Using Maximum Entropy Outlier Analysis to Identify Multinational Corporation Tax Havens

Daniel L. Friesner
North Dakota State University

Andrew M. Brajcich
Gonzaga University

In many practical applications of finance and accounting, reporting mean or total values mask many underlying trends in data, especially when the researcher or manager is interested in understanding how specific observations relate to the data. If researchers or managers are interested in these trends, using inferential statistics that focus on measures of central tendency may not allow for identification of these relationships. This paper uses maximum entropy techniques drawn from ecology literature to provide a robust framework to make empirical inferences about particular groups of observations, which we refer to as "influential observations". The framework may be applied to individual-level data as well as aggregate data. It is also relatively simple to implement using Excel or similar programs. We apply the technique to data furnished by the IRS for 2008, which describes the levels of multinational corporation subsidiary development in a given country. We use the empirical framework to identify those countries in which multinational corporations are establishing corporate subsidiaries to avoid paying U.S. federal income tax.

INTRODUCTION

In most traditional applications of statistics and econometrics, the primary objective is to use a randomly collected sample to make inferences about the underlying population. In those instances, the primary focus is typically placed upon estimating one or more measures of central tendency (usually the mean or the median) and dispersion (usually a variance or standard deviation). The impacts of specific groups of observations on these measures of central tendency or dispersion are generally not explored, except in a residual fashion. In many statistical procedures (including, but not limited to regression analysis), outliers or other related statistical anomalies are usually treated as nuisance observations which bias the measures of central tendency and dispersion, and adjustments are usually made to reduce the impacts of these outliers on the estimates of central tendency and/or dispersion.

However, there are many practical applications in which specific groups of observations (of which outliers and/or other statistical anomalies are special cases) are the primary phenomena of interest. For example, in applications of international finance, the Internal Revenue Service (IRS) is concerned with identifying those multinational corporations that attempt to establish subsidiaries for the purpose of avoiding the payment of U.S. income tax, e.g. through the shifting of income out of the U.S., knowing that a relatively small number of foreign subsidiaries are established solely for these purposes (Brajcich, Friesner, & McPherson, 2013a).[1] Here, the focus of the IRS is not placed upon mean levels of corporate income and tax payments so much as it is interested in identifying statistical anomalies, where a firm might have an extraordinarily high level of subsidiary income with low levels of repatriated earnings to the U.S.-based parent corporation.

Researchers attempting to use secondary data to examine various financial issues face similar challenges. In many cases, publicly available data are aggregated, reporting totals, measures of central tendency or measures of dispersion across various firms in a given industry or across related industries. While those reported statistics are certainly valuable, the researcher might be interested in the distribution of firms within those industries that create such statistics. Within the context of international finance, the researcher might have access to corporate income tax payments and repatriated income from international subsidiaries within a broadly defined industry. The researcher might be interested in determining the extent to which specific sub-industries (or groups of subsidiaries, as defined by higher digit NAICS codes) contribute to this mean. For example, U.S. based multinational corporations (MNCs) in chemical manufacturing industry have established subsidiaries in a variety of different countries (each with its own tax laws and jurisdiction) and manufacture a variety of different products. Moreover, firms that manufacture a specific type of chemical product (say, pharmaceuticals) might find it beneficial to locate their manufacturing facilities in specific countries with beneficial tax rules, other regulatory characteristics and resource markets which facilitate lower production costs or allow the MNC to shield income generated by the subsidiary from the IRS. But if aggregate statistics reported by the IRS merely report MNC activity in the chemical manufacturing industry by country, and similarly report disaggregated total activity in the chemical manufacturing industry by sub-industry, but do not disaggregate data jointly across these two variables, it is difficult for researchers to determine if the distribution of pharmaceutical manufacturing firms are locating their subsidiaries in a different set of countries than other chemical manufacturing MNCs.

