1-1-2005

Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data

Russell Hillberry
University of Melbourne

David Hummels
Purdue University

Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data

Russell Hillberry
University of Melbourne

David Hummels
Purdue University
NBER

January 2005

Abstract: A large literature has shown that geographic frictions reduce trade, but has not clarified precisely why. In this paper we provide some insight into *why* such frictions matter by focusing on precisely *what they matter for*. We ask: what parts of trade do these frictions reduce most? Using data that tracks manufacturers' shipments within the United States on an exceptionally fine grid, we find that the pattern of shipments is extremely localized. Shipments within 5-digit zip codes, which have a median radius of just 4 miles, are 3 times larger than shipments outside the zip code. We decompose aggregate shipments into extensive and intensive margins, and show that distance and other frictions reduce aggregate trade values primarily by reducing the number of commodities shipped and the number of establishments shipping. Extensive margins are particularly important over very short distances. We consider two broad reasons for these facts and conclude that trade in intermediate goods is the most likely explanation for highly localized shipments and the dominant role of the extensive margin. In another significant finding, we find no evidence of state-level home bias when distances are measured precisely and trade is observed over a very fine grid.

Acknowledgements: We thank the editor and two anonymous referees for very helpful comments. We also thank Jack Barron, Tim Cason, Michael Ferrantino, and John Ries for comments on an earlier draft. We thank Arnold Rezneck of the Center for Economic Studies for his help with the data. All errors are the responsibility of the authors. . Much of the research work occurred while Russell Hillberry was employed by the U.S. International Trade Commission. This research was supported by grants from Purdue CIBER and NSF.

Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data

January 2005

Abstract: A large literature has shown that geographic frictions reduce trade, but has not clarified precisely why. In this paper we provide some insight into *why* such frictions matter by focusing on precisely *what they matter for*. We ask: what parts of trade do these frictions reduce most? Using data that tracks manufacturers' shipments within the United States on an exceptionally fine grid, we find that the pattern of shipments is extremely localized. Shipments within 5-digit zip codes, which have a median radius of just 4 miles, are 3 times larger than shipments outside the zip code. We decompose aggregate shipments into extensive and intensive margins, and show that distance and other frictions reduce aggregate trade values primarily by reducing the number of commodities shipped and the number of establishments shipping. Extensive margins are particularly important over very short distances. We consider two broad reasons for these facts and conclude that trade in intermediate goods is the most likely explanation for highly localized shipments and the dominant role of the extensive margin. In another significant finding, we find no evidence of state-level home bias when distances are measured precisely and trade is observed over a very fine grid.

I. Introduction

How does trade respond to geographic frictions? A large literature has shown that frictions associated with distance and state and national borders reduce trade, but has not clarified precisely why. In this paper we provide some insight into *why* spatial frictions matter by focusing on precisely *what they matter for*. We ask: what parts of trade do these frictions reduce most?

To explain, the value of trade used in empirical work on spatial frictions is an aggregation, first over the price and quantity of each unique variety traded, and then, an aggregation of unique varieties shipped between origin and destination. (This is true even for studies that rely on trade disaggregated by commodity.) Empirical work typically assumes that spatial frictions operate by reducing the traded quantities of each variety. This effect is captured perfectly well by aggregated values if prices and the number of traded varieties do not co-vary with frictions. The approach is theoretically consistent with the models invoked most frequently to underpin the empirical work, and in any case, most researchers lack data sufficiently detailed to assume anything else.

We employ the micro-data file of the Commodity Flow Survey (CFS), which reports establishment-level shipments on an exceptionally fine geographic grid within the United States, and provides consistent measures of trade quantities and prices across hundreds of industries and thousands of establishments. These data allow us to decompose bilateral trade values into several components, and to investigate how each component varies over spatial frictions. We find that the number of commodities shipped and the number of unique establishment x destination pairs in which shipments are observed fall dramatically as the destination distance rises. These distance effects are extraordinarily pronounced, with the number of shipments dropping almost an order of magnitude between 1 and 200 miles, and being nearly flat thereafter.

In our most detailed data we show that shipments destined within the same 5-digit zip code, i.e. within a 4 mile radius of the shipper, are 3 times higher than those outside the zip code.

In marked contrast, average value per unique shipment is relatively constant over short distances, and falls only gradually with distance as distance gets large. In other words, the aggregate trade-distance relationship in intra-US trade over short distances is driven entirely by the fact that most establishments ship only to geographically proximate customers (the extensive margin), rather than shipping to many customers in quantities that decrease in distance (the intensive margin). For the small handful of shipments that reach beyond a few hundred miles, average shipment values do fall gradually with distance, but even at larger distances the extensive margin remains the primary channel through which spatial frictions reduce trade.

Why are shipments localized in this particular way? We consider two explanations, each distinguished by what is assumed about the spatial distribution of demand.

Our primary hypothesis is that goods produced at a distance are not purchased because there is no local industrial demand for them. Suppose that region i produces an intermediate good that is useful only as an input into producing a particular final good. If region j does not produce that final good, it has no use for region i's output. To complete the story, the production "match" (the likelihood that region i produces inputs that region j's industry wants to consume) is stronger when the two regions are proximate. In other words, specialized up- and down-stream establishments are sorted geographically to avoid spatial frictions. In the resulting pattern of trade, most shipments occur only between highly proximate location pairs. Further, what appears to be intra-industry trade occurring between these proximate pairs is the exchange of products at different stages of production (i.e. auto parts for finished autos).

Our alternative hypothesis is consistent with most of the theoretical models used to study spatial frictions. In these models consumers substitute away from distant goods because spatial frictions raise their price relative to local goods. The effects of spatial frictions are most pronounced, and imports lowest, when local goods are in plentiful supply. Distance costs could choke off trade entirely, as opposed to merely reducing quantities, if consumers regard local and distant goods as homogeneous, or if consumers regard them as differentiated but must pay a fixed cost to import.

To separate these hypotheses we provide three exercises. First, we combine data on manufacturing at the zip code level with input usage taken from the U.S. input-output structure in order to generate predicted industrial demands. These intermediate demands help explain variation in industry expenditures across regions. Second, we show that the share of industry imports into a region is positively correlated with that industry's output and export shares. This is inconsistent with a model in which distance costs lead consumers to substitute away from distant and toward local goods. Third, recall that the effect of distance on trade operates primarily in a 1,0 manner, shutting off distant shipments entirely as opposed to reducing their quantities. We use probit regressions to show that a region's output mix, and resulting industrial demands, affects the likelihood a commodity is shipped to that region.

This paper contributes to a number of distinct literatures. Very recently, decompositions that separate trade into extensive margins (number of commodities) and intensive margins (value per commodity) have appeared in the literature. Our decomposition method is inspired by Hummels and Klenow (2004) who examine the response of the extensive and intensive margin to exporter characteristics. Other papers focus on trade policy changes[1], and generally show that tariff liberalizations result in trade growth along both extensive and intensive margins. Unlike these papers, we focus on the spatial dimension of trade within a country, rather than external

[1] Hillberry and McDaniel (2002) and Kehoe and Ruhl (2003) examine changes in the numbers of trade commodities in the US after NAFTA, while Klenow and Rodriguez-Clare (1995) show similar results in the aftermath of trade liberalization in Costa Rica.

trade. Further, we can identify individual establishments, rather than observing all establishments that ship a particular commodity aggregated into a single flow. We show that spatial frictions reduce extensive margins defined in terms of both commodities and unique establishments.

This paper contributes to a large literature[2] that employs distance in import demand (gravity) equations, either as a direct object of interest or as a control. Unlike previous work, we decompose aggregate trade flows into multiple components in an effort to examine what parts of trade spatial frictions act upon.

We also contribute to work on "home bias" or border effects in trade. By exploiting the CFS' fine geographic grid, we show that state level home bias, like that identified by Wolf (2000) or Hillberry and Hummels (2003), is an artifact of geographic aggregation. The nature of the data allows us to examine only internal political boundaries (state borders), and not directly estimate the effect of national boundaries as in McCallum (1995) and Anderson and van Wincoop (2003). However, an explanation of why most establishments ship goods only to very near customers within a country may contribute to understanding why those establishments do not ship to customers outside the country.

Finally, we contribute to the literature on equilibrium location choices of firms, and in particular, models that emphasize the role of intermediate demand in forming industrial agglomerations.[3] We do not offer a formal test of these models, nor do we offer a fully articulated model of firm location. Nevertheless, our evidence regarding regional absorption

[2] Disdier and Head (2004) identify 78 published papers with an estimated elasticity of trade with respect to distance.
[3] Krugman and Venables (1995) derive a theoretic model in which suppliers of intermediates co-locate with final goods producers, thereby generating highly localized trade in intermediate goods. Eaton and Kortum (2002) offer another model with this feature, and fit it to international trade data. One notable feature of the Eaton and Kortum model is the existence of an extensive margin in trade. Amiti and Cameron(2003) find that Indonesian firms in closer proximity to intermediate suppliers pay higher wages. Overman, Midelfart-Knarvik and Venables (2001) find that geographic proximity to intermediate supply and to sources of industrial demand help explain the location of industry in Europe.

variation and the reasons that shipments travel very short distances are consistent with these models.

Section II describes the data in greater detail. Section III describes our methodology for and results on decomposing shipments into various components. Section IV provides explanations for these results and related empirical tests. Section V concludes.

II. Data.

The primary data source we use is the raw data file from the 1997 U.S. Commodity Flow Survey (CFS). The CFS is collected every five years by the U.S. Census Bureau, which chooses a stratified sample of U.S. mining, manufacturing, and wholesale establishments[4]. The sampled establishments report characteristics of a random sample of their shipments. Each shipment record contains the shipment's weight (in pounds), value and SCTG commodity classification,[5] an establishment identifier, the shipper's (SIC) industrial classification, the zip code of the shipment's origin and destination, the actual shipping distance between them, and a sampling weight for each shipment.[6]

These are the best available data documenting intra-national shipments, and are substantially better for our purposes than the publicly available CFS data used by Wolf (2000) and Anderson and van Wincoop (2003), or the Statistics Canada data employed by McCallum (1995) and subsequent authors. There are many advantages to using the CFS raw data file. First,

[4] The Census Bureau unit of analysis is establishments, not firms. Each establishment in multi-establishment firms is treated as distinct.

