0.44 | 389,384 |
| 2003 | 0.52 | 0.45 | 392,306 |
| 2004 | 0.53 | 0.45 | 401,803 |
| 2005 | 0.54 | 0.46 | 412,222 |
| 2006 | 0.54 | 0.46 | 414,524 |
| 2007 | 0.54 | 0.45 | 417,724 |
| 2008 | 0.54 | 0.45 | 418,736 |
| 2009 | 0.53 | 0.44 | 407,822 |
| OVERALL | 0.53 | 0.45 | 6,846,859 |

Source: Authors’ calculations from EIU CityData, described in Appendices 1, 2, and 3.

**Table 3. Network Strength—Summary Statistics**

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

| | CHINA | INDIA | JAPAN |
|---|---|---|---|
| | Excluding China, Taiwan and Hong Kong | Excluding India, Bangladesh and Pakistan | Excluding Japan |
| Mean | .0003656 | .0001327 | 0.000000974 |
| SD | .0044493 | .0011422 | 0.00000203 |
| Minimum | 4.47E-12 | 5.75E-14 | 4.00E-12 |
| Maximum | .2330015 | .0588396 | 0.0000222 |
| Obs. | 497,939 | 686,274 | 334,570 |

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

Note: All calculations are based on pooled observations across the one or two exact years for which the ethnic share data pertains, using goods and city pairs for which price dispersion is not missing. For China this is 1990 and 2000; for India, 2001 and 2010; and for Japan, 2000. Indian ethnic share data does not include India, Bangladesh and Pakistan.

**Table 4. Mean Price Dispersion by Chinese Ethnic Network Strength (Deciles)**

**A. All Regions**

| | All Goods | Food Products | Apparel |
|---|---|---|---|
| Decile | Mean (SD) | Mean (SD) | Mean (SD) |
| 1 | 0.71 (0.60) | 0.79 (0.60) | 0.69 (0.59) |
| 2 | 0.59 (0.54) | 0.66 (0.58) | 0.63 (0.53) |
| 3 | 0.56 (0.49) | 0.60 (0.52) | 0.57 (0.48) |
| 4 | 0.50 (0.42) | 0.54 (0.43) | 0.48 (0.40) |
| 5 | 0.52 (0.43) | 0.59 (0.45) | 0.47 (0.39) |
| 6 | 0.50 (0.43) | 0.54 (0.44) | 0.45 (0.39) |
| 7 | 0.41 (0.37) | 0.47 (0.43) | 0.39 (0.36) |
| 8 | 0.46 (0.40) | 0.51 (0.42) | 0.48 (0.42) |
| 9 | 0.54 (0.45) | 0.53 (0.45) | 0.54 (0.45) |
| 10 | 0.53 (0.40) | 0.54 (0.41) | 0.47 (0.37) |
| N per decile | 56,776 | 26,358 | 7,881 |

**B. Excluding China, Taiwan and Hong Kong**

| | All Goods | Food Products | Apparel |
|---|---|---|---|
| Decile | Mean (SD) | Mean (SD) | Mean (SD) |
| 1 | 0.72 (0.60) | 0.79 (0.62) | 0.69 (0.59) |
| 2 | 0.61 (0.55) | 0.69 (0.60) | 0.65 (0.54) |
| 3 | 0.58 (0.51) | 0.64 (0.55) | 0.61 (0.50) |
| 4 | 0.51 (0.43) | 0.55 (0.45) | 0.50 (0.42) |
| 5 | 0.50 (0.43) | 0.55 (0.43) | 0.44 (0.36) |
| 6 | 0.50 (0.41) | 0.57 (0.45) | 0.46 (0.39) |
| 7 | 0.49 (0.42) | 0.55 (0.45) | 0.44 (0.38) |
| 8 | 0.40 (0.37) | 0.45 (0.40) | 0.40 (0.37) |
| 9 | 0.44 (0.37) | 0.47 (0.39) | 0.46 (0.40) |
| 10 | 0.50 (0.41) | 0.50 (0.41) | 0.57 (0.45) |
| N per decile | 49,794 | 24,800 | 6,881 |

Pooled observations from 1990 and 2000.