To date, few tools are available to provide meaningful empirical inferences in situations where the entire population distribution (including the tails of the distribution, in which outliers and other empirical anomalies exist) is the primary focus of the analysis. In this paper, we adapt maximum entropy techniques to provide an appropriate framework to make data-driven inferences about outliers and other empirical anomalies, which we refer to as "influential observations" (Eff, Gschwend, & Johnston, 2008; Ruther et al., 2013). The framework is not only robust, in that it can be applied to individual-level data as well as aggregate data (i.e., cross-tabulations and/or summaries of individual-level data), but it is also relatively simple to implement using Excel or other spreadsheet programs. We apply the technique to data taken from IRS data in 2008 (Form 5471, *Information Return of U.S. Persons With Respect To Certain Foreign Corporations*) which describes the levels of multinational corporation subsidiary development in a given country. We use the empirical framework to identify those countries in which multinational corporations are establishing corporate subsidiaries to avoid paying U.S. federal tax. Additionally, the technique identifies which explanatory variables are most informative in identifying these countries.

The remainder of the paper proceeds in four steps. The next section presents background information and data on our empirical application, which examines MNC activity in the chemical products manufacturing industry. As noted above, we are particularly interested in determining whether pharmaceutical manufacturing firms differ from other chemical manufacturing firms, as well as from other MNCs in general, in their subsidiary location, as well as their attempts to use their international subsidiaries to hide intellectual property and shield taxable income from the IRS. In the third section, we describe the nature and use of the maximum entropy technique, which is adapted from the ecology literature, to address the data limitations discussed in the second section. The fourth section presents our empirical results. The paper concludes by discussing the implications of our findings, by discussing the limitations of the empirical methodology, and by suggesting directions for future work in this area.

BACKGROUND AND DATA

The U.S. taxes corporations at a statutory rate of up to 35%.[2] According to the OECD Tax Database, the U.S. corporate tax statutory rates are the highest among developed countries for 2013 (Organisation for Economic Co-operation and Development [OECD], 2013). In recent years, industrialized nations such as United Kingdom, Japan and Canada, among others, have reduced their statutory rates. While there are current proposals on Capitol Hill to overhaul the U.S. tax system, recent grid lock and budget crises leave the authors with little expectation U.S. tax rates will be decrease in the near future.

As a result of high U.S. tax rates, executives at U.S. publicly traded companies are often subject to pressure from shareholders to reduce tax expense. This has led to a booming business for tax advisors. Where a company operates both within and outside the U.S., opportunities arise to shift profits earned beyond the reach of U.S. tax authorities. (Brajcich, Friesner, & McPherson, 2013b). This is particularly true when the primary income producing assets are intangibles such as patents. A patent is much easier to transfer to an overseas subsidiary than brick and mortar factories. Thus, companies operating in the pharmaceutical and technology industries (whose income producing assets are disproportionally patents and other intangible assets) are often presented with a greater opportunity to transfer foreign earnings beyond the reach of the U.S. We are particularly focused on pharmaceutical companies in this study.

The U.S. taxes its corporations on their worldwide income. Regardless of where the income is earned, a U.S. corporation must pay tax to the U.S. Treasury Department on that income.[3] This may be in addition to any income tax paid in a foreign jurisdiction.[4] However, if a U.S. corporation forms a subsidiary corporation in a foreign jurisdiction, U.S. tax law respects that subsidiary as a separate taxpayer. Having U.S. shareholders alone is not generally sufficient to bring the foreign subsidiary within the jurisdiction of U.S. tax authorities. Thus, forming a subsidiary to accumulate foreign earnings allows U.S. corporations to temporarily or altogether avoid U.S. tax on those earnings. If and when the foreign subsidiary pays a dividend to the U.S. parent, it will then be subject to U.S. tax. This deferral of U.S. tax is subject to many limitations in the Internal Revenue Code (IRC), yet U.S.-based MNCs are still able to effectuate the technique in certain circumstances. Avoiding U.S. tax in the current period is often referred to as an interest free loan from the U.S. government. Not only may tax expense be reduced, but in the case of deferral the taxpayer has use of the monies unless and until it is repatriated to the U.S. and taxed. In the instance where hundreds of millions or even billions are at stake with a tax rate as high as 35%, the value of tax savings can be substantial.