[5] The Standard Classification of Transported Goods (SCTG) classification system is related to the Harmonized System, with some modifications that suit it to studies of transportation. Our data are reported with 5-digit detail, where there are 512 commodity groupings.

[6] The value and weight of shipments are calculated by multiplying reported estimates by the inverse of the sampling weight. Other data available in our sample, but unused in this study, include a flag for shipments of hazardous materials, the shipment mode used to transport the good, a binary variable denoting shipments bound for export, and the destination of export shipments. Unfortunately, export destination was not a focus of the Census Bureau's data collection effort, and reported information on export destination is noisy and incomplete.

the data are drawn from stratified random samples of actual shipments. This is in sharp contrast with the Statistics Canada data, which are imputed from at least ten distinct data sources.[7]

Second, the total value of trade between two locations is an aggregation over shipments. Unlike other studies, we observe the shipment level details and can unpack this aggregation. This allows us to decompose shipment values into the number of shipments and the average value per shipment. We can also go further, separating values into prices and quantities, and separating numbers of shipments into number of commodities shipped and numbers of unique establishment-destination pairs per commodity shipped.

Third, knowing the SIC code of the establishments allows us to distinguish wholesale shipments from producer shipments. This distinction is important because manufacturers and wholesalers serve an entirely different economic function. Typically, products reach consumers through a spatial value chain: producers ship to wholesalers, and wholesalers ship locally to retailers. Previous work has shown wholesale shipments to be localized to a far greater degree than shipments from producers.[8] Including wholesalers in the sample both double counts shipments and imparts a strong bias in favor of highly localized shipments.

Fourth, because we know the 5-digit postal zip code of the origin establishment and the 5-digit zip code of the destination we are able to describe the geography of shipments on a very fine grid.[9] There are 29,194 such zip codes in our data, with a median population of 2802 persons. The median distance between the central places of a zip code and its closest neighbor is 4 miles. Unlike state or national political borders, these zip codes are allocated in rough proportion to population density. Where economic activity is dense we have a finer geographic grid on which to measure it. In many of our exercises, we employ 3-digit zip code detail. There

[7] The Statistics Canada documentation of the imputation algorithm is 48 pages long.
[8] Hillberry and Hummels (2003).
[9] We also know true distances, how far a shipment must travel given existing road and rail lines, and can employ these rather than "as the crow flies" straight line distances.

are 873 such codes, with a median population of 218,432 persons, and central places a median distance of 21 miles from the central place of their nearest neighbor.[10]

This detail is critical for properly identifying spatial frictions in the data. Most work on trade frictions analyzes trade between locations as if they were dimensionless points in space. But locations, whether measured as nations or as intra-national regions (states or provinces), are geographical aggregates of many possible origin/destination pairs. This can create measurement problems in several important instances.

The literature on home bias examines whether shipments are greater within than across borders, controlling for measured distances. It is common to include shipments within a particular location, which requires researchers to approximate internal distances. (If we want to know whether Texas trades with itself to an unusual degree we must first measure how far is Texas from itself.) If constructed internal distance measures overstate how far shipments travel within a particular geographic grouping, estimates will be biased toward finding that political borders reduce trade.[11] To solve this problem it is necessary to measure actual distances on a very fine grid.

Finally, the effects of distance on trade are typically assumed to be log-linear, with estimated elasticities on the order of -1.0. It is perhaps plausible that doubling distance will halve trade when the distances in question are 500 and 1000 miles. Is it also plausible that trade will be halved when the distances in question are 5 and 10 miles? This is not a question that researchers who lack a very fine geographic grid can answer. Putting in higher order polynomials is no help either, unless one has shipments data that covers the relevant range of

---

[10] When measuring distances between 3-digit zip code regions we use the simple average distance of all 5-digit zip code pairs within those 3-digit regions in which trade occurs.

[11] Related, a robust fact from the literature on spatial frictions is that adjacent regions trade more with each other than can be predicted by measured distances alone. This finding may be a simple artifact of measurement. Distances between neighboring states are commonly measured from central place to central place, but much trade may occur between highly proximate origin-destination pairs lying on either side of a border.

distance variation.

We can capture how far a shipment has actually moved (to a precision of 4 miles), rather than impute it from central place distances. These features allow us to separate the effect of very short shipment distances from "home bias" spuriously created by poor measures of internal distance. And, since origin-destination pairs are distributed continuously from 1 to 3000 miles we are able to assess nonlinear effects of distance on shipments with greater precision.

We also employ a private sample of the 1997 Census of Manufactures. We use these data to generate gross output and value added figures for 4-digit SIC industries at the levels of geography as fine as the 5-digit zip code level. This also allows us to calculate predicted industrial demands at fine levels of geographic detail, by multiplying the vector of zip code level production by a matrix of industrial inputs taken from the U.S. Input-output table (Bureau of Economic Analysis) in 1997.

III. Decomposing trade and spatial frictions

The effects of spatial frictions on trade are commonly assessed using a gravity model of trade. As a point of departure, consider a derivation of this model contained in Krugman (1980). In Krugman, consumers have constant-elasticity of substitution (CES) preferences over differentiated varieties produced by monopolistically competitive firms. Consumers buy all varieties in positive quantities. The quantity of a variety from origin region i purchased by consumers in region j is given by

$$(1) \qquad q_{ij} = E_j \left(\frac{p_i^{-\sigma} t_{ij}^{-\sigma}}{\tilde{P}_j} \right)$$

E_j is region j's total expenditure on the differentiated product, assumed to be independent of the origin region. $p_i t_{ij}$ is the price of i's good at destination j, assumed to vary across regions only because of some iceberg trade cost t_{ij}. That is, because demands are CES and because trade costs are iceberg, or proportional to the value shipped, the same factory gate price is charged to all destinations. Finally, σ is the elasticity of substitution among varieties, and $\tilde{P}_j$ a transformation of the region j price index.

Since the data usually employed do not allow the econometrician to observe the quantities of a particular unique variety it is necessary to aggregate. This requires additional assumptions: all varieties in origin region i are symmetric (prices are equal), and destination region j will consume all these varieties in the same quantity. This allows us to first multiply quantity per variety by prices to get observable values and second to multiply by the number of varieties to get total values shipped. This gives

$$(2) \qquad T_{ij} = p_i n_i q_{ij} = E_j n_i \left(\frac{p_i^{1-\sigma} t_{ij}^{-\sigma}}{\tilde{P}_j} \right)$$

Suppose we are interested in assessing the effect of spatial frictions such as distance on the aggregate value of bilateral trade. The relationship is the same in equations (1) and (2). That is, bilateral variation in spatial frictions affects aggregate trade values only because they affect q_{ij}.[12]

Most studies of spatial frictions stop here simply because they cannot separately identify the individual components in equation (2), except perhaps to estimate commodity level versions of it. Our data, however, allow a closer look. We examine a less restrictive econometric model in which quantities, prices and the number of varieties vary across destinations and can co-vary

[12] Apart from the direct effect operating through t_{ij}, Anderson and van Wincoop (2003) note that there are indirect effects of barriers on exporter prices and importer price indices. However, these effects are constant for a given region, do not vary over partners, and will be controlled for in our estimates. Our focus is on the left hand side of (2), where q_{ij} is the only component of T_{ij} with bilateral variation.

with spatial frictions.

(3) $T_{ij} = p_{ij} n_{ij} q_{ij}$

Prices (exclusive of shipping costs) may vary across destinations if demand is not CES (so that a quantity change induced by frictions also changes the relevant demand elasticity and therefore markups), or if transport costs are not of the iceberg form.[13]

The literature has suggested several reasons why the number of traded varieties might co-vary with spatial frictions. The first and simplest explanation is that the goods produced in locations i and j are homogeneous. If production costs in the two locations are sufficiently similar, or the trade costs sufficiently large, these goods will not be traded. Of course, the further away are trade partners, the more likely goods are to fall into this non-traded category.[14] An example of this phenomenon might be Coca-Cola bottlers, who set up multiple establishments around the country producing identical goods for local shipment.

A slightly more complicated version of this story assumes that consumers view output as differentiated by region, not homogeneous. Starting with the usual CES utility function, demand per variety is given by equation (1). If shipment requires a fixed cost per variety, then spatial frictions may reduce quantities exported to the point that firms can no longer cover fixed costs. In this case, the number of traded varieties will depend negatively on the size of spatial frictions (i.e. be decreasing in distance). As discussed in the introduction, the idea that not all varieties will be traded has substantial support when looking at international data, and there is some evidence that this effect is correlated with ad-valorem costs (tariffs and transport costs). The literature has much less to say about whether the fixed cost story is relevant for shipments within

[13] Hummels and Skiba (2004) find strong support for the latter contention, and show that factory gate prices for a particular exporter and product positively co-vary with shipping costs across destinations.
[14] See Dornbusch, Fischer, and Samuelson (1977) for this result in a two-location world, and the generalization to the n-country case in Eaton and Kortum (2002).

a particular country[15], or whether this effect contributes in an important way to the overall effect of spatial frictions on trade.[16]

A third explanation turns on the idea that not all varieties are consumer goods. Suppose that region i produces an intermediate good that is useful only as an input into producing a particular final good. If region j does not produce that final good, it has no use for region i's output. To complete the story, the production "match" (the likelihood that region i produces inputs that region j's industry wants to consume) is stronger when the two regions are proximate. If production is footloose and demand for specific inputs is geographically concentrated, up- and downstream firms will co-locate (see Eaton and Kortum (2002) or Krugman and Venables (1996)).

