Firm Heterogeneity and the Structure of U.S. Multinational Activity: An Empirical Analysis*

Stephen Ross Yeaple

University of Pennsylvania

March 7, 2005

Keywords: Multinational Enterprise, Firm Heterogeneity, Foreign Direct Investment, Production Location Decisions.

JEL Codes: F0, F1, L1, L2

Abstract: Adapting trade models to account for within-industry firm heterogeneity is one of the fast growing areas within International Trade theory. A key feature of these models is that the volume of trade and foreign direct investment (FDI) between countries is a function of both an *extensive* margin (number of firms) and an *intensive* margin (sales per firm). Using the firm-level data of the Bureau of Economic Analysis, we study the importance of these margins for our understanding of the behavior of individual U.S. multinationals and for our understanding of the aggregate cross-country structure of FDI. We use the model of Helpman, Melitz, and Yeaple (2004) to formulate and test several implications of models firm heterogeneity and international commerce. Our results reveal that the behavior of U.S. multinationals is closely consistent with the model, thereby highlighting the importance of this class of models to our understanding the geographic structure of production.

* The statistical analysis of firm level data on U.S. Multinational Corporations reported in this study was conducted at the International Investment Division, U.S. Bureau of Economic Analysis, under arrangements that maintained legal confidentiality requirements. Views expressed are those of the authors and do not necessarily reflect those of the Bureau of Economic Analysis. We thank Bill Zeile for his comments and Jonathan Eaton for his encouragement.

Multinational enterprises, those firms that own and control production activities in multiple countries, play a dominant role in the conduct of international commerce. The value of the foreign sales of the approximately 2,400 U.S. multinational enterprises (MNEs) surveyed by the Bureau of Economic Analysis (BEA) in its 1999 benchmark survey swamp the value of export sales from the United States. Moreover, these same firms accounted for 63 percent of U.S. exports and 37 percent of U.S. imports (Mataloni and Yorganson, 2002). The disproportionate role played by MNEs in the economy is further highlighted by the fact that virtually all privately funded R&D in the United States is conducted by MNEs.

In recent years, International Trade theorists have developed models to understand the concentration of international commerce in a small number of firms (see for instance, Melitz 2003, and Helpman, Melitz, and Yeaple, 2004). By allowing firms within an industry to differ in their productivity, these models afford the insight that firms "sort" into their mode of serving foreign markets. A key implication of this sorting is that trade and foreign direct investment (FDI) volumes between countries vary along both an *intensive* margin (sales per firm) and an *extensive* margin (the number of firms selling their product abroad). These two margins can be defined both *within-firm* (the number of countries in which a firm produces and average volume of sales per country) and *within-country* (the number of firms that produce locally and average sales per firm.)

In this paper, we conduct a thorough investigation of the relevance of the extensive margin to our understanding of FDI patterns. Using firm-level FDI data from the 1994 benchmark survey of BEA we first generate several stylized facts concerning the importance of the *extensive* margin both *within-firm* and *within-country*. In particular, we show that the most productive U.S. MNE (i) own affiliates in a larger set of countries and (ii) generate larger sales

revenue in those countries in which they operate. With respect to aggregate cross-country FDI volumes, we show that the variation explained by the *extensive* margin, a defining characteristic of models of firm heterogeneity, is almost as important as the variation explained by the *intensive* margin.

We then specify a simplified version of the model presented in Helpman, Melitz, and Yeaple (2004) that is consistent with the stylized facts. We use the model to motivate two econometric analyses of the structure of U.S. multinational operations. The first of the two analyses exploits variation in the cross-country investment behavior of individual U.S. multinationals. We show that the probability that a firm invests in a given foreign country is increasing in that firm's productivity, but the importance of firm productivity varies across countries in a manner consistent with the model. In particular, firm productivity plays a less important role in the investment decision of firms in those countries most attractive to multinational firms.

In the second of the two analyses we assess the role of firm heterogeneity in the aggregate structure of FDI across countries. Using a technique derived from the theoretical model, we decompose the volume of FDI across countries into an *extensive* margin (the number of firms investing in a location), a *composition* margin (the average productivity of the firms investing in that location), and a *scale* margin (the residual size-of-affiliate). Our results indicate that a larger number of U.S. multinationals enter large, developed countries and that the revenue per firm in these markets is disproportionately large even after accounting for differences in the composition of firms investing there. Further, we show that country characteristics that induce the entry of a larger number of U.S. MNEs are also associated with a change in the composition of investing firms toward those with below-average productivity. This latter result is broadly

consistent with the sorting of heterogeneous firms into mode of international commerce: an expansion in the number of firms investing in a location necessarily involves attracting progressively less productive firms. The two empirical analyses, one exploiting *within-firm* variation and the other exploiting *within-country* variation in the data, both provide strong support for the predictions of Helpman, Melitz, and Yeaple (2004).

Our results complement a rapidly growing literature that seeks to assess the extent to which firms sort into different modes of serving foreign markets on the basis of their productivity. Examples of recent research that explores the relative productivity of firms that engage in FDI relative to firms that engage in exporting are Girma et al (2003) and Head and Ries (2003).

Our analysis of FDI is similar in spirit to two recent papers that seek to understand the role of the firm heterogeneity in the structure of trade volumes. In particular, our analyses complement those of Eaton, Kortum, and Kramarz (2004), who explore the relevance of models of firm heterogeneity for understanding *export volumes* across countries. The model used in our analysis is similar to that used by Helpman, Melitz, and Rubinstein (2004) to understand trade volumes. In that paper, the *extensive* margin is not directly observed in their aggregate data and must be inferred directly. Our firm-level data allows direct inference.

Our analysis also bears some similarity to the work of Buch et al (2004) and Feinberg (2004) who show that the composition of firms investing in different locations varies systematically with country characteristics. Finally, our results are relevant to a literature that seeks to understand the within-country organization of production (see, for instance, Holmes and Stevens, 2002).

The remainder of the paper is divided into six sections. In the first section, we describe our firm-level FDI data. In the second, we generate several stylized facts concerning both the manner in which individual firms expand abroad and the aggregate structure of multinational production across countries. In the third section, we present a simplified version of the model derived in Helpman, Melitz, and Yeaple (2004) that is consistent with the stylized facts and that can be used to better organize empirical analyses of the role of firm heterogeneity and FDI. In the fourth section, we estimate a discrete choice model of FDI motivated directly by our model. In the fifth section, we conduct a parallel analysis that makes use of a theory-inspired decomposition of the structure of multinational sales data across countries. The final section concludes.