The data used in this empirical application was acquired from the IRS Statistics of Income webpage (Internal Revenue Service, 2013). It includes the quantity, assets, receipts, earnings, taxes paid, distributions, subpart F income and related party transactions as reported to the IRS on Form 5471, *Information Return of U.S. Persons With Respect To Certain Foreign Corporations*, as filed during the year 2008. The data represent the cumulative amount reported on all Forms 5471 received by the IRS and are categorized by NAICS Industrial Sector. This information reporting form generally does not require a payment of tax by the filer. Rather, U.S. taxpayers that own a controlling interest in certain foreign corporations are required to report the results of operations and related party transactions of any controlled foreign corporation (CFC) by attaching Form(s) 5471 to their U.S. federal income tax return. Subpart F income represents income earned by the CFCs that despite being earned beyond the jurisdictional arm of U.S. tax authorities is taxed on the U.S. return of the CFC's parent. The Internal Revenue Code defines various categories of subpart F income, as well as other "tainted" income. Generally, these provisions are aimed at limiting U.S. taxpayers ability to defer the payment of U.S. income tax through the U.S. of foreign subsidiaries and are consequently referred to as anti-deferral provisions.[5] The NAICS codes represent industry data presented according to the North American Industry Classification System. All amounts reported by the IRS are in U.S. Dollars. Conversions made from foreign currencies are made on the Form 5471 by the taxpayer in accordance with IRS regulations. The data is also classified by geographic regions as defined by the IRS Statistics of Income office. Not all countries in the world are represented. A list of countries included in each region can be found in the Appendix at Table 6.

EMPIRICAL METHODOLOGY

The techniques used in this manuscript are drawn from ecology and population studies literatures (Eff, Gschwend, & Johnston, 2008; Ruther et al., 2013). In such instances there is a certain amount of information that is known, and collected from a population. The goal is to use that known information to generate estimates about how that information is further disaggregated across sub-populations, for which further sampling is infeasible. Using the example illustrated previously, international finance researchers may know the distribution of MNC subsidiaries by NAICS code, and may know the distribution of MNC subsidiaries by geographic region, but may not know the joint distribution of subsidiaries by both region and industry. In this section, we describe how the concept of maximum entropy can be applied to MNC activity.

Let N denote the total number (or count) of observations in a dataset. In the previous example, N denotes the number of subsidiaries owned by U.S. MNCs. Let i = 1,…,I denote the number of mutually exclusive and collectively exhaustive industries into which subsidiaries can be categorized, and let X_i denote the absolute frequency of MNC subsidiaries in the ith category.[6] It follows that $\phi_i = \frac{X_i}{N}$ denotes the relative frequency (or sample proportion) of MNC subsidiaries in the ith industry classification. Similarly, let j = 1,…,J denote the set of collectively exhaustive and mutually exclusive geographic areas in which MNC subsidiaries are located, and let Y_j, denote the absolute frequency of subsidiaries in region j. Then the sample proportion (or relative frequency) of firms located in region j is given by $\rho_j = \frac{Y_j}{N}$. Lastly, define $Z_{ij} = X_i \cap Y_j$, such that $p_{ij} = \frac{Z_{ij}}{N}$ denotes the joint relative frequency between X_i and Y_j. Then the entire distribution of responses (expressed as absolute frequencies) can be characterized using an IxJ cross-tabulation (or matrix) of the following form:

$$\begin{bmatrix} Z_{11} & \cdots & Z_{1J} & X_1 \\ \vdots & \ddots & \vdots & \\ Z_{I1} & \cdots & Z_{IJ} & X_I \\ Y_1 & & Y_J & N \end{bmatrix} \tag{1}$$

where the elements on the outside of the matrix represent totals, whether row totals (the Xs), column totals (the Ys) or the grand total (N), and the interior elements (the Zs) represent intersections between the two groups. Matrix (1) can equivalently be expressed in relative frequency form by dividing every element in the matrix by the constant N:

$$\begin{bmatrix} p & \cdots & p_{1J} & \phi_1 \\ \vdots & \ddots & \vdots & \\ p_{I1} & \cdots & p_{IJ} & \phi_I \\ \rho_1 & & \rho_J & 1 \end{bmatrix} \tag{1b}$$