III.A Decomposition methodology

To examine equation (3) empirically we decompose the aggregate value of trade flowing from region i to region j as follows. Indexing unique shipments by s, $T_{ij} = \left(\sum_{s=1}^{N_{ij}} P_{ij}^{s} Q_{ij}^{s}\right)$ is the total value of shipments from region i to region j, N_{ij} is the number of unique shipments (the extensive margin), and $\overline{PQ}_{ij}$ is the average value per shipment (the intensive margin), or

$$(4) \qquad T_{ij} = N_{ij} * \overline{PQ}_{ij}$$

"Unique" in this case means that we have at most one observation per origin establishment x SCTG commodity code x 5-digit destination zip code triplet. If the same establishment ships the

[15] This explanation requires that the fixed costs be paid for each destination, which seems plausible in an international context, where fixed costs of trade are often described as costs associated with search and information or establishment of distribution networks. To apply this story in the context of our data requires that fixed costs be zip code specific and operating at a radius of 4 miles from the shipper's establishment.

[16] Evans (2001) shows that industries that export a smaller share of output exhibit larger border effects in US-Canadian trade. However, this is an indirect inference as Evans cannot discern in the trade data the presence or absence of shipments by individual firms, only aggregate values.

same product to the same 5-digit destination zip code more than once, we aggregate and count that as a single shipment. In this way we treat as similar 10 shipments of $100, or 1 shipment of $1000. In the data sample we employ, the modal and median number of shipments per unique establishment x commodity x destination is one.[17] "Region" is defined flexibly, using zip codes of origin and destination. In some cases we will use the finest geographic detail (5-digit zips), while in others we will use less detail (typically 3-digit zip codes).

There can be multiple unique shipments within an origin-destination region pair. We can further decompose the number of shipments into N_{ij}^{k}, the number of distinct SCTG commodities shipped, and N_{ij}^{F}, the average number of shipments between a unique origin establishment and a unique destination zip code per commodity.

$$(5) \qquad N_{ij} = N_{ij}^{k} * N_{ij}^{F}$$

When we are measuring regions at the 5-digit zip code, $N_{ij}^{F} > 1$ means that we observed more than one unique establishment per commodity in region i shipping to zip code j. If we are measuring regions at the 3-digit zip code level, $N_{ij}^{F} > 1$ could result from seeing more than one unique establishment per commodity and/or having multiple (5-digit) destination regions within the 3-digit region j.

Finally, we decompose the average value per shipment into average price and average quantity per shipment

$$(6) \qquad \overline{PQ}_{ij} = \frac{\left(\sum_{s=1}^{N_{ij}} P_{ij}^{s} Q_{ij}^{s}\right)}{N_{ij}} = \frac{\left(\sum_{s=1}^{N_{ij}} P_{ij}^{s} Q_{ij}^{s}\right)}{\left(\sum_{s=1}^{N_{ij}} Q_{ij}\right)} \frac{\left(\sum_{s=1}^{N_{ij}} Q_{ij}\right)}{N_{ij}} = \overline{P}_{ij} * \overline{Q}_{ij}$$

Our units are weight (pounds) for all commodities. By using this common unit we are able to

[17] The mean is approximately 2, and is driven by a small number of cases in which we observe a large number of repeated shipments.

aggregate over dissimilar products, and to compare prices (per pound) across all commodities.

We now have total trade between two regions, decomposed into four component parts. Taking logs we have for the first level decomposition

(7) $\ln T_{ij} = \ln N_{ij} + \ln \overline{PQ}_{ij}$

and for the second level decomposition,

(8) $\ln T_{ij} = \ln N_{ij}^{k} + \ln N_{ij}^{F} + \ln \overline{P_{ij}} + \ln \overline{Q_{ij}}$

III.B. Decomposition Results

We begin by analyzing how each component in equation (8) covaries with the distance between origin-destination pairs. We use a kernel regression estimator to provide a smoothed estimate of the relationship between (levels) of each element in (8) and distance shipped.[18]

Figure 1 shows a kernel regression of T_{ij} on distance. Value declines very rapidly with distance, dropping off almost an entire order of magnitude between 1 and 200 miles, and is nearly flat thereafter. This figure demonstrates that there is a significant advantage to observing trade on a very fine geographic grid. Research that employs country level trade data ignores the pronounced effect of distance on trade over very short distances. Even work that employs state level trade data and imputes intra-state distances has the potential to misstate very short distance effects.

Figure 2 shows kernel regressions on distance of the two components of total value captured in equation (7). The number of unique shipments drops very rapidly over distance, at roughly the same rate as total value. In contrast, average value per shipment shows no clear diminution over space. The plot rises then falls slightly over distance, with similar values at 50

[18] We use the Gaussian kernel estimator in STATA, calculated on n=100 points, and allowing the estimator to calculate and employ the optimal bandwidth.

and 3000 miles.

Finally, Figure 3 shows kernel regressions on distance of the four components of total value captured in equation (8). The number of shipments drops off over space due both to reductions in the number of commodities shipped and number of shipments per commodity. Note also that the distance profiles of these components are flatter than that of the total number of shipments. The graph showing average values per shipment masks interesting variation in its components. Price per pound rises, and average shipment weight over distance falls.

Next we use linear regression analysis to decompose the effect of spatial frictions on total shipment values. The advantage of this technique relative to the kernels is three-fold: we can control for other covariates; we can address the significance of variables that have appeared in the literature such as state borders; and by estimating how each component varies over space, we can precisely gauge its contribution to the overall decline in trade values over distance.

We take logs of total trade and each of its components given in equations (7) and (8), and use each in turn as the dependent variable. We regress each on a vector of spatial variables X after differencing all variables by origin region and destination region means. Referencing equation (2), differencing in this manner eliminates variation in output and shipment prices specific to origin regions, and expenditures and price levels that are specific to destination regions, leaving only bilateral variation in the variables.[19] The vector X includes logs of distance (in miles), the square of log distance, and dummy variables that take a value of one if the flow took place within the same zip code (ownzip) or state (ownstate). Including distance is standard in gravity equations. We employ the other terms to investigate nonlinear effects of shipments over very short distances, and also to see if political boundaries (state lines) that pose no

---

[19] Anderson and vanWincoop (2003) highlight the dangers of omitting prices and price indices as they are likely to be correlated with trade frictions. We follow Hummels (2001) in differencing the data to control for these variables.

apparent impediment to trade survive better measurement of shipment distances.

Because OLS is a linear operator, we can regress $\ln T_{ij}$ on a variable in X, and its components on the same variable in X, and the resulting coefficients will have a useful additive property. For example, using the decomposition in (7), regressions of $\ln T_{ij}$, $\ln N_{ij}$, and $\ln \overline{PQ}_{ij}$ on distance will yield coefficients $\beta_T = \beta_N + \beta_{pq}$, where the subscripts refer to the dependent variable. If doubling distance cuts shipments in half, some part of this effect may occur through reducing the number of shipments, and some part through reducing the average value per shipment. We can assess the importance of each by looking at the contribution of each component to the total, e.g. if $\beta_N / \beta_T = .75$ then 75% of the total impact of distance on the value of shipments comes through a reduction in the number of shipments.

Table 1 reports results of regressions of $\ln T_{ij}$ and its components on spatial variables, with regions measured at the 3-digit zip code level. All our spatial variables are significantly correlated with total value of shipments. Shipments are steeply declining in distance, but this effect substantially flattens the further out shipments travel. In the final column of Table 1 we evaluate the distance elasticity, including the quadratic term, at the sample mean distance of 523 miles. For aggregate trade, the distance elasticity is a third smaller at the sample mean than at one mile. In addition, intra-state and intra-zip shipments are larger than those outside. Intra-state shipments are exp(0.55)=1.73 times larger than interstate shipments, a magnitude that is similar to Hillberry and Hummels (2003) estimates using states as the geographic unit of measure.[20] In addition, shipments within the same 3-digit code are exp(1.05)=2.86 times larger than shipments outside the zip code. Putting these effects together, shipments inside the same 3

[20] Wolf (2000) finds an intra-state effect three times larger, but his data include wholesaling shipments. As in the preferred specifications of Hillberry and Hummels (2003), Table 2, we use only manufacturing data, but our geographic data are 3-digit zip codes rather than states.

digit zip are 4.95 times larger than those outside the state, after controlling for distance.

The subsequent rows demonstrate which components of aggregate trade are responsible for the spatial effects. Looking at the decomposition in (7), 96 percent of the aggregate, first-order distance effect and 88 percent of the aggregate ownstate effect comes from the number of shipments, while the aggregate ownzip effect is smaller than the effect operating through number of shipments. In other words, average value per shipment is actually rising over short distances, given the negative ownzip coefficient, while the number of shipments drops off very rapidly.[21]

At larger distances, both the aggregate distance elasticity, and the fractional contribution of the extensive margin to it, are falling. For example, at the sample mean of 523 miles, the extensive margin represents 62 percent of the aggregate distance elasticity.[22] It should be noted that at this point nearly all of the effect of distance on shipment numbers has already been felt. This can be seen most simply by inspecting the kernel regressions in Figures 1 and 2. Starting from N > 200 for the most proximate pairs, there are only a handful of unique shipments still being observed at distances beyond a few hundred miles. While the average value of these shipments does decline at a similar rate as the number of shipments from this point, their contribution to the overall trade-distance profile is quite minor.

The substantial contribution of the extensive margin to falling trade values (especially over short distances) is the first major result of this paper. It is important to emphasize that the predominance of the extensive margin is not caused by over-sampling short shipments. There is no correlation between Census' sampling weights and the distance shipped. And we have aggregated sample records in which there are multiple small shipments by the same establishment, in the same SCTG commodity, to the same destination.