I. Data

In this study we use the confidential enterprise-level data collected by the Bureau of Economic Analysis (BEA). The BEA conducts annual surveys of U.S. Direct Investment Abroad where U.S. direct investment is defined as the direct or indirect ownership or control by a single U.S. legal entity of at least ten percent of the voting securities of an incorporated foreign business enterprise or the equivalent interest in an unincorporated foreign business enterprise. A U.S. multinational entity (MNE) is the combination of a single U.S. legal entity that has made the direct investment, called the U.S. parent, and at least one foreign business enterprise, called the foreign affiliate. The International Investment and Trade in Services Survey Act requires that all firms larger than a certain size file detailed financial and operating items for the parent firm and each affiliate.[1]

[1] The language in this paragraph is taken from Desai, Foley, and Hines (2004).

From the BEA's database we have extracted firm-level data for the benchmark year of 1994. The BEA's reporting requirements are more stringent in these benchmark survey years, requiring firms with even small foreign levels of production to report. We include in our sample only those affiliates that are majority-owned by their U.S. parent and which are involved in either manufacturing or distribution. For each affiliate, we observe the country in which that affiliate is located and the value of the sales of that affiliate to customers located in that country. For each parent firm, we observe the value of their sales in the U.S. market. We will use a parent firm's share of the U.S. market as a proxy for a firm's relative productivity.[2] For each foreign affiliate in the dataset, we were able to identify its parent firm, thereby forming a measure of the scope and scale of each U.S. multinationals international production network. When the industry of the parent firm and the affiliate were not the same, we classified the affiliate by the industry of the parent firm.[3] When a single parent owned multiple affiliates in the same country, we aggregate the sales of these affiliates, constructing a single firm-country observation.

II. A First Look at the Structure of U.S. Multinational Enterprise

Before proposing a theoretical framework that relates firm productivity to the structure of its multinational operations, we first describe important features of the data. Our focus is on the relative importance of the *extensive* and *intensive* margins in the structure of U.S. multinational activity. These margins can be defined in two separate contexts. First, they can be used to understand the manner in which individual firms expand abroad. In this case, the *within-firm extensive margin* is the number of foreign locations in which a firm owns affiliates while the

[2] In most oligopolistic models a firm's relative size or market share in a given location is monotonic in a firm's relative productivity. This will be true of the model considered in this paper.
[3] In 1994, the BEA's industry classification system was based on the 1987 SIC classifications where the typical BEA industry lay between the two and three-digit SIC industry. Given the small number of firms in some three digit BEA industries, we aggregate the data to the level of 2-digit SIC industry.

within-firm intensive margin is the average sales per location of firms that own affiliates abroad. Second, these margins can be used to understand the structure of aggregate U.S. multinational activity across countries. In this case, the *within-country extensive margin* is the number of firms that invest in a particular country while the *within-country intensive margin* is the average sales per U.S. multinational within that country.

II.1 Within-Firm Extensive and Intensive Margins

We first analyze the relative importance of *within-firm extensive* and *intensive* margins of U.S. MNE. A particularly striking feature of the data is that most U.S. MNE own affiliates in only a handful of countries while a small number of U.S. MNE have far flung foreign operations. This heterogeneity in this *extensive* margin is illustrated in Figure 1, which shows the extent of geographic expansion of the approximately 1,500 multinational manufacturing firms. On the horizontal axis are ten categories corresponding to the number of countries in which an individual parent firm owns a foreign affiliate. On the vertical axis is the number of firms in these categories. More than a third of U.S. multinationals invest in no more than one foreign location and less than a third operate a foreign affiliate in more than five.

To describe further the characteristics of U.S. multinational expansion abroad, we now conduct a series of simple regression analyses. To do so, we must first introduce notation. Let ω index individual firms and let k index industries. Let Ω_{US}^{k} be the set of all U.S. multinational firms active in industry k, and let Ω_{i}^{k} be the set of all U.S. multinational firms in industry k that operate an affiliate in country i. Finally, let M_{ω} be the number of countries in which U.S. multinational ω operates an affiliate and let $R_{\omega US}$ be the revenue of the U.S. parent firm of U.S. MNE ω in the U.S. market.

To gauge the extent to which U.S. MNE expand abroad through the *extensive* margin relative to the *intensive* margin consider the following decomposition:

$$M_\omega\left(\frac{\sum_i R_{\omega i}}{M_\omega}\right)=\left(\frac{\sum_i R_{\omega i}}{\sum_{\omega'\in\Omega_{US}^k}\sum_j R_{ij}}\right)\left(\sum_{\omega'\in\Omega_{US}^k}\sum_i R_{\omega' i}\right).$$

On the left-hand side are total foreign sales of multinational ω to all countries decomposed into number of countries and average sales per country. On the right-hand side is the same sales expressed as a share of all U.S. MNE sales in the same industry and the total volume of U.S. MNE sales in the same industry. This decomposition motivates the regression:

$$\log M_\omega=\alpha_0+\alpha_1\log\left(\frac{\sum_i R_{\omega i}}{\sum_{\omega'\in\Omega_{US}^k}\sum_i R_{\omega' i}}\right)+\alpha_2\log\left(\sum_{\omega'\in\Omega_{US}^k}\sum_i R_{\omega' i}\right)+\mu_i.$$

The coefficient α_1 is a gauge of the relative importance of adding more foreign locations in explaining the share of that firm in its total U.S. foreign affiliate sales, while $(1-\alpha_1)$ summarizes the importance of affiliate size in explaining MNE ω's share of total U.S. foreign affiliate sales.

We report only the results for a sample in which manufacturing industries are pooled:

$$\log M_\omega=0.41*\log\left(\frac{\sum_i R_{\omega i}}{\sum_{\omega'\in\Omega_{US}^k}\sum_i R_{\omega' i}}\right)+0.47*\log\left(\sum_{\omega'\in\Omega_{US}^k}\sum_i R_{\omega' i}\right),\ \text{N=1,401;}\ R^2=0.68.$$

(0.02) (0.02)

The standard errors, shown in parentheses below the coefficient estimates, allow for heteroskedascity and clustering by parent industry. Approximately 40 percent of U.S. MNEs' expansion abroad is accomplished through an increase in the number of countries in which a firm operates (i.e. the data shown in Figure 1), while the remaining 60 percent is explained by a larger quantity of sales per foreign location. Average sales per location are at least as skewed toward a handful of firms as are the number of locations in which firms own local affiliates.

We now consider the relationship between a firm's productivity (as measured by the U.S. market share of the parent firm) and the number of foreign locations in which that parent firm owns an affiliate (the *extensive* margin) by regressing the logarithm of the number of countries in which firm ω owns a local affiliate ($\log M_\omega$) on the logarithm of that firm's U.S. sales ($\log R_{\omega US}$) and a full set of industry dummies. We find

$$\log M_\omega = \underset{(0.01)}{0.12} * \log R_{\omega US}, \quad N = 1{,}502; \quad R^2 = 0.29$$

The standard error (shown in parenthesis) is calculated allowing for clustering by firm. A ten-percent increase in the U.S. sales of the parent firm relative to the industry mean is associated with 1.2 percent more foreign locations in which the MNE opens affiliates. Put another way, a one-standard deviation increase in parent firm sales in the U.S. is associated with a 0.6 standard deviation increase in the number of countries in which the firm has affiliates.[4] The firms in Figure 1 that operate in many locations are more productive that those that operate in a few.