In general, information on the Xs, the Ys and N is known, while the Zs and ps are unknown. Specific elements of Z can be considered as "influential observations" and the corresponding p can be considered as an estimate of the likelihood that such influential observations occur. The goal of the analysis is to estimate or impute the interior elements of the cross-tabulation. Ruther et al. (2013) posits the following maximum entropy formulation to estimate the interior cells:

$$\begin{aligned} &\underset{p_{ij}\forall i,j}{maximize} \quad -\textstyle\sum_{i=1}^{I}\sum_{j=1}^{K} p_{ij} log_2(p_{ij}) \\ \text{subject to} \quad &\textstyle\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij} = 1 \\ &N\textstyle\sum_{i=1}^{I} p_{ij} = N\rho_j = Y_j \qquad \forall\, j = 1,..,J \\ &N\textstyle\sum_{j=1}^{J} p_{ij} = N\phi_i = X_i \qquad \forall\, i = 1,..,I \end{aligned} \tag{2}$$

where the objective function ($-\sum_{i=1}^{I}\sum_{j=1}^{K} p_{ij} log_2(p_{ij})$) is known as the "entropy" of the cross-tabulation and $\log_2$ refers to the base 2 logarithm.

The problem formulated in (2) seeks to maximize the entropy of the system, subject to satisfying basic laws of probability (that each of the joint probabilities in a row or column sum to the corresponding row or column, and that the entire sum of joint probabilities is one) and the information that is known about the population of interest (i.e., the marginal distribution of the population according to each variable of interest: X and Y). In the case where one of the probabilities is zero, it is also standard to assume that 0*ln(0) = 0 (Golan, Judge and Miller, 1996; p. 8).

Shannon (1948) demonstrated that the entropy of the system reaches a unique maximum where $p_{ij} = \frac{1}{I+J}\ \ \forall i, j$; that is, where the joint probabilities are uniformly distributed.[7] Within the context of information theory, this implies an assumption of "ignorance" about the distribution of MNC activity across the two variables, X (industry) and Y (geographic region). Put differently, when entropy is at its maximum, there is no prior information to suggest that any one outcome is more or less likely to occur than any other possible outcome, or that a MNC operating in a specific industry is any more or less likely to place a subsidiary in a specific geographic region. The role of the constraints is to introduce information that is known, and constrains or prevents the system from achieving maximum entropy. The more binding the constraints, the further entropy is from its unconstrained maximum, and the more existing information (or less ignorance) there is about the data generating process, which in this case is where MNCs in specific industries are more or less likely to locate their subsidiaries.

The problem formulated in (2) can also be implemented in a number of ways. In its simplest form, the problem in (2) can be treated as a simple spreadsheet modelling exercise, where no statistical foundation for the problem and its solution are assumed.[8] Alternatively, Jaynes (1957, 1982) shows that, in the event of independent and identical random sampling, and as the population size approaches infinity, each joint probability is asymptotically normally distributed with mean δ_i and standard deviation σ. This, in turn, implies that the ratio of two entropies (say, a constrained entropy solution relative to its unconstrained maximum) is proportional to a chi-square distribution.[9] Golan, Judge and Perloff (1996) extend this concept to show how, for individual observations, the joint probabilities (p_{ij}s) can be cast within a discrete choice regression framework, and modeled as a function of one or more exogenous covariates (Q_{ij}, V_{ij}, etc.) and that (under independent and identically distributed sampling) the resulting coefficient estimates are asymptotically equivalent to those of the analogous discrete choice logit model.