---

[21] The elasticity of average value per shipment with respect to distance is -0.345 at the sample mean distance of 523. At 100 and 1000 miles, the estimated elasticities are -0.269 and -0.375, respectively.

[22] The extensive margin accounts for 74 and 57 percent of the trade elasticity at 100 and 1000 miles.

Turning to the second-level decomposition of equation (8), we see a pattern similar to that in the kernel regressions. The decline in number of shipments over space is coming equally from two places. A larger number of commodities are shipped between proximate geographic pairs, and proximate geographic pairs see a larger number of unique (establishment x destination zip code) shipments per commodity

Also similar to the kernels, we see that while average value per shipment does not change much over distance, its components do. Increases in shipment distance correspond to increases in average price per pound and decreases in average pound shipped. Examining the quadratic terms we see that both effects grow stronger at greater distances. However, the two effects are highly significant and opposite – mostly canceling one another out in the average value of shipments measure.

Viewed in isolation, the sharp declines in average shipment weights over larger distances might appear to justify traditional theoretic and empirical treatments that assume "iceberg" trade costs operating on q_{ij} as in equation (1). However these models also predict that f.o.b. prices will be constant over space. Our evidence of offsetting price rises is more consistent with the work of Hummels and Skiba (2004), who show that if delivered prices are additive in trade frictions, $p^* = p + t$, and transportation costs are increasing in product weight, average prices will rise with distance while average quantities fall, the pattern we see in our data. The additive trade frictions model is broadly consistent with a pattern of changes in the composition of the traded bundle, i.e. goods with low value to weight ratios like cement travel shorter distances than goods with high value to weight ratios like electronics. In later robustness checks, we address whether this compositional change occurs within- or between-industries.

It is also useful to project the components of equation (8) on spatial variables when employing the full 5-digit zip code detail provided in the data. The estimates are done in the same manner, except that the variables are now differenced relative to their 5-digit origin and destination zip code means, and ownzip now refers to shipments within the same 5-digit zip code. The median radius of a 5-digit zip code is 4 miles.

Table 2 reports results for the 5-digit zip code data that are markedly different from the 3-digit data in Table 1. Concentrating on the total value regression, the regression fit is much worse in all regressions, with an R^2=.01. The elasticity of trade value with respect to distance is -0.13, 1/10th that of the estimates from Table 1, and the coefficient on the own state variable is actually negative but very close to zero. The only variable that comes through in a similar fashion to Table 1 is ownzip; shipments are exp(1.1)=3 times higher within a 5-digit zip code (median radius = 4 miles) than outside.

Turning to the decomposition, we see more differences relative to the 3-digit data. While average shipment values are high for very short distances (positive ownzip), they are thereafter increasing in distance, and average values are higher outside states than within. Shipment numbers are still declining in distance, but the elasticities are about 1/5th as large as before.

We see two reasons for the large difference between the 3- and 5-digit data. The first comes from employing a more precise distance measure, which shows up in the estimates on the ownstate and ownzip variables. Earlier studies using CFS data at the state level (Wolf 2000, Hillberry and Hummels 2003) find strong evidence of state-level home bias. We find that, after controlling for distance and ownzip effects, there is no evidence of state home bias in the 5-digit zip code data. In fact, the coefficient is negative. Related, the strength of the ownzip effect is the same whether the zip code in question has a median radius of 4 miles (5-digit zips) or 30 miles (3-digit zips). This suggests to us that shipments are extremely localized, dropping off rapidly over even very short distances. When measuring spatial frictions on broader geographic aggregates, even aggregates as small as 3-digit zip codes, these sharp distance declines are captured by "home bias" dummy variables. This seems likely to confound any effort to attribute "home bias" to plausible and measurable border barriers. The absence of state-level home bias over fine levels of geographic aggregation is the second major result of this paper.

The other difference between the 3- and 5-digit data is the much smaller effect of distance on trade, and it is closely related to the dominant role played by number of shipments. In both cases, average shipment values are mostly flat over distance, and the numbers of shipments are driving the spatial composition of trade. When evaluated at the 5-digit zip code level there are over 800 million origin-destination pairs. Our data show shipments between less than 1 percent of those pairs, and the remaining pairs are dropped from the regression. When we aggregate geographically we are adding together many 5-digit pairs contained within the 3-digit pairs. As a result roughly half of the 3-digit zip code origin-destination pairs contain shipments.

The lesson we take from this extends to broader geographic aggregation, to the level of states, regions, and perhaps countries. Suppose that there is a low probability of a shipment being observed between any 5-digit pair, with that probability dropping with distance. As we aggregate geographically, an essentially binary dependent variable (presence or absence of shipments) sums into a continuous variable (number of shipments) that co-varies strongly with distance. Much of that co-variation is lost through censoring of zero values at the 5-digit level. But when we aggregate from 5-digit to 3-digit, we see more commodities are shipped, and more establishments shipping to more destinations. We return to this probabilistic view of trade flows in Section IV.

III.C. Robustness: Commodity Level Decompositions

There may be pronounced differences across commodities in their response to geographic frictions. We repeat the decomposition in equation (7) for each of 20 2-digit SIC industries[23] using regions defined again at the 3-digit zip code level. The value of shipments in these regressions is the value of 2-digit industry shipments, and the decomposition follows (1.4)-(1.8). By mean-differencing the data we now remove variation in output, prices, and price levels that are specific to origin-industry and destination-industry.

In the interests of brevity we omit a full table of regression output, but the important results from Table 1 go through at the commodity level. Most of the covariation between total values shipped and spatial frictions is driven by the number of shipments. The regression fit is very low for all dependent variables except those with the number of shipments as the dependent variable. Focusing on these regressions, we find that the coefficients match the sign pattern from Table 1 for all explanatory variables in all commodities. In 16 of the 20 industry level regressions the average value per shipment is either insignificantly related to distance, or is rising in distance.

The one result from the aggregate regressions that does not hold up in the industry level regressions is the spatial pattern of prices and quantities. In the aggregate we found that quantities (weight) were falling over distance, while prices (per pound) were sharply rising. In the commodity regressions, prices are either constant or falling over distance in 17 of 20 regressions, while weight is either constant or rising in distance for 17 of the 20. These results suggest that the effect of distance on average price and average weight from the aggregate estimates is driven by broad cross-industry, rather than within-industry, substitution in the traded bundle.

As a final robustness check, we examined whether industry characteristics could explain

[23] We organize the data by the SIC of the shipping establishment rather than the SCTG code of the commodity shipped in order to use SIC industry characteristics. In these SIC industry decompositions, we continue to investigate the number of SCTG commodities shipped per establishment x destination pair.

either the strength of spatial variables (i.e. the coefficients on distance) or the relative contribution of each component of trade to the aggregate value. We constructed elements of equation (8) for each ij pair for each 4-digit SIC industry. We then pooled over all industries so that an ij pair has as many observations as the number of 4-digit industries traded between them. We regressed each element of (8) on spatial variables as well as an interaction between spatial variables and industry characteristics including measures of scale (average number of employees per establishment), tradability (the share of the transportation sector in industry gross output), and intermediate / final goods status (the share of final consumption in industry sales).[24] While some interactions were statistically significant, they did not result in economically significant changes in coefficients.

What have we learned from these regressions? There are four main messages which appear robust to our choices of geographic and commodity aggregation. First, spatial frictions matter over even very short distances. At the most extreme, shipments destined within the same 5-digit zip code, i.e. within a 4 mile radius of the shipper, are 3 times higher than those outside the zip code. Second, spatial frictions primarily reduce trade by reducing the number of shipments. This effect is especially pronounced over short distances. Third, but related, average shipment values are flat over short distances, and are only moderately sensitive to spatial frictions at larger distances. Fourth, state-level home bias observed in more coarse geographic data disappears when own 5-digit zip code dummies are included, and distance is measured precisely.

[24] These regressions are similar to those conducted by Chen (2004), who interacts commodity characteristics with the border dummy in European data. Our left hand side variables also include components of trade value, and our data offer both finer geography and greater industry detail.

IV. Explanations

We noted above that most empirical work on spatial frictions is motivated by theoretical models in which frictions affect only quantity per variety traded, and not the number of varieties. However, the evidence in Section III suggests that for intra-national flows spatial frictions work almost entirely through reducing the number of unique shipments, with little effect on the average value traded per unique shipment.

The previous literature suggests a few hypotheses that could explain this fact. Our primary hypothesis is that goods produced at a distance are not purchased because there is no local industrial demand for them. That is, region i produces an intermediate good that is not employed by industry in importing region j. Models of co-location, such as Krugman and Venables (1995) and Eaton and Kortum (2002), suggest that this is more likely to occur when the two regions are distant. Our alternative hypothesis is that consumers substitute away from distant goods because spatial frictions raise their price relative to local goods. Distance costs could choke off trade entirely if consumers regard local and distant goods as homogeneous, or if consumers regard them as differentiated but must pay a fixed cost to import.

To sharpen the distinction, under the alternative hypothesis, consumers would buy a good produced at a distance if the price became competitive with local alternatives. Under our primary hypothesis of industrial inputs, there might be no demand for the good at any price. That is, paper mills locate near forests, and auto assemblers near auto part plants. But an auto assembler's demand for uncut logs would not rise even if the assembler located near forests.

One approach to separating these hypotheses might be to study the industrial classification system, categorize goods as intermediate or final, and examine whether they respond differently to frictions. The problem is that, at the level we see the data, all industries contain both intermediate and final goods. And undoubtedly, elements of both hypotheses can be found in the data. Accordingly, we do not think of the exercises that follow as sharp tests of any model. Rather they are intended to shed some light about some of the likely sources of the Section III decomposition results.