**Table 5. Mean Price Dispersion by Indian Ethnic Network Strength (Deciles)**

**Excluding India, Bangladesh and Pakistan**

| | All Goods | Food Products | Apparel |
|---|---|---|---|
| Decile | Mean (SD) | Mean (SD) | Mean (SD) |
| 1 | 0.57 (0.45) | 0.61 (0.50) | 0.55 (0.44) |
| 2 | 0.59 (0.46) | 0.65 (0.49) | 0.61 (0.47) |
| 3 | 0.55 (0.44) | 0.59 (0.45) | 0.54 (0.44) |
| 4 | 0.52 (0.42) | 0.54 (0.44) | 0.54 (0.44) |
| 5 | 0.48 (0.39) | 0.52 (0.41) | 0.48 (0.38) |
| 6 | 0.47 (0.38) | 0.49 (0.38) | 0.51 (0.42) |
| 7 | 0.46 (0.37) | 0.49 (0.38) | 0.48 (0.38) |
| 8 | 0.42 (0.40) | 0.44 (0.36) | 0.44 (0.36) |
| 9 | 0.43 (0.35) | 0.47 (0.36) | 0.45 (0.37) |
| 10 | 0.43 (0.35) | 0.44 (0.34) | 0.50 (0.41) |
| N per decile | 68,628 | 33,468 | 9,188 |

Pooled observations across 2001 and 2010.

**Table 6. Mean Price Dispersion by Japanese Ethnic Network Strength (Deciles)**

**Excluding Japan; data from 2000**

| | All Goods | Food Products | Apparel |
|---|---|---|---|
| Decile | Mean (SD) | Mean (SD) | Mean (SD) |
| 1 | 0.71 (0.60) | 0.79 (0.60) | 0.69 (0.59) |
| 2 | 0.59 (0.54) | 0.66 (0.58) | 0.63 (0.53) |
| 3 | 0.56 (0.49) | 0.60 (0.52) | 0.57 (0.48) |
| 4 | 0.50 (0.42) | 0.54 (0.43) | 0.48 (0.40) |
| 5 | 0.52 (0.43) | 0.59 (0.45) | 0.47 (0.39) |
| 6 | 0.50 (0.43) | 0.54 (0.44) | 0.45 (0.39) |
| 7 | 0.41 (0.37) | 0.47 (0.43) | 0.39 (0.36) |
| 8 | 0.46 (0.40) | 0.51 (0.42) | 0.48 (0.42) |
| 9 | 0.54 (0.45) | 0.53 (0.45) | 0.54 (0.45) |
| 10 | 0.53 (0.40) | 0.54 (0.41) | 0.47 (0.37) |
| N per decile | 56,776 | 26,358 | 7,881 |

**Table 7. Chinese Network Effects—"Within" (Fixed Effects) Estimator, 1990 and 2000**

**All regressions exclude China, Taiwan and Hong Kong**
**Dependent variable: |ln$P_i$ - $P_j$|**
**Group variable: city-pair by good**
**Cluster standard errors, by group variable, in parentheses**

| | **All Goods** | **Food** | **Apparel** |
|---|---|---|---|
| ln(Chinese Network) | -0.020 (0.001) | -0.014 (0.001) | -0.038 (0.002) |
| ln(Wage Ratio) | 0.073 (0.001) | 0.079 (0.002) | 0.056 (0.004) |
| ln(Rent Ratio) | 0.080 (0.002) | 0.098 (0.002) | 0.048 (0.004) |
| ln(VAT Ratio) | 0.076 (0.018) | 0.055 (0.025) | 0.061 (0.040) |
| ln(Tariff Ratio) | 0.888 (0.018) | 0.985 (0.023) | 0.660 (0.039) |
| Number of observations | 342,012 | 216,746 | 54,124 |
| Number of groups | 211721 | 135565 | 32222 |
| R-square: within | 0.084 | 0.097 | 0.054 |
| (between) | 0.122 | 0.140 | 0.105 |
| (overall) | 0.116 | 0.136 | 0.072 |