There are several advantages to using a maximum entropy approach. First, the approach focuses on estimating the entire distribution of responses, rather than on specific parameters that reflect a measure of variation or central tendency of a distribution. This allows the researcher to take a very general approach to empirical inference (by focusing on the entire distribution of absolute and/or relative frequencies) or it can focus on specific probabilities of interest to the researcher. In this analysis, we are interested in those values for the p_{ij}s that are essentially outliers, or influential observations. Within the context of the empirical application, we seek to determine whether pharmaceutical manufacturing MNCs are different from MNCs in other industries (i.e., whether they are influential observations) in their use of CFCs to amass income and wealth in overseas tax havens. Second, the maximum entropy approach allows much more flexibility in determining what information (both prior information about the p_{ij}s themselves, or on the formation of the p_{ij}s via covariates) is known or unknown at the time of estimation. Lastly, the approach can be used for confirmatory purposes (i.e., estimating probabilities and comparing them to prior expectations or known qualities) or for imputation purposes (i.e. to simply determine the probabilities with very little prior information or data on the p_{ij}s). Since, in the case of MNC subsidiary location by industry, we have little knowledge about the joint distribution of MNC subsidiaries across these two variables, we choose to employ as little prior information as possible and to impute the joint distribution.

In this particular analysis, we adopt the simplest, spreadsheet modelling formulation of the maximum entropy problem (2). That is, we solve the problem in (2) using nonlinear programming methods and do so without any additional assumptions (statistical or otherwise) about the formation of these probabilities.

We do so based on several considerations. One, which is noted above, is that we wish to impute the joint distribution using very limited available information. This limited information essentially precludes the use of advanced regression and/or the imposition of complicated prior distributions on the joint distribution. To account for the relationship between other variables and the location and/or use of MNCs by industry classification, it is possible to repeat this imputation analysis based on other joint distributions. For example, one could impute the joint distribution of dividends paid to the parent corporation by the subsidiary across different MNC industry classifications, and subsequently compare the imputed joint distribution of MNC subsidiary activity by industry code and location to the imputed joint distribution of MNC subsidiary activity by industry code and dividend payments.

Second, and consistent with most IRS-collected data, the total number of MNC subsidiaries (or realizations of any variable based on the activities of these subsidiaries) is finite, and observations in the data are sampled without replacement. Hence, it is unlikely that the data meet the assumptions necessary to establish a formal statistical foundation for the maximum entropy estimates as outlined by Jaynes (1957, 1982). As Friesner, Mittelhammer and Rosenman (2013) note, this would require the imposition of an additional constraint to equation (2) which explicitly incorporates the use of a hypergeometric data generating process. Since this would destroy any parsimony inherent in the model (see point three below), a decision was made to avoid imposing discrete distributional assumptions of this nature. Instead, when multiplying probabilities by the respective sample size, we simply round the results to yield integer values where necessary. This results in joint frequency and probability estimates (whether absolute or relative) that may be measured with error. Such imputed estimates could be used as dependent variables in subsequent regression analyses, but cannot directly be included as explanatory variables without adjusting for this measurement error (Greene, 2000).[10]

Lastly, since the joint frequencies that we wish to impute are most commonly used by accounting and financial practitioners, there is a need to use a parsimonious approach that can be replicated as practical needs dictate. The value of using a simple spreadsheet model is that it is possible to replicate each of our analyses using Microsoft Excel's Solver Add-in Package (Microsoft Corporation, Redmond, WA), which is accessible to the typical accounting and financial professional.

RESULTS

The spreadsheet model used to generate this study's results employs Excel's Solver platform (via its GRG Nonlinear routine), and is available from the lead author upon request. Table 1 contains the imputed joint distribution (both as absolute frequencies/counts and well as relative frequencies/probabilities) for MNC subsidiary activity by location and sub-industry. As noted earlier, a primary consideration is the location of pharmaceutical and medicinal manufacturing-based subsidiary activities relative to other subsidiary activity. Thus, while the entire distribution of imputed estimates will be discussed, pharmaceutical and medicinal manufacturing will be emphasized relative to other industry classifications. When examining Table 1, it is clear that the majority of subsidiaries are located in Europe (42.9 percent), followed by Asia (23.5 percent), Latin America (13.7 percent) and other countries in the western Hemisphere (Canada, etc.; 13.0 percent). Activity in Africa (2.1 percent), Oceania (4.2 percent) and in U.S. territories (1.0 percent) is minimal. Similarly, most MNC subsidiaries operate in the services sector (33.9 percent), followed by distribution and transportation (20.0 percent), the production of miscellaneous goods (16.1 percent), and financial services (12.2 percent), respectively. With regard to pharmaceutical and medicinal manufacturing, we see that slightly less activity (relative to MNC subsidiary activity overall) [11] is taking place in Europe (41.4 percent), Asia (23.0 percent) and in Other Western Hemispheric countries (12.5 percent), and slightly more activity is occurring in Latin America (15.4 percent) and Oceania (5.0 percent).