We provide three exercises. First, we examine variation across regions in industry absorption to see if absorption is related to industrial structure. Second, we relate variation across regions in import shares to regional variation in output and export shares. Third, the decomposition results suggest that spatial frictions reduce trade primarily by eliminating shipments entirely. We use probits to assess whether industrial demands affect the likelihood that a good will be shipped between two regions.

IV.A. Spatial Variation in Absorption and Imports

The Commodity Flow Survey data provides a record of commodity shipments into and out of zip codes within the US. By summing over all incoming shipments, we are uniquely able to construct absorption as the sum of expenditures on each product in each region. The simple hypothesis is that regional industrial structure should help explain the pattern of bilateral shipments. Regions that are large producers of autos should be more likely destinations for shipments of auto parts.

Define total *actual* absorption of a commodity k in region j using the sum of all CFS shipments (T) from all source regions i, to absorbing region j (including i=j) in that commodity.

(9) $$E_j^k = \sum_i T_{ij}^k$$

Commodity k is defined as a 4 digit SIC category. We also define $e_j^k = E_j^k / \sum_k E_j^k$ as the share of k in j's total absorption. There is a tremendous amount of variation across regions in these absorption shares. To show this we calculate, for each 4 digit SIC industry, the mean and standard deviation of e_j^k over all 3-digit zip codes. The *least* variable SIC industry has sd/mean = 1.19, that is, the region that is one standard deviation above the mean has an e_j^k 119% larger than the mean. The median SIC industry has sd/mean = 4.11.

We seek to explain observed absorption, e_j^k, with predicted absorption $\tilde{e}_j^k$, a variable we construct from the 1997 U.S. Benchmark Input-output tables, and selected measures of regional output and consumer expenditure. Data availability differs depending on whether we define region at the state or 3-digit zip code level, and our measure of predicted absorption differs accordingly. When constructing predicted absorption at the 3-digit zip code region, we employ manufacturing value added for each zip code x 4-digit SIC industry taken from the non-public Census of Manufactures data. When constructing predicted absorption at the state level, we use personal income and Gross State Product by industry taken from the BEA. Gross State Product data are somewhat more aggregated than the CoM data at the industry level, but provide us with value added for non-manufacturing activities in addition to manufacturing.

We construct predicted industrial demands by first calculating the ratio U^{kl}/Y^l from the input-output table, where U^{kl} is the value of industry k shipments in industry *l* gross output, and Y^l is industry *l* value added.[25] We multiply this ratio by industry *l* value added in a region, Y_j^l. Summing over all industries *l* in region j gives a measure of predicted industrial demand:

(10) $$\tilde{E}_j^k = \sum_l Y_j^l \frac{U^{kl}}{Y^l}$$

For 3-digit zip codes the summation runs only over manufacturing industries, and provides our

[25] We first concord Input-Output industries to the 4-digit SIC level. The concordance is complicated by a one-to-many problem between I-O sectors and SIC industries. We address this problem by allocating 1/n of I-O sector use to each of n SIC sectors in those lines where there is a one to many problem.

complete measure of predicted absorption. For states this summation runs over all manufacturing and services industries. To this industrial demand we add personal consumption expenditures, constructed by multiplying final consumption expenditure shares for each industry taken from the IO table by each state's personal income. We then express the level of j's predicted absorption of k in shares, as above.

We regress the share of industry k in region j's absorption (taken from the Commodity Flow Survey), on imputed share of industrial demands for k in j (constructed from data on output, incomes, and input-output tables).

(11) $e_j^k = \alpha^k + \beta \tilde{e}_j^k + \varepsilon_j^k$.

We conduct the regression in levels, pool over all commodities[26] and include a commodity fixed effect.[27]

An important feature of our data is that the right and left hand sides of the estimating equation come from different datasets. The more common approach to measuring absorption is to calculate output less net exports, which means that the level of output enters both e_j^k and $\tilde{e}_j^k$. In our data the two are not correlated by construction.

Results are reported in Table 3. The first two columns define a region as a state. The first column includes all expenditure, and the second column excludes personal consumption expenditures from imputed absorption, leaving only intermediate demands. The third column defines a region as a 3-digit zip code, and calculates imputed demands arising only from manufacturing output.

Our estimates suggest that imputed industrial demands do help predict regional

---

[26] We also estimated equation (11) separately for each 4 dig SIC sector. Results are qualitatively similar, though in general the coefficients are larger and the regression fit is somewhat worse.

[27] The fixed effect soaks up some of the variation due to concordance problems, and serves as a control for relative industry size. Regressions that omit industry fixed effects yield similar coefficients, but a lower regression fit.

expenditure patterns. The coefficient on imputed industrial demands is positive, and precisely estimated, but less than one in every case.[28] Despite data limitations, we find that the production structure of US regions, and their resulting industrial demands, are strongly related to idiosyncrasies in the absorption patterns of those regions. This result is in marked contrast to the literature on international shipments. Harrigan (1995) finds no relationship between the commodity structure of a country's imports and the output mix of the country.

IV.B. Regional import shares: coals to Newcastle?

Our alternative hypothesis, that consumers substitute away from "foreign" goods because distance costs raise their price relative to local goods, contains a useful implication. Holding constant demand, consumers are more likely to import distant goods if local goods are unavailable. Put another way, imports should be negatively related to local output of the good.[29]

In contrast, industry imports may be positively correlated with local output in that industry if spatial variation in absorption is driven by industrial demands. To explain, simple inspection of input-output relationships shows that a substantial portion of industrial expenditure on intermediate inputs goes toward the purchase of own-industry output. For example, SIC industry 3711 "Motor Vehicles and Passenger Car Bodies" includes both finished cars and unfinished car parts like bodies and chassis. As a consequence, a region may have both high output (cars) and high imports (chassis) in a particular SIC industry.

To examine this, we construct the total value of industry k imports into region j as $M_j^k = \sum_{i \neq j} T_{ij}^k = E_j^k - T_{jj}^k$, or industry absorption less shipments from j to itself. We then

---

[28] While we would expect a coefficient of one, mismeasurement of predicted demands caused by concordance problems in the IO table could cause attenuation bias toward zero.

[29] In a homogeneous goods model, fix the demand curve and the foreign supply price, and shift the supply curve in. This results in a fall in the quantity supplied locally and a rise in imports. In a model with CES differentiated goods, reducing the number of varieties of local goods will raise the price index, making foreign goods more attractive at a given price, increasing imports.

express this as a share of j's total imports, $m_j^k = M_j^k / \sum_k M_j^k$ and regress it on industry k's share of gross output in region j, $g_j^k = GO_j^k / \sum_k GO_j^k$, as reported in the Census of Manufactures.

$$m_j^k = \alpha^k + \beta g_j^k + u_j^k \tag{12}$$

We pool the regression over all 4-digit SIC industries, including a fixed effect for each industry to capture relative industry size. Regions are defined flexibly as 2, 3, and 5 digit zip codes.

Our results are reported in table 4. At all three levels of geographic aggregation, we find that regions that are large producers of a given industry tend to be large importers of that industry's products. The effect is stronger when importing regions are defined over larger areas. The estimated coefficient on 2-digit shipments is 0.132, and on 5-digit zip codes is 0.074.[30] These positive coefficients indicate a pervasive "coals to Newcastle" effect in U.S. manufacturing shipments.

Positive estimates of β in (1.9) could be consistent with a model in which region j has, for some unknown reason, idiosyncratically large consumer demand for some constant returns to scale industry k. In this case, satisfying local consumption might require high levels of both local output and imports. To address this, we repeat the estimates, replacing gross output shares in region j with export shares out of region j, taken from the CFS. The results, also reported in Table 4, reveal very similar results. Regions that export idiosyncratically large shares of a given industry also import idiosyncratically large shares of that industry.[31] While the strength of the coals-to-Newcastle effect is smaller when exports are the independent variable, the coefficients remain positive and statistically significant.

---

[30] The regressions include many cases where either the production or import share of industry k was zero. This is less likely at the two digit zip code level, and probably is responsible for the larger coefficient in the more aggregated regression.

[31] Recall that we have excluded wholesale shipments from the data, so we are not observing wholesalers' re-exports of industry k output.

The coals-to-Newcastle effect we find is related to, and could be caused by, the home market effect proposed by Krugman(1980), and identified empirically in work by Davis and Weinstein (2003), Hanson and Xiang (2004) and Head and Ries (2001). The home market effect works as follows. If region j has idiosyncratically large demand for industry k whose production is subject to increasing returns to scale, output of k in j may rise by more than the demand shock. In principle this could result in region j simultaneously having idiosyncratically high absorption, imports, and exports in industry k. However, the unanswered question in the home market effect literature is where, precisely, these idiosyncratic demand shocks come from. One natural explanation is that intermediate demands are a key source of the demand variation observed in empirical studies of the home market effect.[32]

IV.C. Probits.

Recalling Table 1, the aggregate trade-distance relationship in intra-US trade is largely driven by the fact that most establishments ship only to geographically proximate customers, rather than shipping to many customers in values that decrease in distance. Conditional on a shipment taking place, its values are largely unaffected by spatial frictions. In this section we turn to a probit analysis of the likelihood that trade in a given origin-destination pair is observed. The most general specification is as follows:

$$\text{(13)} \quad \Pr(I_{ij}^k = 1) = \Phi\begin{pmatrix} \beta_0^k + \beta_1^k \ln Dist_{ij} + \beta_2^k \left(\ln Dist_{ij}\right)^2 + \beta_3^k Ownzip + \beta_4^k Ownstate \\ + \beta_5^k \ln GO_i^k + \beta_6^k \ln Pop_j^k + \beta_7^k \ln \tilde{E}_j^k + \beta_8^k \ln GO_j^k + \varepsilon_{ij}^k \end{pmatrix},$$

where $I_{ij}^k = 1$ if industry k ships from region i to region j.