[4] The mean and standard deviation are 0.2 and 0.38 for the dependent variable and 12.3 and 1.84 for the independent variable. When each industry was considered separately (results not shown), we found a positive relationship

We now consider the *intensive* margin within the multinational firm as a function of its productivity by regressing the logarithm of a local affiliate of firm ω's market share in country i ($\log R_{\omega i} / \sum_{\omega' \in \Omega_i^k} R_{\omega' i}$) on the logarithm of its parent firm's market share in the United States ($\log R_{\omega US} / \sum_{\omega' \in \Omega_i^k} R_{iUS}$). The results are

$$\log\left(R_{\omega i} / \sum_{\omega' \in \Omega_j^k} R_{\omega' j} \right) = 0.68 * \log\left(R_{\omega US} / \sum_{\omega' \in \Omega_i^k} R_{\omega' US} \right); \quad N = 1{,}388; \quad R^2 = 0.33 .$$
$$(0.02)$$

The standard errors shown in parenthesis allows for clustering by firm. On average, a one-percent increase in the relative size of a parent firm's U.S. sales is associated with 0.68 percent increase in the relative size of its affiliate's sales to host country customers.

In summary, our analyses uncover two important facts concerning the expansion strategies of U.S. MNE: (i) the *extensive* and *intensive* margins are of roughly equal importance in understanding the structure of a firm's international operations, and (ii) as firms become more productive they expand in both dimensions.

II.2 Within-Country Extensive and Intensive Margins

We now use an accounting methodology similar to one suggested by Eaton et al (2004) to quantify the contribution of the *extensive* and *intensive* margins to the total sales of U.S. MNEs across countries. The sales of all U.S. MNEs in market i (R_{iUS}) can be written:

---

between a firm's market share in the U.S. and the number of its affiliates abroad although the coefficients vary considerably in terms of their economic and statistical significance across industries.

$$R_{iUS} = M_{iUS}\overline{R}_{iUS} = \lambda_{iUS}S_i,$$

where M_{iUS} is the number of affiliates of U.S. parent firms operating in country i, $\overline{R}_{iUS}$ is the average sales per affiliate in that market, λ_{iUS} is the share of U.S. affiliates in total country i sales, and S_i is the total value of sales of all firms active in market i. By regressing $\log(M_{iUS})$ on $\log(\lambda_{iUS})$ and $\log(S_i)$, we gauge how much of the variation in U.S. MNE market shares is due to variation in number of affiliates and how much to variation in average sales per firm. We conduct this decomposition using industry-country output data from the Industrial Statistics Database maintained by UNIDO for a total of 40 countries and twenty-one 3-digit ISIC industries[5]. The results are presented in Table 1. Column one corresponds to a specification in which all data were aggregated to the level of manufacturing, and the remaining two columns correspond to specifications in which the data is disaggregated by industry and pooled.

First, we interpret the coefficient estimates in column one: holding fixed a country's market size, a higher U.S. multinational market share is due to a 61 percent increase in the number of affiliates and a 39 percent increase in the average sales per affiliate. Further, holding fixed the market share of U.S. affiliates, a larger market reflects 56 percent more affiliates and 44 percent more sales per firm. These results are similar to those presented for French exports in Eaton et al (2004): much of the variation of US MNE sales across countries can be attributed to variation in the number of firms participating in the market (the *extensive* margin). To explore the possibility that aggregate sample masks the effect of variation in industrial mix of FDI across

[5] There are 26 non-petroleum industries in the INSTAT database. To ease concordance, we combined industries 321 and 322, 323 and 324; and 361 and 369. The countries are Argentina, Australia, Austria, Canada, Chile, China, Colombia, Costa Rica, Denmark, Ecuador, Egypt, Finland, France, Greece, Honduras, Hong Kong, India, Indonesia, Ireland, Israel, Italy, Japan, Korea, Malaysia, Mexico, Netherlands, New Zealand, Norway, Peru, Portugal, Singapore, South Africa, Spain, Sweden, Switzerland, Taiwan, Thailand, Turkey, United Kingdom, and Venezuela.

countries, we disaggregate by industry, including a full set of industry dummies to pick up industry effects. As shown in column 2, industry heterogeneity is important: the *extensive* margin is smaller at 41 percent, and the *intensive* margin appears more important (59 percent of variation in market share). Finally, because many countries do not receive U.S. FDI in many industries, we consider a Heckman selection correction specification (column 3).

In summary, variation in the market shares of U.S. multinational affiliates can be attributed to both to the *extensive* (more affiliates) and intensive (more sales per affiliate) margins, but the *intensive* margin plays a somewhat larger role. Holding fixed the market share of U.S. multinational enterprises, an increase in a foreign country's output leads to equal changes in the *extensive* and *intensive* margins. In the next section, we present a model of firm heterogeneity and FDI that will help explain these facts and that motivates a more formal econometric analysis of the structure of multinational enterprise.

III. Conceptual Framework

In this section, we provide a rudimentary sketch of the framework presented in Helpman et al (2004).[6] We use the model to motivate two empirical analyses, one that makes use of the investment behavior of individual firms and the other that makes use of cross-country variation in aggregate U.S. MNE sales.

There is a continuum of firms of mass N in a country called "home." Each firm is capable of producing a single, and unique, variety of a differentiated product. The firm wishes to sell this product in each of I countries. The demand for variety ω in country i is

$$x_i(\omega) = A_i p_i(\omega)^{-\sigma}$$

[6] Many of the ideas in the recent literature on firm heterogeneity and international trade have existed in an informal sense for a long time. See Horst (1972) for an argument with some of the implications of the model presented in this section.

where A_i is the demand level in country i, $\sigma > 1$ is the elasticity of substitution across varieties, and $p_i(\omega)$ is the price charged for variety ω in country i. If $c_i(\omega)$ is the constant marginal cost of serving country i for the producer of variety ω, and $\alpha = \sigma/(\sigma - 1)$, then firm ω's optimal pricing decision yields revenue

$$R_i(\omega) = \alpha^{1-\sigma} A_i c_i(\omega)^{1-\sigma}. \tag{1}$$

Firms vary in their productivity φ: a firm of productivity φ requires $1/\varphi$ bundles of inputs to produce one unit of output. The empirical distribution of φ is given by the cumulative distribution function $G(\bullet)$. We assume that a bundle of factors (measured in efficiency units) in country i is w_i while the cost of a bundle of inputs in home is normalized to unity.