**Table 8. Indian Network Effects—"Within" (Fixed Effects) Estimator, 2001 and 2009**

**All regressions exclude India, Bangladesh and Pakistan**
**Dependent variable: $|\ln(P_i/P_j)|$**
**Group variable: city-pair by good**
**Cluster standard errors, by group variable, in parentheses**

| | All Goods | Food | Apparel |
|---|---|---|---|
| ln(Indian Network) | 0.004 (0.000) | -0.008 (0001) | 0.028 (0.001) |
| ln(Wage Ratio) | 0.039 (0.001) | 0.038 (0.001) | 0.041 (0.002) |
| ln(Rent Ratio) | 0.057 (0.001) | 0.047 (0.001) | 0.085 (0.003) |
| ln(VAT Ratio) | 0.392 (0.013) | 0.203 (0.018) | 0.520 (0.030) |
| ln(Tariff Ratio) | 0.134 (0.010) | 0.113 (0.013) | 0.066 (0.025) |
| Number of observations | 476,166 | 297,138 | 73,767 |
| Number of groups | 269,024 | 170,134 | 40,869 |
| R-square: within | 0.029 | 0.024 | 0.067 |
| between | 0.068 | 0.123 | 0.010 |
| overall | 0.058 | 0.102 | 0.015 |

**Table 9. Chinese and Indian Network Effects Combined—"Within" (Fixed Effects) Estimator, 1990-2009**

**All regressions exclude China, Taiwan and Hong Kong and also India, Bangladesh, and Pakistan**
**Dependent variable: $|\ln(P_i/P_j)|$**
**Group variable: city-pair by good**
**Cluster standard errors, by group variable, in parentheses**

| | All Goods | Food | Apparel |
|---|---|---|---|
| ln(Chinese Network) | -0.003 (0.001) | -0.010 (0.001) | 0.012 (0.002) |
| ln(Indian Network) | 0.006 (0.001) | -0.013 (0.001) | 0.040 (0.002) |
| ln(Wage Ratio) | 0.033 (0.000) | 0.033 (0.000) | 0.030 (0.001) |
| ln(Rent Ratio) | 0.064 (0.000) | 0.068 (0.001) | 0.065 (0.001) |
| ln(VAT Ratio) | 0.234 (0.006) | 0.131 (0.007) | 0.236 (0.013) |
| ln(Tariff Ratio) | 0.073 (0.003) | 0.092 (0.004) | 0.119 (0.007) |
| Number of observations | 3,303,174 | 2,084,307 | 519,354 |
| Number of groups | 201,359 | 129,391 | 30,332 |
| R-square: within | 0.028 | 0.030 | 0.039 |
| (between) | 0.156 | 0.163 | 0.004 |
| (overall) | 0.078 | 0.099 | 0.005 |

**Note**: Chinese network strength data measured in 1990 was imputed to years 1990-1999, and network strenghth measured in 2000 was imputed to 2000-2009; Indian network strength measured in 2001 was imputed to 1990-2001, and network strength measure in 2010 was imputed to 2002-2009.

### Table 10. Price Dispersion and Chinese Network Strength, 1990-2009, OLS Estimation

**All regressions exclude China, Taiwan and Hong Kong**
**Dependent variable: $|\ln(P_i/P_j)|$**
**Robust standard errors in parentheses**

| | All Goods | Food | Apparel |
|---|---|---|---|
| ln(Chinese Network) | -0.009 (0.000) | -0.009 (0.000) | -0.010 (0.001) |
| ln(Distance) | 0.021 (0.000) | 0.023 (0.000) | 0.021 (0.000) |
| ln(Wage Ratio) | 0.098 (0.001) | 0.107 (0.001) | 0.103 (0.002) |
| ln(Rent Ratio) | 0.069 (0.001) | 0.064 (0.001) | 0.102 (0.001) |
| ln(VAT Ratio) | 0.242 (0.005) | -0.198 (0.009) | -0.452 (0.010) |
| ln(Tariff Ratio) | 0.043 (0.004) | 0.089 (0.005) | 0.012 (0.008) |
| Border | 0.014 (0.001) | 0.015 (0.001) | -0.015 (0.003) |
| Common language | -0.040 (0.001) | -0.044 (0.001) | -0.056 (0.001) |
| Common currency | -0.055 (0.001) | -0.054 (0.001) | -0.054 (0.002) |
| Internal EU | -0.035 (0.001) | -0.025 (0.001) | -0.006 (0.002) |
| Internal US-CAN | -0.011 (0.001) | -0.008 (0.001) | -0.004 (0.002) |
| Number of observations | 3,616,026 | 2,285,452 | 566,384 |
| R-squared | 0.215 | 0.238 | 0.393 |
| RMSE | 0.385 | 0.385 | 0.345 |