Table 2 contains the imputed estimates for the dollar value of MNC subsidiary assets by location and industry. As was found in Table 1, the majority of subsidiary assets are located in Europe (62.5 percent), followed by other countries in the Western Hemisphere (Canada, etc.; 19.2 percent), Asia (10.7 percent), Latin America (4.2 percent), Oceania (2.7 percent), Africa (1.0 percent) and in U.S. territories (0.2

percent). When classified by industry, MNC subsidiary assets are placed in the financial services sector (46.4 percent), followed by the services sector (28.1 percent), the production of miscellaneous goods (7.3 percent), distribution and transportation (6.1 percent) and raw materials and energy production (3.2 percent). Pharmaceutical and chemical manufacturing comprises only 1.7 percent of subsidiary assets. The distribution of pharmaceutical and medicinal manufacturing assets (relative to all MNC subsidiary activity) are less likely to be placed in Europe (54.7 percent) and in Other Western Hemispheric countries (9.1 percent). Relatively more assets are being placed in Asian (21.0 percent) and Latin American (11.2 percent) subsidiaries.

Estimates of MNC subsidiary income (Current Earnings and Profit after Taxes, or CEPAT) are presented in Table 3. Subsidiary income was mostly earned by European subsidiaries (53.9 percent), followed by other countries in the Western Hemisphere (Canada, etc.; 25.2 percent), Asia (10.0 percent), Latin America (5.7 percent), Oceania (3.5 percent), U.S. territories (1.0 percent) and Africa (1.0 percent). MNC subsidiaries operating in the services (37.8 percent) and financial (18.1 percent) sectors accounted for most of the subsidiary income. Pharmaceutical and chemical manufacturing comprises only 4.2 percent of subsidiary income. The distribution of pharmaceutical and medicinal manufacturing subsidiary incomes (relative to the entire collection of MNC subsidiaries) are much more likely to come from European MNCs (71.0 percent) and Asian subsidiaries (15.4 percent), and much less likely to come from subsidiaries in Other Western Hemispheric countries (5.0 percent).

Table 4 examines the distribution of dividends paid by subsidiaries to their parent corporation(s), disaggregated by location and industry. For MNCs taken cumulatively, most dividends are being paid by European subsidiaries (40.0 percent), followed closely by subsidiaries in Other Western Hemisphere countries (38.0 percent). Remaining percentages include Asian MNCs (8.3 percent), those from Lain America (8.1 percent), Oceania (3.1 percent), Africa (2.3 percent) and U.S. territories (0.2 percent). But when examining (and conditioning upon) only MNC subsidiary activity in the pharmaceutical and medicinal manufacturing sector, a very different pattern emerges. European subsidiaries continue to be the largest source of dividends at 38.8 percent. However, Asian subsidiaries paid 34.5 percent of the dividends in this sector. Latin American MNCs paid 13.9 percent of dividends, while Other Western Hemisphere subsidiaries paid only 9.3 percent of the dividends in this sector. Remaining dividends flowed from Oceania (2.2 percent), Africa (1.0 percent) and U.S. territories (0.3 percent). Clearly, pharmaceutical manufacturers are much more reliant on their Asian subsidiaries to generate and repatriate earnings from their subsidiaries. This is consistent with recent literature which suggests that pharmaceutical manufacturers are moving many of their most lucrative operations overseas to capture lower drug development and approval costs (Masri, Ramirez, Popescu, & Reggie, 2012; Fiedler, Bebber, & Oetjen, 2013; Yadav, 2013).