To serve country i, a firm has two options: export to country i from the home country or build an affiliate in country i.[7] Firms that export face an iceberg-type transport costs of $\tau_i > 1$. A firm that serves country i via a local affiliate avoids transport costs but must incur a fixed set-up cost f_i. The profits that would accrue to a firm of productivity φ if it exports to country i is

$$\pi_i^X(\varphi) = \frac{\alpha^{1-\sigma}}{\sigma} A_i \left(\frac{\varphi}{\tau_i} \right)^{\sigma-1}. \tag{2}$$

The profits that would accrue to this firm if it instead engaged in multinational production in country i is

[7] We abstract from the possibility that a firm serves a country from an affiliate in a different foreign country.

$$\pi_i^M(\varphi) = \frac{\alpha^{1-\sigma}}{\sigma} A_i \left(\frac{\varphi}{w_i}\right)^{\sigma-1} - f_i \tag{3}$$

Equations (2) and (3) are graphed in Figure 2 under the assumption that $w_i < \tau_i$. From Figure 2, it is clear the profits to export exceed the profits to multinational production for low productivity firms while the opposite is true for high productivity firms: avoiding transport cost is more important to the high productivity firms that wish to sell a large number of units in country *i*. If $w_i < \tau_i$, then a firm with productivity φ will invest abroad if $\pi_i^M \geq \pi_i^X$. Using equations (2) and (3), this condition becomes

$$\frac{A_i}{\sigma\alpha^{\sigma-1}}\left(w_i^{1-\sigma} - \tau_i^{1-\sigma}\right)\varphi^{\sigma-1} \geq f_i . \tag{4}$$

Equation (4) holds with equality for the least productive firm that finds it profitable to invest in country *i*, which can be written

$$\overline{\varphi}_i = \alpha\left[\frac{\sigma f_i}{A_i(w_i^{1-\sigma} - \tau_i^{1-\sigma})}\right]^{\frac{1}{\sigma-1}} . \tag{5}$$

An increase in $\overline{\varphi}_i$ is associated with a reduction in the number of firms investing abroad as fewer firms are sufficiently productive to make multinational production superior to exports. From equation (5) we see that a decrease in the market-size adjusted fixed cost f_i / A_i, a decrease in

the host country factor prices w_i, or an increase in the transport cost τ_i is associated with a lower $\overline{\varphi}_i$ and relatively larger volume of multinational production. Note that the more productive is a firm, the larger the number of countries for which the payoff to foreign direct investment will exceed the payoff to exporting from the home country.

Our discussion has yielded two empirically relevant predictions regarding the structure of international production across heterogeneous firms: (i) more productive firms earn larger revenues in every market in which they produce relative to less productive firms, and (ii) more productive firms own local affiliates in a larger set of countries than less productive firms. Both predictions are consistent with the features of the data presented in the previous section.

We now turn our attention to the predictions of the model regarding the structure of aggregate multinational activity by country. By aggregating over firms that invest in country *i*, we can construct a measure of the total volume of the sales in that country by the home's multinationals. Letting R_i be the total revenue of the affiliates of home firm in country *i*, we find

$$R_i = \left\{N\left[1-G(\overline{\varphi}_i)\right]\right\} \times \left\{\frac{1}{\left[1-G(\overline{\varphi}_i)\right]}\int_{\overline{\varphi}_i}^{\infty} \varphi^{\sigma-1} dG(\varphi)\right\} \times \left\{\left(w_i \alpha\right)^{1-\sigma} A_i\right\}, \qquad (6)$$

where $\overline{\varphi}_i$ is defined in equation (5). Equation (6) organizes the volume multinational sales across countries into three components, each shown in separate curly brackets. In the first bracket, is the *extensive* margin, or the total number of home country firms that own an affiliate in country *i*. In the second curly bracket is the *composition component of the intensive margin*, or the average productivity, of home country firms that own an affiliate in country *i*. In the final curly bracket is the *scale component of the intensive margin*, which is a country-specific, size-of-affiliate effect.

It is clear from equation (6) that the term in the first curly brackets is decreasing in $\overline{\varphi}_i$, while the term in the second curly brackets is increasing in $\overline{\varphi}_i$. This relationship yields another prediction: the average productivity of multinational affiliates in a country that attracts a large number of affiliates is lower than in a country attracts fewer multinationals. This prediction follows directly from the observation that the cutoff productivity is below the average productivity of firms engaged in FDI. Below, we assess the empirical analog of equation (6) using U.S. MNE data.

IV. Firm Productivity and the Decision to Investment

In this section, we use equation (4) to specify a discrete choice model of a firm's FDI decisions across countries. While all the firms in our sample are multinationals (they invest in at least one other country), there is substantial variation in the number of countries in which these firms invest. To operationalize equation (4) we assume that the fixed cost facing firm ω in country *i* is partially stochastic: $f_i = F_i \exp(\varepsilon_{i\omega})$, where F_i is a non-stochastic component to the cost of investing in country *i* and $\exp(\varepsilon_{i\omega})$ is a stochastic firm-country component. As now interpreted, the logarithm of equation (4) yields an equation with the form

$$\beta \ln Z_i + \gamma \ln \Phi_\omega \geq \varepsilon_{i\omega}, \qquad (7)$$

where Z_i is a vector of country characteristics associated with market size, factor prices, transport and fixed costs of investment, and Φ_ω is a proxy for firm productivity φ_ω. By making appropriate distributional assumptions over $\varepsilon_{i\omega}$ we can estimate (7) as either a Logit or Probit.

The dependent variable, *Invest*, is equal to one if a firm owns an affiliate in a particular country and equal to zero otherwise.

We now turn to the independent variables. Our choice of country variables is guided by the existing literature that emphasizes market size and level of development as key determinants of foreign direct investment. The first two variables *GDPPC* and *POP* are the logarithm of real GDP per capita and the logarithm of the population size. These variables measure the level of development and the size of the market respectively. As measures of the ease of investing in a foreign location we include as regressors *DIST*, *ADJ*, and *ENGLISH*, which are the logarithm of the distance from the U.S., a dummy for Canada and Mexico (adjacent), and the share of the population that speaks English respectively. To capture a wide range of host country institutions that affect the cost of production we include a measure of "social infrastructure," *SOCINF*, which was constructed by Hall and Jones (1999). Finally, our measure of firm productivity (*MKTSH*) is a firm's share of the U.S. market. Descriptive statistics for the dependent and independent variables can be found in Table 2.