**Note**: Chinese network strength data measured in 1990 is imputed to years 1990-1999, and network strenghth measured in 2000 is imputed to 2000-2009

**Table 11. Price Dispersion and Indian Network Strength, 1990-2009, OLS Estimation**

**All regressions exclude India, Bangladesh and Pakistan**
**Dependent variable: $|\ln(P_i/P_j)|$**
**Robust standard errors in parentheses**

| | All Goods | Food | Apparel |
|---|---|---|---|
| ln(Indian Network) | -0.012 (0.000) | -0.014 (0.000) | -0.010 (0.000) |
| ln(Distance) | 0.022 (0.000) | 0.027 (0.000) | 0.012 (0.001) |
| ln(Wage Ratio) | 0.102 (0.001) | 0.117 (0.001) | 0.081 (0.002) |
| ln(Rent Ratio) | 0.077 (0.001) | 0.073 (0.001) | 0.107 (0.001) |
| ln(VAT Ratio) | 0.331 (0.005) | 0.142 (0.008) | -0.173 (0.009) |
| ln(Tariff Ratio) | 0.079 (0.003) | 0.101 (0.004) | 0.042 (0.007) |
| Border | 0.012 (0.001) | 0.019 (0.001) | -0.013 (0.003) |
| Common language | -0.044 (0.001) | -0.053 (0.001) | -0.029 (0.001) |
| Common currency | -0.061 (0.001) | -0.057 (0.001) | -0.061 (0.002) |
| Internal EU | -0.016 (0.001) | -0.008 (0.001) | -0.005 (0.002) |
| Internal US-CAN | 0.012 (0.001) | 0.026 (0.001) | 0.016 (0.002) |
| Number of observations | 4,097,441 | 2,566,686 | 644,338 |
| R-squared | 0.164 | 0.198 | 0.393 |
| RMSE | 0.378 | 0.380 | 0.345 |

**Note**: Indian network strength data measured in 2001 is imputed to years 1990-2001, while network strength measured in 2010 is imputed to 2002-2009

**Table 12. Price Dispersion and Japanese Network Strength, 1990-2009, OLS Estimation**

**All regressions exclude Japan**
**Dependent variable: $|\ln(P_i/P_j)|$**
**Robust standard errors in parentheses**

| | All Goods | Food | Apparel |
|---|---|---|---|
| ln(Japanese Network) | -0.018 (0.000) | -0.018 (0.000) | -0.013 (0.000) |
| ln(Distance) | 0.019 (0.000) | 0.021 (0.000) | 0.010 (0.001) |
| ln(Wage Ratio) | 0.090 (0.001) | 0.097 (0.001) | 0.095 (0.002) |
| ln(Rent Ratio) | 0.071 (0.001) | 0.063 (0.001) | 0.090 (0.001) |
| ln(VAT Ratio) | 0.154 (0.005) | 0.107 (0.008) | -0.251 (0.009) |
| ln(Tariff Ratio) | 0.174 (0.004) | 0.227 (0.005) | 0.048 (0.008) |
| Border | -0.016 (0.001) | -0.013 (0.001) | -0.032 (0.003) |
| Common language | -0.039 (0.001) | -0.051 (0.001) | -0.033 (0.001) |
| Common currency | -0.061 (0.001) | -0.062 (0.001) | -0.054 (0.002) |
| Internal EU | -0.046 (0.001) | -0.039 (0.001) | -0.030 (0.002) |
| Internal US-CAN | 0.033 (0.001) | -0.027 (0.001) | -0.016 (0.002) |
| Number of observations | 4,111,426 | 2,600,619 | 638,707 |
| R-squared | 0.178 | 0.178 | 0.372 |
| RMSE | 0.385 | 0.389 | 0.344 |