Table 5 combines information from Tables 1-4 to examine assets, income and dividends paid on a per MNC subsidiary basis. For the purposes of simplicity, we aggregate all non-pharmaceutical/medicinal sectors into a single sector, to more compactly gauge differences between the “influential observations” in the pharmaceutical/medicinal manufacturing industry versus all other MNC subsidiary activity. Panel A examines the magnitude of assets per MNC subsidiary, disaggregated by location and industry. Pharmaceutical subsidiaries have 5-6 times the value of assets per subsidiary in Latin American countries than do other types of firms. They also have more than four times as many assets per subsidiaries for African operations, more than four times as many assets per subsidiary in Asian locations, and nearly twice as many assets per subsidiary in European, Oceanic and U.S. territory-based subsidiaries compared to other types of industries. The magnitude of assets per subsidiary in Other Western Hemisphere countries is comparable to those of other industries. Clearly, these results indicate that pharmaceutical and medicinal manufacturing firms tend to concentrate a larger amount of firm wealth in a smaller number of overseas subsidiaries than other types of firms.

Panel B examines CEPAT per MNC subsidiary, disaggregated by industry and location. Pharmaceutical companies which locate facilities in Asia and Europe generate income streams that are eight to nine times as large (per MNC subsidiary) than those of other industries combined. Pharmaceutical MNCs located in Latin American countries generate earnings that are nearly six times as large as their counterparts in other industries, while in Oceania pharmaceutical subsidiaries generate more earnings than other firms by a factor of 1.7 to 1. Lastly, pharmaceutical subsidiaries located in other countries generate returns that are relatively comparable to non-pharmaceutical manufacturing firms.

The third and final panel (Panel C) in Table 5 disaggregates dividends repatriated (or paid) to the parent corporation by location and industry. Once again, the amount of dividends paid per MNC subsidiary is larger for pharmaceutical manufacturers than for non-pharmaceutical manufactures, particularly in Latin America, Europe, Asia and Oceania. However, the discrepancies in repatriated earnings by pharmaceutical and medicinal manufacturers relative to other types of firms are not nearly as large as the discrepancy in earnings per subsidiary. Pharmaceutical manufacturers whose subsidiaries are based in Europe repatriate about three times as many dividend dollars per subsidiary as other times of firms. In Latin America, pharmaceutical manufacturers repatriate approximately five times as many dividend earnings as other firms. In U.S. territories, the ratio is 4:1, while in Oceania, the ratio is approximately 2:1. The one exception is in Asian countries, where pharmaceutical subsidiaries repatriate nearly 15 times as many earnings per subsidiary as non-pharmaceutical firms. With the exception of the subsidiaries located in Asia, the relatively large pharmaceutical manufacturer earnings per subsidiary, coupled with lower rates of repatriation, suggests (but does not prove) that pharmaceutical companies are shielding larger amounts of corporate income and assets abroad relative to other types of firms. The two exceptions are in Asia and in U.S. territories, where the repatriation flows by pharmaceutical manufacturers appear to outpace MNC subsidiaries in other industries.

CONCLUSIONS

In this paper, we present a simple model that can be used to impute or estimate the impacts of "influential observations" when such data are unavailable, or when standard methods of estimation do not shed significant light on these observations. The methodology is robust, in that, depending on the information available to the researcher, it may have a traditional statistical foundation, or may be used as a non-statistical spreadsheet modeling tool. In either case, it is straightforward to employ the technique using simple data analysis tools such as Excel.