The baseline results, corresponding to equation (7), are shown in column one of Table 3.[8] The coefficients on the six country variables are statistically significant and conform to results found elsewhere. U.S. MNEs are more likely to enter countries that are highly developed (*GDPPC*) and have large populations (*POP*). U.S. MNEs are also more likely to enter countries that are "close" to the U.S. in both a physical distance sense (*DIST* and *ADJ*) and in a linguistic sense (*ENGLISH*). Finally, countries with good social infrastructure, and therefore are more likely to have liberal economic policies, attract many U.S. MNE (*SOCINF*).

Turning to our measure of firm productivity (*MKTSH*), we find that the probability that a firm invests in any given country is sharply increasing in its share of the U.S. market. This result

[8] The standard errors allow for dependence within firms.

is consistent with our finding in section II that firms with large U.S. market shares invest in a larger number of foreign countries. As predicted by the model, the more productive the firm the larger the expected size of its multinational network.[9]

We now take our analysis a step further by introducing interaction terms between the firm productivity variable and the six country characteristics. The logit results for this specification are shown in column (2). The coefficient estimates for each country characteristic keep their sign and remain statistically significant. The estimates on the interaction terms are statistically significant for each country characteristic and, with the exception of the coefficient on *MKTSH*DIST*, all have the opposite sign of the country characteristic alone.

To understand these results, we carefully interpret the negative coefficient on *GDPPC* and *MKTSH*. Our interpretation of the other interaction coefficients is similar. An increase in a country's real GDP per capita makes U.S. MNE more likely to invest there, but the strength of this effect is decreasing in a firm's productivity. More productive firms are less likely to be near the cutoff productivity level and so are less affected by a country's GDP per capita. Put another way, a firm's relative productivity (*MKTSH*) is a less important determinant of whether a firm invests in a given location if that location features a relatively high real GDP per capita. To show that our results are not sensitive to our distributional assumption, the results of estimating the model as a PROBIT are shown in column (3). The estimates shown are marginal effects evaluated at independent variable means.

In summary, the results are highly suggestive of a country-varying "productivity" cutoff between those high productivity firms that invest abroad frequently and those low productivity firms that invest abroad rarely, i.e. there is an *extensive* margin that varies across countries.

[9] A Probit specification yielded remarkably similar results. In the interest of space we omit them here.

V. Extensive and Intensive Margins Within-Country

We now turn our attention to the cross-country structure of U.S. multinational production with an explicit focus on the importance of the *extensive* margin on one hand, and the productivity *composition* of multinational production across countries on the other. To do so, we define a decomposition technique that is the empirical analog to equation (6):[10]

$$\sum_{\omega \in \Omega_i} R_{\omega i} = M_i \times \left(\frac{1}{M_i} \sum_{\omega \in \Omega_i} R_{\omega US} \right) \times \left(\frac{\sum_{\omega \in \Omega_i} R_{\omega i}}{\sum_{\omega \in \Omega_i} R_{\omega US}} \right). \tag{8}$$

On the left-hand side of equation (8) is the total revenue of the affiliates of U.S. MNE located in country *i* to customers in country *i*. On the right-hand side there are the three components into which these sales can be decomposed. The first is the number of all U.S. MNE that have an affiliate located in country *i*. This is the *extensive* margin. The second is the average size of U.S. MNE in terms of parent firm sales to U.S. customers of those MNE that invest in country *i*. This is the *composition component of the intensive* margin, and it measures the relative productivity of firms that invest in country *i*. The last term in parentheses is the ratio of the aggregate local sales of the affiliates of U.S. MNE investing in country *i* to the aggregate local sales of the parents of these same firms in the U.S. market. This is the *scale component of the intensive* margin, and it captures differences in affiliate sizes holding fixed the composition of firms.[11] We rewrite this relationship as

[10] Industry superscripts have been suppressed.
[11] The scale and composition margin are decompositions of the *intensive* margin or average sales per affiliate.

$$r_i \quad = \quad n_i \times c_i \times s_i .$$

We investigate how these three components of U.S. MNE activity are related to country characteristics. Let $y_i \in \{r_i, n_i, c_i, s_i\}$. For each component of FDI, we estimate:

$$\log y_i = \chi \log Z_i + \mu_i \ , \tag{9}$$

where Z_i is a vector of country characteristics often associated with FDI volumes. Because the decomposition (equation (8)) is separable in logarithms and OLS is linear operator, the coefficient estimates on the set of country characteristics for the three components of FDI must sum to the coefficient estimate obtained by regressing FDI volume (r_i) on the same country characteristics. Since there are a significant number of country-industry pairs in which there is zero U.S. MNE activity, we estimate (9) using the standard Two-step Heckman selection correction.[12] We use the same set of country characteristics Z_i as in the previous section and include a full set of industry indicator variables.[13]

The results (coefficients on industry dummies suppressed) are shown in Table 4. The first two columns correspond to the aggregate sales and probit selection equations, while the remaining three columns correspond to the three components of FDI. Beginning with the total sales specification, shown in the first column, we see that U.S. multinational sales are primarily made in large developed markets as indicated by the positive and statistically significant coefficients on *GDPPC* and *POP*. U.S. FDI is also more prevalent in those countries in which it

[12] The results presented below are robust to a variety of estimation techniques, including ordinary least squares (see Appendix Table 3) and median regression. They are a bit stronger in the median regression due to a few outliers that we have not dropped from the dataset.

[13] To see a list of countries and descriptive statistics for $\log(y_i)$, see the appendix.

is relatively easy for U.S. firms to enter: those that are geographically close to the U.S. (negative coefficient on *DIST* and positive coefficient on *ADJ*) and those where English is the dominant language (positive coefficient on *ENGLISH*). Finally, U.S. MNEs are more active in countries with quality social infrastructure (*SOCINF*).

Turning to the *extensive* margin, consider the coefficients in the second and third columns of Table 4.[14] The signs of the coefficient estimates match those of the FDI sales equation and are highly consistent across both the PROBIT component (column 2) and the OLS component (column 3). A larger number of firms locate affiliates in large, developed, and easy-to-enter countries. The coefficients in the *extensive* margin equation are large relative to those in the FDI sales equation indicating that the *extensive* margin is a large component of the variation in U.S. multinational sales across countries. For instance, roughly half of the expansion in FDI sales associated with *GDPPC* is due to expansion in the *extensive* margin.