**Note**: Japanese network strength data is measured in 2000 and imputed to all years 1990-2009.

**Table 13. Price Dispersion and Chinese, Indian and Japanese Network Strength**

**1990-2009, OLS Estimation**
**All regressions exclude China, Taiwan and Hong Kong; India, Bangladesh and Pakistan; and Japan.**
**Dependent variable: $|\ln(P_i/P_j)|$**
**Robust standard errors in parentheses.**

| | **All Goods** | **Food** | **Apparel** |
|---|---|---|---|
| ln(Chinese Network) | 0.006 (0.000) | 0.006 (0.000) | 0.016 (0.000) |
| ln(Indian Network) | -0.001 (0.000) | -0.003 (0.000) | -0.001 (0.000) |
| ln(Japanese Network) | -0.027 (0.000) | -0.023 (0.000) | -0.036 (0.001) |
| ln(Distance) | 0.024 (0.000) | 0.026 (0.000) | 0.016 (0.001) |
| ln(Wage Ratio) | 0.094 (0.001) | 0.106 (0.001) | 0.071 (0.002) |
| ln(Rent Ratio) | 0.063 (0.001) | 0.059 (0.001) | 0.095 (0.001) |
| ln(VAT Ratio) | 0.065 (0.005) | -0.136 (0.009) | -0.378 (0.010) |
| ln(Tariff Ratio) | 0.046 (0.004) | 0.085 (0.005) | 0.106 (0.009) |
| Border | -0.000 (0.001) | 0.003 (0.001) | -0.026 (0.003) |
| Common language | -0.025 (0.001) | -0.037 (0.001) | -0.012 (0.001) |
| Common currency | -0.050 (0.001) | -0.048 (0.001) | -0.055 (0.002) |
| Internal EU | -0.015 (0.001) | -0.006 (0.001) | 0.008 (0.002) |
| Internal US-CAN | 0.000 (0.001) | 0.010 (0.001) | 0.006 (0.002) |
| Number of observations | 3,102,005 | 1,965,754 | 483,751 |
| R-squared | 0.133 | 0.144 | 0.279 |
| RMSE | 0.348 | 0.347 | 0.317 |

**Note**: Chinese network strength data measured in 1990 is imputed to years 1990-1999, and network strength measured in 2000 is imputed to 2000-2009; Indian network strength data measured in 2001 is imputed to years 1990-2001, while network strength measured in 2010 is imputed to 2002-2009; for Japan, network strength is measured in 2000 and imputed to all years 1990-2009.

## Table 14. Economic Significance of Ethnic Network Effects on Price Dispersion

| Table | Ethnic Group | Goods | Estimator | Elasticity | Mean Ethnic Network Strength | SD of Ethnic Network Strength | +1 SD: Percent Equiv | Effect on Price Ratio, Percent |
|---|---|---|---|---|---|---|---|---|
| 7 | China | All | FE | -0.020 | 0.0003656 | 0.0044493 | 1217 | -24.3 |
| 7 | China | Food | FE | -0.014 | 0.0003656 | 0.0044493 | 1217 | -17.0 |
| 7 | China | Apparel | FE | -0.038 | 0.0003656 | 0.0044493 | 1217 | -46.2 |
| 8 | India | All | FE | 0.004 | 0.0001327 | 0.0011422 | 861 | 3.4 |
| 8 | India | Food | FE | -0.008 | 0.0001327 | 0.0011422 | 861 | -6.9 |
| 8 | India | Apparel | FE | 0.028 | 0.0001327 | 0.0011422 | 861 | 24.1 |
| 9 | China | All | FE | -0.003 | 0.0003656 | 0.0044493 | 1217 | -3.7 |
| 9 | India | All | FE | 0.006 | 0.0001327 | 0.0011422 | 861 | 5.2 |
| 12 | Japan | All | OLS | -0.018 | 0.0000010 | 0.0000020 | 208 | -3.8 |
| 12 | Japan | Food | OLS | -0.018 | 0.0000009 | 0.0000019 | 209 | -3.8 |
| 12 | Japan | Apparel | OLS | -0.013 | 0.0000010 | 0.0000021 | 207 | -2.7 |

## 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.

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).