The first four independent variables – distance, distance squared, ownzip and ownstate –

[32] Presumably, this effect would go away if trade data, and input-output linkages, were sufficiently granular that we could separate specific final goods from the specific intermediate inputs they employ.

are the spatial frictions that appear in our decomposition regressions from Section III. We also include a "supply" variable, GO_i^k, the value of industry k gross output in origin region i, as measured by the Census of Manufactures, and control for the size of demand by including the population of the destination region (POP_j^k).

Our primary interest lies in variables that might help explain the composition of demand. We include predicted industrial demand ($\tilde{E}_j^k$) as constructed in Section IV.A to see if destination region output mix, and the implied demand for inputs, helps predict the presence of shipments. Following the "Coals to Newcastle" results from the section IV.B., we also experiment with including the gross output of industry k in *destination* region j (GO_j^k). If distance operates through consumer substitution toward local varieties, greater output of commodity k in destination region j should lower the likelihood that k is shipped to j. A positive coefficient on ln(GO_j^k) would be consistent with the coals to Newcastle effect operating, in part, through the extensive margin.

We conduct our analysis at the 3-digit zip code level of geography, estimating probit models pooled over all observations within a two-digit SIC category.[33] Coefficients on spatial frictions match those from Table 1 for all commodities. We report estimated coefficients on $\ln(GO_i^k)$, $\ln(Pop_j)$, $\ln(\tilde{E}_j^k)$ and $\ln(GO_j^k)$ in Table 5.[34] As expected we find that increases in origin region i output and destination region j population raise the propensity of two regions to trade. In 13 of 20 industries, the coefficient on our constructed industrial demands is positive and statistically significant, as expected. The coefficient estimate on destination region output is

[33] It would be interesting to perform this estimate at the 5-digit zip code level, or for more disaggregated SIC data, but this would have exceeded memory constraints.
[34] We also estimated (4.3) excluding destination region output. This had no impact on the spatial frictions, but generally raised the estimated coefficient on industrial demands, making it positive and statistically significant in two more cases.

positive and significant in 18 of 20 sectors. Similar to the evidence on import shares, having more industry k output in the destination region increases, rather than decreases the likelihood of shipments. These results provide further support for our hypothesis that trade in intermediate inputs is responsible for the large extensive margin.

V. Conclusions and Implications

We employ a unique data set, the Commodity Flow Survey, which allows us to track shipments on a very fine geographic grid within the United States. We use these data to shed light on why spatial frictions matter for trade by isolating which components of trade they most affect. Our decomposition provides several new findings.

First, spatial frictions matter, and indeed have the greatest impact, over very short distances. Shipments destined within the same 5-digit zip code, i.e. within a 4 mile radius of the shipper, are 3 times higher than those outside the zip code. Trade values fall an order of magnitude between 1 and 200 miles, and are relatively flat thereafter.

The highly nonlinear effect of distance on trade may help explain some results in the "home bias" literature. While our estimates on 3-digit zip code data reveal that intra-state shipments are significantly higher than cross-state shipments, this effect disappears entirely when shipment distances are measured more accurately using 5-digit zip codes. We instead find that the "borders" between 5-digit zip codes represent a sizeable barrier to trade. We consider these zip-code effects the *reductio ad absurdum* of the home bias literature. While one can imagine many barriers to trade that operate at national borders, it is harder to conceive of what barriers plausibly operate at state borders, and harder still to imagine those associated with 5-digit zip codes. Our results suggest that "home bias", at least in state borders, is an artifact of geographic aggregation. Since shipments drop off extraordinarily rapidly over very short distances, attempts to measure border effects on larger geographical groupings are nearly certain to ascribe the nonlinear effects of distance to "home bias" dummy variables.

Second, and in contrast with the theoretical models typically used to underpin gravity models of trade, we show that spatial frictions reduce trade primarily by reducing the number of shipments. That is, the aggregate trade-distance relationship in intra-US trade is driven by the fact that most establishments ship only to geographically proximate customers, rather than shipping to many customers in quantities that decrease in distance. Conditional on a shipment taking place, its value is largely unaffected by spatial frictions.

To explain these results, we examine two hypotheses about why spatial frictions matter. The first hypothesis, consistent with the modeling framework in most of the literature on spatial frictions, focuses on final demand for consumer goods. It asserts that demand is uniform but goods are not imported because local substitutes are available at lower prices. The second hypothesis focuses on intermediate inputs. It asserts that demand is not uniform, and goods are not imported because there is no local industrial demand for them.

We find three pieces of evidence to support the intermediate inputs hypothesis. We show that industry level absorption varies considerably across regions, and is ably predicted by industrial structure and demand for intermediate inputs. Similarly, the likelihood of a particular shipment occurring is closely related to industrial structure: goods are more likely to be imported into a region when industrial demand for that good is high Finally, and in contrast to models where local consumer goods substitute for expensive "foreign" alternatives, regions import intensively those industries in which they have both high output, and high exports. That is, we observe intra-industry trade even at the level of 5-digit zip codes. This seems most plausibly explained by the intra-industry exchange of intermediate inputs for assembled outputs.

These results complement work showing that relatively few firms ship internationally

(Bernard and Jensen, 1999; Bernard et al 2003), and that conditional on exporting, firms ship to relatively few destinations (Eaton, Kortum, Kramarz 2004). Our data show that shipping to few, local, markets also characterizes shipments within a country. These authors emphasize cost advantages of firms, showing that the most productive firms export, and export to more destinations, while we emphasize specialized industrial demands. We leave to future work whether there might be gains from trade between these explanations. That is, can differences in productivity help explain which firms ship to more markets within a country, and can specialization in input demands shed light on why so few firms export internationally?

Bibliography

Amiti, Mary and Lisa Cameron 2003, "Economic Geography and Wages," University of Melbourne, *mimeo*.

Anderson, James and Eric van Wincoop 2003, "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, pp. 170-92.

Bernard, Andrew, Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum, 2003, "Plants and Productivity in International Trade", *American Economic Review*, 93, pp. 1269-1290.

Bernard, Andrew and Bradford Jensen. 1999, "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics*, 47, pp 1-25.

Chen, Natalie 2004, "Intra-national versus international trade in the European Union: why do national borders matter?" *Journal of International Economics,* 63, pp. 93-118.

Davis, Donald R. and David E. Weinstein 2003, "Market Access, Economic Geography, and Comparative Advantage: An Empirical Test," *Journal of International Economics,* 59, pp. 1-23.

Disdier, Anne-Célia and Keith Head 2004, "The Puzzling Persistence of the Distance Effect on Bilateral Trade," University of British Columbia, *mimeo*.

Dornbusch, Rudiger, Stanley Fisher, and Paul Samuelson 1977, "Comparative Advantage, Trade and Payments in a Ricardian Model with a Continuum of Goods," *American Economic Review*, 67, pp. 823-839.

Eaton, Jonathan and Samuel Kortum, 2002, "Technology, Geography, and Trade," *Econometrica,* 70, pp. 1741-1779.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz, 2004, "An Anatomy of International Trade: Evidence from French Firms" mimeo. NYU.

Evans, Carolyn L. 2001, "Border Effects and the Availability of Domestic Products Abroad" Federal Reserve Bank of New York, *mimeo.*

Hanson, Gordon H. and Chong Xiang 2004, "The Home Market Effect and Bilateral Trade Patterns," *American Economic Review*, 94, pp. 1108-1129.

Harrigan, James 1995, "The Volume of Trade in Differentiated Intermediate Goods: Theory and Evidence," *Review of Economics and Statistics,* 77(2), (May) pp. 283-293.

Head, Keith and John Ries 2001, "Increasing Returns Versus Product Differentiation as an Explanation for the Pattern of US-Canada Trade, *American Economic Review*, 91, pp. 858-876.

Hillberry, Russell and David Hummels 2003, "Intra-national Home Bias: Some Explanations," *The Review of Economics and Statistics*, 85(4), pp. 1089-1092.

Hillberry, Russell and Christine McDaniel 2002, "A Decomposition of North American Trade Growth since NAFTA," USITC Working Paper 2002-12-A (July).

Hummels, David 2001, "Toward a Geography of Trade Costs," Purdue University, *mimeo*.

Hummels, David and Peter Klenow 2004, "The Variety and Quality of A Nation's Exports," *American Economic Review,* forthcoming.

Hummels, David and Alexandre Skiba 2004, "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian Allen Conjecture" *Journal of Political Economy*, 112(6), pp. 1384-1402.

Kehoe, Timothy J and Kim J Ruhl 2003, "How Important is the New Goods Margin in International Trade?" Federal Reserve Bank of Minneapolis Staff Report 324.

Klenow, Peter and Andres Rodriguez-Clare 1997, "Quantifying Variety Gains from Trade Liberalization," Stanford University, *mimeo*.

Krugman, Paul 1980, "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review,* 70, pp. 950-959.

Krugman, Paul and Anthony Venables 1996, "Integration, Specialization, and Adjustment," *European Economic Review,* 40, pp. 959-967.

McCallum, John 1995, "National Borders Matter: Canada-U.S. Regional Trade Patterns," *American Economic Review,* 85, pp. 615-623.

Midelfart-Knarvik, Karen-Helene, Henry Overman and Anthony Venables 2001, "Comparative advantage and Economic Geography; Estimating the Location of Production in the EU," CEPR Discussion Paper 2618.

Romer, Paul 1994, "New Goods, Old Theory and the Welfare Costs of Trade Restrictions" *Journal of Development Economics,* 43, pp. 5-38.

Wolf, Holger C. 2000, "Intra-national Home Bias in Trade," *The Review of Economics and Statistics,* 82(4) (November), pp. 555-563.