Now consider the coefficient estimates corresponding to the *composition component of the intensive margin* shown in fourth column. With the exception of *DIST*, all of these coefficients have exactly the opposite signs of those in the *extensive* margin equation: those country characteristics associated with a large number of U.S. entrants are also associated with the entry of U.S. MNE with lower productivity. For instance, Canada and Mexico enjoy FDI from approximately 65 percent more U.S. firms, but the average productivity of these firms is approximately 35 percent less than in other countries. Further, countries where the dominant language is English receive 60 percent more entrants than those that speak no English, but average productivity of firms investing in English-speaking countries is 18 percent lower. The countries that are easier for U.S. firms to enter attract relatively more marginal firms. These results are consistent with the observation by Horst (1972) and by Feinberg (2003) that these

[14] It is natural to consider the probit coefficients with the *extensive* margin.

easy-to-enter countries (such as Canada) are "starter" foreign locations for U.S. MNE. The difference in the response to unobserved country characteristics also appears in these results via the coefficients on the Mills ratio in columns 3 and 4.[15]

Now consider the coefficients corresponding to the *scale component of the intensive* margin shown in the last column. Of the six covariates only the coefficients on *GDPPC* and *POP* are statistically significant. U.S. multinational affiliates located in large, developed countries have a larger scale than those in small, developing countries. The results are consistent with a story in which the ability to sell relatively more in large, developed countries (i.e. a larger potential scale) attracts a larger number of U.S. multinational firms. The similarity in the coefficient estimates for the extensive and scale margins is reminiscent of a result of Holmes and Stevens (2002), who show that U.S. counties with a high concentration of an industry's plants also tend to have plants that are disproportionately large.

VI. Conclusion

Using the firm-level data of the BEA, we have presented a thorough set of analyses of the importance of the *extensive* and *intensive* margins both *within-firms* and *within-countries* in the structure of U.S. multinational activity. We began our study by generating a set of stylized facts: the most productive firms (i) own affiliates in a larger set of countries and (ii) generate larger sales revenue in those countries in which they operate. Further, the *extensive* margin is of almost equal importance to the *intensive* margin in the cross-country variation in U.S. MNE activity.

[15] When each industry is considered separately, the tendency for the coefficients in the extensive and composition component regressions to have different signs is somewhat less stark (see Appendix Table 4) This is in part due to larger standard errors associated with smaller sample size. Nevertheless, the same basic message appears consistently and is particularly pronounced in certain industries, such as Rubber and Plastic Products (SIC 30), Fabricated Metal Products (SIC 34), Machinery (SIC 35), and Instruments (SIC 38).

We then specified a simplified version of the firm heterogeneity, sorting, and foreign direct investment model presented in Helpman et al (2004). In addition to being shown to be consistent with the stylized facts, the framework was used to motivate two econometric analyses. The first exploited within-firm variation in a discrete choice framework. We showed that the probability that a firm invests in a given foreign country is increasing in that firm's productivity, but that firm productivity plays a less important role in the investment decision of firms in those countries most attractive to multinational firms. In the second analysis we used a technique derived from the theoretical model to decompose the volume of FDI across countries into an *extensive* margin (the number of firms investing in a location), a *composition* margin (the average productivity of the firms investing in that location), and a *scale* margin (the residual size-of-affiliate). We showed that country characteristics that induce the entry of a larger number of U.S. MNEs are also associated with a change in the composition of investing firms toward those with below-average productivity, i.e. as predicted by the model an expansion in the number of firms investing in a location involves attracting progressively less productive firms. The two empirical analyses, one exploiting *within-firm* variation and the other exploiting *within-country* variation, both provide strong support for the predictions of Helpman et al (2004).

While our results highlight the empirical importance of the sorting of heterogeneous firms into mode of foreign market access, one drawback of our model is that it ignores the possibility that firms use one country as an export platform to another. Future empirical work should be directed toward better understanding the tradeoff facing heterogeneous firms in their strategy serving a region: consolidate production in a single location versus serving each country from a local affiliate.

References:

Buch, Claudia, Jorn Kleinert, and Farid Toubal. 2003. "Heterogeneity and FDI: Evidence from German Firm-Level Data." Mimeo, Institute for World Economics.

Desai, Mihir, Fritz Foley, and James Hines. 2004. "Dividend Policy inside the Firm." Mimeo Harvard University.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2005. "Dissecting Trade: Firms, Industries, and Export Destinations." Forthcoming *American Economic Review*.

Feinberg, Susan. 2003. "The Expansion and Location Patterns of U.S. Multinationals." Mimeo University of Maryland.

Girma, S., R. Kneller, and M. Pisu. 2003. "Exports versus FDI: An Empirical Test." GEP Research Paper 03/21.

Hall, Robert and Charles Jones. 1999. "Why do Some Countries Produce so Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114(1): 83-116.

Head, K. and J. Ries. 2003. "Heterogeneity and the FDI versus Export Decision of Japanese Manufacturers." *Journal of the Japanese and International Economies*.

Helpman, E., M. Melitz, and Y. Rubinstein. 2004. "Trading Partners and Trading Volumes." Mimeo, Harvard University.

Helpman, E., M. Melitz, and S. Yeaple. 2004. "Exports versus FDI with Heterogeneous Firms." *American Economic Review* 94(1): 300-16.

Holmes, T. and J. Stevens. 2002. "Geographic Concentration and Establishment Scale." *Review of Economics and Statistics* 84(4): 682-690.

Horst, T. 1972. "Firm and Industry Determinants of the Decision to Invest Abroad: An Empirical Study." *Review of Economics and Statistics* 54(3): 258-266.

Mataloni, R. and D.Yorganson. 2002. “Operations of U.S. Multinational Companies: Preliminary Results for the 1999 Benchmark Survey.” *Survey of Current Business*, March: 24-54.

Melitz, Marc. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica* 71: 1695-1725.