## 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**

*Apples (1 kg)*
*Bacon (1 kg)*
*Bananas (1 kg)*
*Beef, entrecote (1 kg)*
*Beef, filet mignon (1 kg)*
*Beef, ground (1 kg)*
*Beef, roast (1 kg)*
*Beef, shoulder (1 kg)*
*Bread, white (1 kg)*
*Butter (500 g)*
*Carrots (1 kg)*
*Cheese, imported (500 g)*
*Chicken, fresh (1 kg)*
*Chicken, frozen (1 kg)*
Coca-Cola (1 l)
*Cocoa (250 g)*
Cocoa, beverage (500 g)
Coffee, ground (500 g)
Coffee, instant (125 g)
Cornflakes (375 g)
*Eggs (12)*
*Fish, fresh (1 kg)*
*Fish, frozen (1 kg)*
*Flour, white (1 kg)*
*Ham, whole (1 kg)*
*Lamb, chops (1 kg)*
*Lamb, leg (1 kg)*
*Lamb, stewing (1 kg)*
*Lemons (1 kg)*
*Lettuce (one)*
*Margarine (500 g)*
Milk, pasteurized (1 l)
*Mushrooms (1 kg)*
*Olive oil (1 l)*
*Onions (1 kg)*
*Orange juice (1 l)*
*Oranges (1 kg)*
*Peaches, canned (500 g)*
*Peanut or corn oil (1 l)*
*Peas, canned (250 g)*
Pineapples, sliced (500 g)
Pork chops (1 kg)
Pork loin (1 kg)
Potatoes (2 kg)
Rice, white (1 kg)

**Food (cont.)**

*Spaghetti (1 kg)*
*Sugar, white (1 kg)*
*Tea bags (25)*
*Tomatoes (1 kg)*
*Tomatoes, canned (250 g)*
*Veal chops (1 kg)*
*Veal fillet (1 kg)*
*Veal roast (1 kg)*
*Water, mineral (1 l)*
*Water, tonic (1 l)*
*Yogurt, natural (150 g)*

**Alcohol**

Beer, local (1 l)
Beer, quality (1 l)
Cognac, French (700 ml)
Gin (700 ml)
Liqueur (700 ml)
Scotch whisky (700 ml)
Vermouth (1 l)
Wine, common (750 ml)
Wine, fine (750 ml)
Wine, superior (750 ml)

**Clothing**

Boys dress trousers
Boys jacket
Childs jeans
Childs shoes, dress
Childs shoes, sport
Girls dress
Mens raincoat
Mens shirt
Mens shoes
Mens suit
Socks, wool
Women's dress
Womens panty hose
Womens raincoat
Womens shoes
Womens sweater

**Social**

Tennis balls (6)

**Household Goods**

Batteries, D-LR20 (2)
Dishwashing soap (750 ml)
Frying pan, Teflon
Insecticide spray (330 g)
Laundry detergent (3 l)
*Light bulb, 60 watt (2)*
*Soap (100 g)*
Toaster, electric
Toilet tissue (2 rolls)

**Personal Care Goods**

*Aspirin (100 tablets)*
Lipstick, deluxe
Lotion, hand (125 ml)
*Razor blades (5)*
*Shampoo (400 ml)*
*Tissues, facial (100)*
*Toothpaste (120 g)*

**Recreation**

Compact disc album
Film, Kodak color (36 exp.)
News magazine, Time
Newspaper, daily local
Newspaper, international
Novel, paperback (bookstore)
Personal computer (64 MB)
Television, color (66 cm)

**Tobacco**

Cigarettes, local (20)
Cigarettes, Marlboro (20)
Pipe tobacco (50 g)

**Automobiles**

Car, compact (high)
Car, compact (low)
Car, deluxe (high)
Car, deluxe (low)
Car, family (high)
Car, family (low)
Car, low-priced (high)
Car, low-priced (low)
Gas, regular unleaded (1 l)

**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 prices are from the EIU's April survey, while 1990-2008 are from the November survey.

**Chinese population data** for 1990 comes from Poston *et al* (1994); we use the Chinese population shares Poston lists as coming from "around 1990," which, for the countries in our sample, are from the 1989-1991 range. Data for 2000 comes from Kumugai (2007).

**Japanese population data** is from Kumagai (2007), and dates from 2000.

**Indian population data** for 2001 and 2010 is from http://www.moia.gov.in/.

**Distance** is calculated 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*.