Figure 1: Kernel Regressions

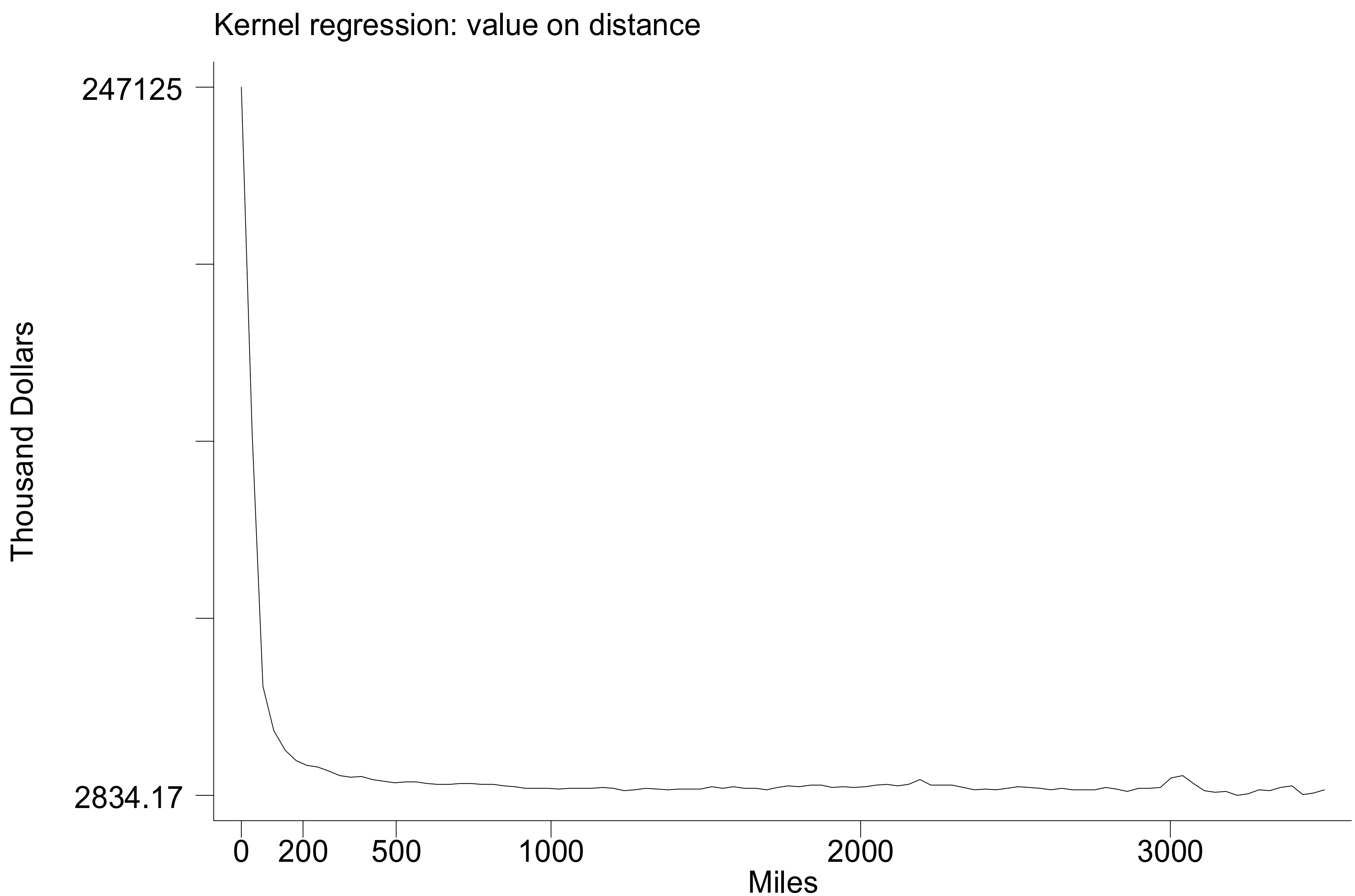

Figure 2: Kernel Regressions

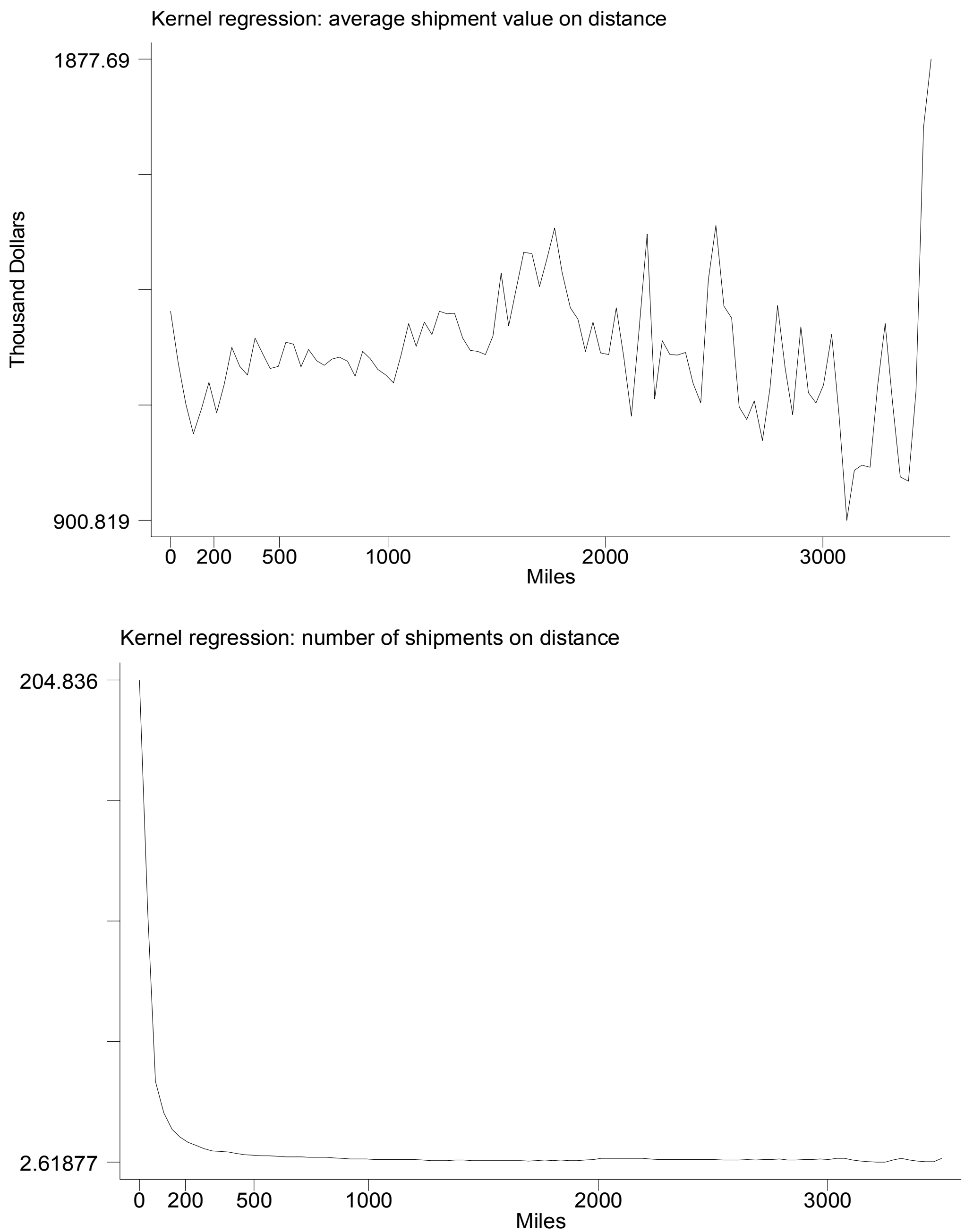

Figure 3: Kernel Regressions

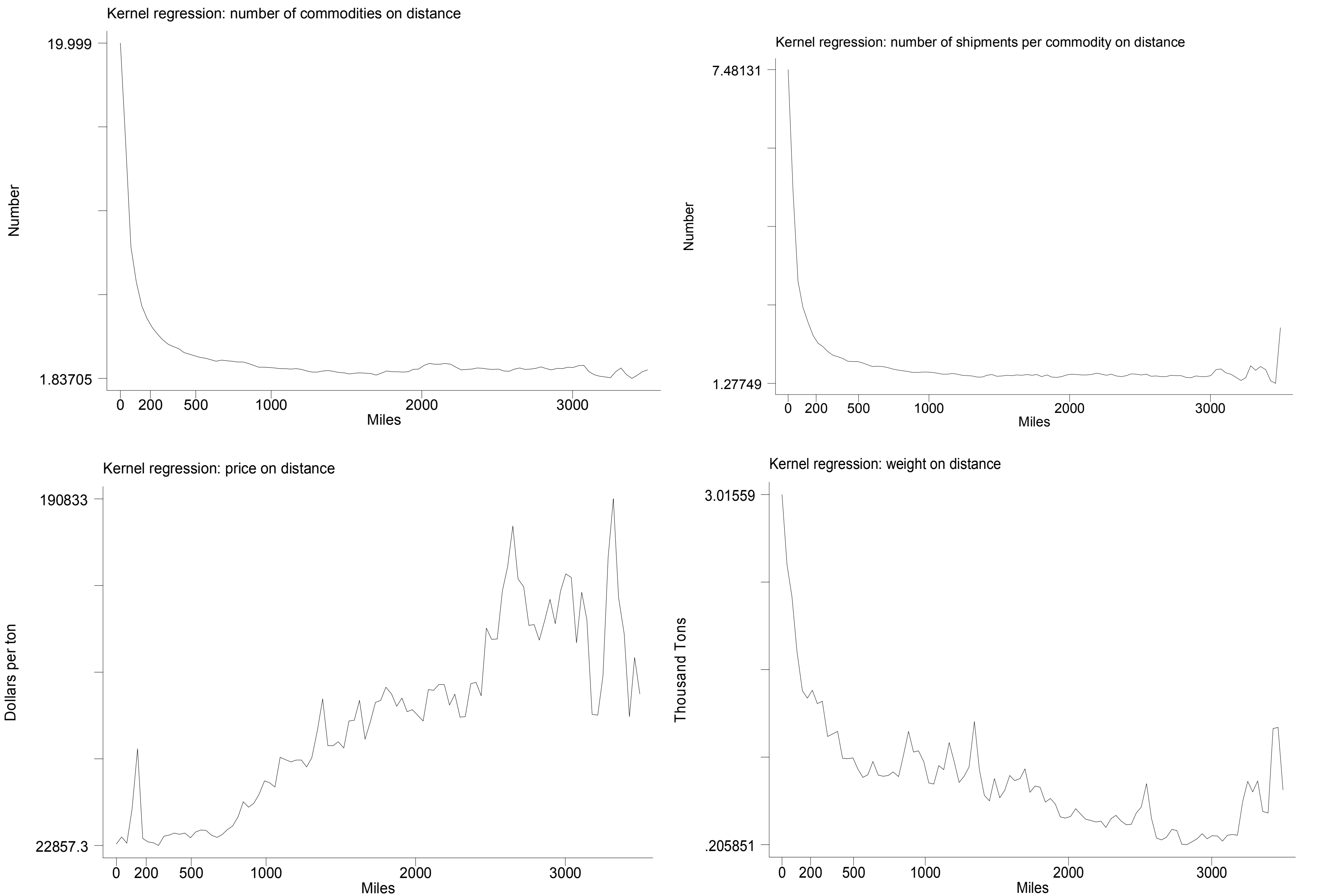

Table 1. Decomposing Spatial Frictions
(3-digit zip code data)

	dist	$dist^2$	ownzip	ownstate	constant	Adj. R^2	N	ε_D
value (T_{ij})	**-1.348** **(0.045)**	**0.035** **(0.004)**	**1.051** **(0.084)**	**0.549** **(0.020)**	**-20.992** **(0.141)**	0.13	333021	-0.910
# of shipments (N_{ij})	**-1.292** **(0.016)**	**0.058** **(0.001)**	**1.232** **(0.031)**	**0.481** **(0.007)**	**-7.021** **(0.051)**	0.36	333029	-0.566
# of trading pairs (N_{ij}^F)	-0.638 (0.010)	0.037 (0.001)	0.813 (0.018)	0.161 (0.004)	-2.907 (0.031)	0.16	333029	-0.175
# of commodities (N_{ij}^k)	-0.653 (0.011)	0.021 (0.001)	0.419 (0.022)	0.319 (0.005)	-4.114 (0.036)	0.32	333029	-0.391
avg. value ($\overline{PQ}_{ij}$)	**-0.057** **(0.038)**	**-0.023** **(0.003)**	**-0.182** **(0.071)**	**0.069** **(0.017)**	**-13.971** **(0.119)**	0.02	333021	-0.345
avg. price ($\overline{P}_{ij}$)	0.389 (0.035)	0.032 (0.003)	-0.232 (0.065)	-0.385 (0.016)	2.909 (0.110)	0.14	333021	0.790
avg. weight ($\overline{Q}_{ij}$)	-0.445 (0.054)	-0.055 (0.004)	0.051 (0.101)	0.454 (0.024)	-16.880 (0.169)	0.12	333021	-1.135

Notes:
1. Regression of (log) shipment value and its components from equations (7) and (8) on geographic variables. Dependent variables in left hand column. Coefficients in right-justified rows sum to coefficients in left justified rows.
2. Standard errors in parentheses.
3. ε_D is the elasticity of trade with respect to distance, evaluated at the sample mean distance of 523 miles.

Table 2. Decomposing Spatial Frictions
(5-digit zip code data)

	dist	$dist^2$	ownzip	ownstate	constant	Adj. R^2	N	ε_D
value (T_{ij})	**-0.137** **(0.009)**	**-0.004** **(0.001)**	**1.102** **(0.030)**	**-0.024** **(0.007)**	**-13.393** **(0.026)**	0.01	1290788	-0.187
# of shipments (N_{ij})	**-0.294** **(0.002)**	**0.017** **(0.000)**	**0.883** **(0.008)**	**0.043** **(0.002)**	**-1.413** **(0.007)**	0.10	1290840	-0.081
# of trading pairs (N_{ij}^F)	-0.159 (0.002)	0.008 (0.000)	0.540 (0.007)	0.029 (0.002)	-0.888 (0.006)	0.05	1290840	-0.059
# of commodities (N_{ij}^k)	-0.135 (0.001)	0.009 (0.000)	0.342 (0.003)	0.014 (0.001)	-0.525 (0.003)	0.10	1290840	-0.022
avg. value ($\overline{PQ}_{ij}$)	**0.157** **(0.008)**	**-0.021** **(0.001)**	**0.219** **(0.028)**	**-0.067** **(0.006)**	**-11.980** **(0.024)**	0.00	1290788	-0.106
avg. price ($\bar{P}_{ij}$)	-0.032 (0.007)	0.036 (0.001)	-0.115 (0.024)	-0.154 (0.006)	0.021 (0.020)	0.08	1290788	0.419
avg. weight ($\bar{Q}_{ij}$)	0.189 (0.011)	-0.058 (0.001)	0.334 (0.037)	0.087 (0.009)	-12.001 (0.031)	0.05	1290788	-0.537

Notes:
1. Regression of (log) shipment value and its components from equations (7) and (8) on geographic variables. Dependent variables in left hand column. Coefficients in right-justified rows sum to coefficients in left justified rows.
2. Standard errors in parentheses.
3. ε_D is the elasticity of trade with respect to distance, evaluated at the sample mean distance of 523 miles.