Figure 1: Number of Countries in which Manufacturing Firms have Affiliates

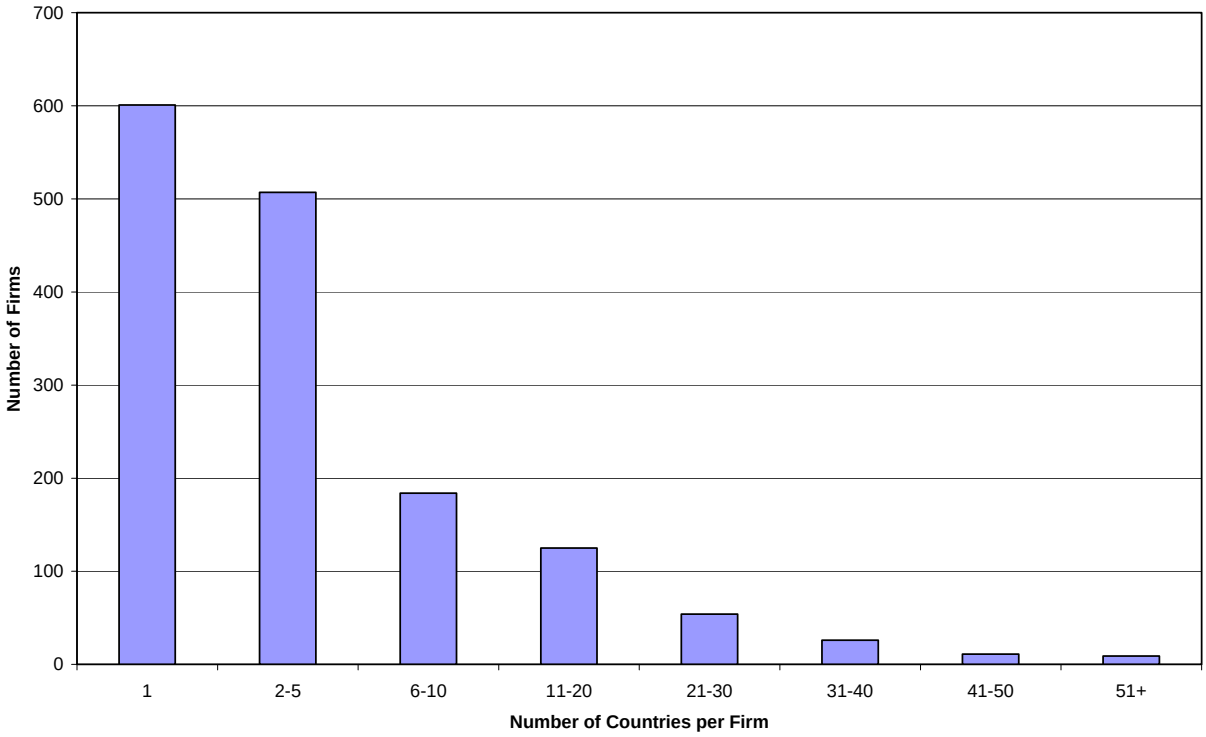

Figure 2

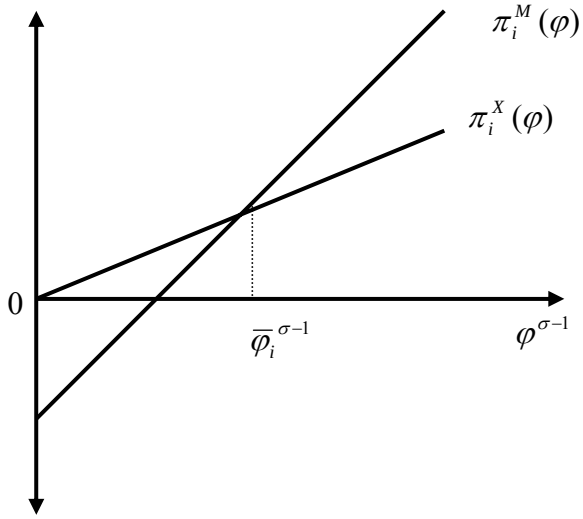

Table 1: A Decomposition of U.S. Affiliate Sales into Intensive and Extensive Margins

	(1) Aggregate Data	(2) Pooled Industry Data OLS	(3) Pooled Industry Data Heckman
λ_{kUS}	0.61	0.41	0.41
	(0.06)	(0.07)	(0.02)
X_k	0.56	0.50	0.46
	(0.05)	(0.04)	(0.02)
N	40	515	783
R^2	0.89	0.77	
χ^2			1349

Standard errors are shown in parentheses. The standard errors in column one and two are robust to heteroskedasticity. The standard errors in column two have been corrected for clustering at the country level. A full set of industry indicators were included in the pooled specifications. The selection equation in the Heckman specification include total industry sales by country, real GDP, distance from U.S., English as the dominant language and a full set of industry indicators.

Table 2: Descriptive Statistics

Variable	Mean	Standard Dev.
Invest	0.05	0.22
GDPPC	9.37	0.79
POP	15.38	1.74
ENG	0.14	0.32
SOCINF	0.60	0.25
DIST	8.81	0.61

Table 3: The Decision to Invest Abroad: Country and Firm Characteristics and the Extensive Margin

Dependent Variable = Invest

	(1) LOGIT	(2) LOGIT	(3) PROBIT
GDPPC	0.951 (0.048)	0.599 (0.079)	0.008 (0.001)
POP	0.561 (0.012)	0.394 (0.024)	0.005 (0.0005)
DIST	-0.152 (0.041)	-0.417 (0.090)	-0.006 (0.003)
ENGLISH	1.066 (0.040)	0.483 (0.091)	0.008 (0.001)
ADJ	0.855 (0.091)	0.744 (0.211)	0.014 (0.006)
SOCINF	2.313 (0.086)	1.520 (0.209)	0.021 (0.003)
MKTSH	0.442 (0.025)	3.311 (0.342)	0.029 (0.004)
GDPPC MKTSH*		-0.116 (0.020)	-0.0008 (0.0002)
POP MKTSH*		-0.046 (0.006)	-0.0003 (0.00007)
DIST MKTSH*		-0.087 (0.022)	-0.001 (0.0003)
ENGLISH MKTSH*		-0.125 (0.022)	-0.001 (0.0003)
ADJ MKTSH*		-0.066 (0.055)	-0.0008 (0.0006)
SOCINF MKTSH*		-0.230 (0.056)	-0.002 (0.0007)
N Log-Likelihood	164,182 -23188	164,182 -22990	164,182 -23021

Standard errors in parentheses allow for clustering by firm. Probit results in column (3) are marginal effects evaluated at the mean of the dependent variables. Constant term is suppressed.

Table 4: The Determinants of FDI and each of its Components

	(1) Sales (r)	(2) Probit	(3) Extensive (n)	(4) Composition (c)	(5) Scale (s)
GDPPC	2.24 (0.13)	1.22 (0.13)	1.24 (0.07)	-0.43 (0.08)	1.41 (0.13)
POP	0.93 (0.05)	0.60 (0.05)	0.57 (0.03)	-0.09 (0.03)	0.45 (0.04)
DIST	-0.29 (0.11)	-0.27 (0.12)	-0.16 (0.06)	-0.10 (0.07)	-0.08 (0.10)
ENGLISH	0.51 (0.15)	0.34 (0.20)	0.60 (0.08)	-0.18 (0.10)	0.16 (0.14)
ADJ	0.52 (0.30)	1.02 (0.61)	0.65 (0.16)	-0.35 (0.19)	0.13 (0.28)
SOCINF	0.71 (0.34)	2.00 (0.35)	1.41 (0.18)	-0.08 (0.22)	-0.49 (0.32)
Mills	0.40 (0.21)		0.72 (0.10)	-0.72 (0.13)	0.51 (0.20)

N = 1140. Industry dummy variable coefficients have been suppressed.