Table 3. Predicting Absorption with Industrial Demands.

	Region = state		Region = 3-digit zip
Imputed absorption includes	Manufacturing + services + personal consumption	Manufacturing + services	Manufacturing
$\tilde{e}_j^k$	0.684 (0.020)	0.412 (0.012)	0.313 (0.002)
Adj. R^2	0.47	0.46	0.142
N	16000	16000	395819

Notes:
1. Estimation of equation (11). Dependent variable is industry k share of region j absorption.
2. Industry k fixed effects included.
3. Standard errors in parentheses. All coefficients significant at the 1% level.

Table 4. Coals to Newcastle? Predicting regional import shares

	5-digit zip code		3-digit zip code		2-digit zip code	
gross output share	0.074		0.110		0.132	
	(0.0003)		(0.001)		(0.002)	
export share		0.045		0.069		0.119
		(0.0003)		(0.001)		(0.002)
adjusted R^2	0.026	0.022	0.124	0.111	0.398	0.393
N	5200440	5200440	390939	390939	44394	44394

Notes:

1. Estimation of equation (12). Dependent variable is industry k share of region j imports.
2. "Export" shipments are shipments that leave the zip code, but are bound for U.S. destinations.
3. SIC fixed-effects included
4. Standard errors in parentheses. All coefficients significant at the 1% level.

Table 5. Probit estimation

SIC	Origin gross output	Destination population	Imputed industrial demand	Destination own-sector gross output	Log-likelihood	N
20	0.314 (0.001)	0.282 (0.002)	0.009 (0.000)	0.024 (0.001)	-281011.82	4,790,782
21	0.083 (0.005)	0.233 (0.015)	0.005 (0.005)	0.020 (0.005)	-6554.09	37,057
22	0.227 (0.001)	0.294 (0.004)	0.005 (0.001)	0.016 (0.001)	-122903.58	1,431,521
23	0.274 (0.001)	0.283 (0.003)	0.000 (0.001)	0.008 (0.001)	-224152.68	3524023
24	0.412 (0.002)	0.208 (0.003)	0.016 (0.000)	0.021 (0.001)	-188309.69	3,257,916
25	0.346 (0.001)	0.359 (0.003)	-0.001 (0.000)	0.000 (0.001)	-199739.27	2,239,096
26	0.280 (0.002)	0.344 (0.003)	0.004 (0.001)	0.011 (0.001)	-165056.37	1,847,676
27	0.296 (0.001)	0.245 (0.003)	0.027 (0.001)	0.019 (0.001)	-256926.12	3,497,694
28	0.259 (0.001)	0.275 (0.003)	0.026 (0.001)	0.013 (0.001)	-260554.40	3,271,360
29	0.220 (0.002)	0.249 (0.006)	-0.026 (0.001)	0.015 (0.001)	-39150.69	576,079
30	0.331 (0.001)	0.301 (0.003)	0.012 (0.001)	0.020 (0.001)	-252998.37	2,642,778
31	0.250 (0.002)	0.237 (0.005)	0.000 (0.001)	0.018 (0.001)	-49465.01	625,710
32	0.315 (0.002)	0.304 (0.003)	-0.019 (0.001)	-0.002 (0.001)	-167122.59	3,027,507
33	0.252 (0.002)	0.302 (0.003)	0.010 (0.001)	0.014 (0.001)	-135426.31	1,594,647
34	0.326 (0.001)	0.293 (0.002)	-0.009 (0.000)	0.018 (0.001)	-387176.33	6,089,828
35	0.297 (0.001)	0.229 (0.002)	0.002 (0.000)	0.048 (0.001)	-461600.67	6,797,738
36	0.253 (0.001)	0.345 (0.002)	0.010 (0.000)	0.008 (0.001)	-276317.64	4,047,346
37	0.195 (0.001)	0.265 (0.003)	0.017 (0.000)	0.004 (0.001)	-145136.50	2,068,732
38	0.254 (0.001)	0.337 (0.003)	0.012 (0.001)	0.004 (0.001)	-203245.82	2,590,116
39	0.306 (0.001)	0.302 (0.003)	0.005 (0.000)	0.006 (0.001)	-236787.91	2,545,089

Notes: Estimation of equation (13). The specification also included dist, dist2, ownzip and ownstate as independent variables. Standard errors in parentheses.