Appendix Table 1: Country Coverage

Argentina	Italy
Australia	Jamaica
Austria	Japan
Bangladesh	Korea
Belgium	Malaysia
Bolivia	Mexico
Brazil	Netherlands
Canada	New Zealand
Chile	Nicaragua
China	Norway
Colombia	Pakistan
Costa Rica	Papua New Guinea
Denmark	Paraguay
Dominican Republic	Peru
Ecuador	Philippines
El Salvador	Portugal
Finland	Romania
France	Singapore
Germany	South Africa
Greece	Spain
Guatamala	Sri Lanka
Haiti	Sweden
Honduras	Switzerland
Hong Kong	Taiwan
Iceland	Thailand
India	Trinidad & Tobago
Indonesia	United Kingdom
Ireland	Uruguay
Israel	Venezuela

Appendix Table 2: Components of U.S. Multinational Sales Abroad

	Mean	Standard Deviation
R (sales)	11.6	2.4
N (number)	1.8	1.4
C (composition)	14.9	1.2
S (scale)	-5.1	1.5

Appendix Table 3: The Determinants of FDI and each of its Components

	(1) Sales (r)	(2) Extensive (n)	(3) Composition (c)	(4) Scale (s)
GDPPC	1.62 (0.22)	0.75 (0.08)	-0.39 (0.07)	1.26 (0.21)
POP	0.69 (0.09)	0.38 (0.05)	-0.07 (0.02)	0.39 (0.08)
DIST	-0.24 (0.17)	-0.15 (0.10)	0.01 (0.08)	-0.12 (0.13)
ENGLISH	0.44 (0.33)	0.51 (0.19)	-0.25 (0.09)	0.22 (0.20)
ADJ	0.30 (0.38)	0.41 (0.24)	-0.17 (0.14)	0.00 (0.29)
SOCINF	-0.09 (0.52)	0.70 (0.24)	0.01 (0.26)	-0.76 (0.50)
R-sq F-stat	0.31 125	0.34 73	0.06 26	0.30 100

N = 662. Estimated via OLS. Robust standard errors allowing for clustering are shown in parantheses.

Appendix Table 4: Margins and Country Characteristics, by Industry

Ind.	margin	GDPPC	POP	DIST	ENG	ADJ	SINF	N	R^2	F-Stat
20	n	**0.85**	**0.48**	**-0.33**	-0.15	0.22	0.52	57	0.75	**40**
	c	0.29	0.11	-0.10	-0.29	**-0.84**	-0.26	57	0.53	**2.3**
	s	**0.87**	**0.24**	0.23	**0.28**	1.05	-0.69	57	0.40	**8.3**
21	n	**0.62**	**0.47**	**-0.60**	**0.82**	-0.92	0.74	28	0.72	**19**
	c	-0.57	**0.45**	-0.25	**1.06**	0.62	**2.94**	28	0.46	**4.2**
	s	**2.58**	0.37	-0.90	1.40	**-3.36**	-5.24	28	0.43	**5.6**
23	n	-0.29	**0.28**	**-0.84**	**1.02**	-0.72	1.15	26	0.65	**29**
	c	0.37	-0.04	**0.49**	**-0.82**	0.37	-0.51	26	0.66	**18**
	s	**1.17**	**0.67**	-0.51	0.50	-0.79	-0.60	26	0.75	**11**
25	n	0.15	0.30	-0.35	0.42	0.25	1.61	24	0.52	**17**
	c	-0.13	-0.08	-0.14	0.08	**-0.84**	-0.38	24	0.35	**4.4**
	s	**1.81**	0.13	**-0.84**	0.25	-0.14	**-3.17**	24	0.70	**28**
26	n	**0.87**	**0.42**	-0.23	**0.49**	0.43	**0.98**	48	0.72	**56**
	c	0.10	0.03	-0.12	**-0.66**	**-0.89**	**-0.78**	48	0.54	**11**
	s	**1.71**	**0.59**	**0.59**	-1.02	**1.69**	-1.51	48	0.57	**6.1**
27	n	**1.36**	**0.45**	0.27	**1.18**	**1.21**	**0.75**	36	0.84	**64**
	c	**-0.42**	-0.06	0.17	**-0.30**	0.06	0.41	36	0.38	**12**
	s	**1.54**	**0.45**	0.07	**0.58**	0.18	-0.38	36	0.60	**15**
28	n	**0.95**	**0.49**	-0.19	0.28	0.23	**1.52**	55	0.78	**52**
	c	-0.03	0.06	0.02	0.10	**-0.34**	0.14	55	0.08	**2.2**
	s	**0.65**	**0.36**	-0.13	-0.04	0.42	-0.25	55	0.54	**62**
30	n	**1.32**	**0.49**	0.06	**0.85**	**0.92**	0.47	41	0.75	**39**
	c	**-0.71**	**-0.09**	-0.17	-0.27	-0.47	-0.51	41	0.57	**24**
	s	**1.68**	-0.02	0.16	-0.10	1.03	-2.33	41	0.23	**3.6**
33	n	**0.85**	**0.43**	-0.04	**1.45**	0.83	**1.68**	29	0.76	**28**
	c	-0.16	**0.10**	0.01	**-0.46**	-0.06	0.01	29	0.28	**3.7**
	s	1.46	-0.26	0.09	**1.27**	0.50	-1.91	29	0.30	**2.2**
34	n	**0.99**	**0.54**	**-0.46**	**0.63**	0.26	**1.39**	47	0.81	**56**
	c	**-0.35**	**-0.09**	0.02	**-0.40**	-0.16	0.20	47	0.50	**28**
	s	**1.29**	**0.45**	-0.05	-0.18	0.32	**-1.04**	47	0.72	**100**
35	n	**1.48**	**0.61**	0.03	**0.59**	0.65	**1.49**	53	0.89	**95**
	c	**-0.49**	**-0.26**	0.13	-0.00	0.03	-0.52	53	0.41	**7.5**
	s	**1.33**	**0.56**	0.03	-0.34	-0.07	**1.27**	53	0.76	**23**
36	n	**1.27**	**0.59**	0.11	**0.45**	**1.04**	**1.74**	42	0.83	**30**
	c	0.10	**0.30**	0.27	**-0.99**	-0.09	0.46	42	0.37	**5.9**
	s	0.48	-0.02	-0.41	1.09	-0.05	0.64	42	0.33	**7**
37	n	**1.19**	**0.54**	0.27	**0.56**	**1.66**	**1.56**	42	0.85	**32**
	c	-0.48	-0.05	**-0.80**	0.11	-0.13	1.93	42	0.17	1.1
	s	**1.32**	**0.61**	0.01	0.34	0.16	-0.79	42	0.53	**8.6**
38	n	**1.44**	**0.61**	**-0.40**	0.40	-0.31	**1.55**	50	0.88	**54**
	c	**-0.61**	**-0.32**	**-0.38**	-0.24	**-0.48**	-0.68	50	0.73	**46**
	s	**1.09**	**0.55**	0.18	0.38	0.36	0.80	50	0.82	**45**

Estimated via OLS, Bold indicates statistical significance (robust std errors) at 10 percent level.