As an application, we used data taken from IRS data in 2008 (Form 5471, *Information Return of U.S. Persons With Respect To Certain Foreign Corporations*) to identify those countries in which MNC are establishing corporate subsidiaries to defer or avoid paying U.S. federal tax. We were particularly interested in determining whether pharmaceutical manufacturing firms differ from other chemical manufacturing firms in their subsidiary location, as well as their attempts to use their international subsidiaries to hide intellectual property and shield taxable income from the IRS. We find evidence that such differential incentives exist. Pharmaceutical and medicinal manufacturers are more likely than other types of MNC subsidiaries to establish a small number of subsidiaries in Latin American and Oceania. Perhaps more importantly, we find evidence that pharmaceutical manufacturers are likely to place a much larger amount of assets in a typical subsidiary than other types of subsidiaries. These subsidiaries, as a general rule, also earn greater income per subsidiary and repatriate less per subsidiary. Thus, pharmaceutical manufacturers are slowly accumulating substantial monetary and non-monetary resources in these subsidiaries. As long as those resources remain held abroad (especially in countries where tax rates are lower than the U.S. corporate tax rate, which is true for most developing countries) they provide the parent corporation with a tax haven.

While our application demonstrates the utility of the maximum entropy modeling approach, and provides some interesting policy inferences, we must emphasize that our results should be viewed with a degree of caution. We have assumed very little about the data generating process which characterizes the formation of the MNC data that was reported by the IRS. We have also failed to include information about control variables other than subsidiary location and industry. It may be the case that, by including other salient information into the model, other explanations (such as corporate tax rate differentials, differences in pharmaceutical manufacturing and regulatory process, access to resources and labor markets, etc.) might account for these differences. Future research is necessary to include these missing factors and either verify or refute the preliminary inferences found in this manuscript.

ENDNOTES

1. Such examples are not limited to applications in international finance. For example, hospital administrators are interested in identifying those (relatively few) Medicare or Medicaid patients who are most at risk of acquiring secondary infections at the hospital, and ultimately being readmitted to the hospital at a later date, in which case the hospital could be assessed a financial penalty by the patient's insurer (McHugh, Berez, & Small, 2013). In community development, there is an interest in determining how different characteristics of particular neighborhood (say, household income and spending habits) influence those of the community as a whole (Eff, Gschwend, & Johnston, 2008; Ruther, Maclaurin, Leyk, Buttenfield, & Nagle, 2013).
2. It is important to note that while the U.S. statutory rate is among the highest in the industrialized world, the U.S. effective tax rate is significantly lower. According to the Government Accounting Office, U.S. corporations paid an average effective tax rate of 12.6% in 2010 (Government Accounting Office, 2013). This is in part the result of some of the international tax planning techniques discussed in this paper.
3. The fact that a corporation is a taxable entity in the U.S. has given rise to policy arguments against the double taxation of corporate income. Corporations are currently taxed once when they earn income and shareholders are again taxed when that income is distributed in the form of a dividend.
4. This double taxation, once in the U.S. and once by a foreign jurisdiction, may be alleviated to some degree by the foreign tax credit.
5. Additional limitations are found in transfer pricing rules promulgated by the U.S. Treasury Department.
6. In what follows, we refer generically to X as a variable, and Xi, as a particular set of realizations of data (absolute frequencies) for variable X. Similar language will be used to define other variables.
7. In the absence of any other constraints and data that arise from a well-defined random sampling process, the solution to the maximum entropy problem should identify joint probabilities that are the products of the associated marginal probabilities in the contingency table (Good, 1963). As constraints are added to the problem, or as the data do not follow the statistical assumptions underlying Good's solution, the joint probabilities may deviate from the product of the marginal probabilities.
8. Because the entropy objective has a unique maximum and the formulation of the constrained optimization problem described in (2) is relatively simple, it is possible to solve (2) analytically (rather than numerically) and apply data to that solution to identify the optimal matrix of ps (Golan, Judge and Miller, 1996; pp. 10-11). This would be an advantageous approach in the absence of an efficient and effective nonlinear programming algorithm or when unreliable data or infeasible constraints prevent an effective and efficient algorithm from converging to a unique solution. The analytical approach, however, becomes less feasible as additional constraints are added to the problem, and/or as the matrix (1b) becomes larger. We use the numerical approach to solving (2) on the grounds that practitioners are familiar with spreadsheet modelling tools, and are much more likely to employ the numerical approach in practice. Again, given a well-defined problem and adequate data, both approaches should identify the same solution.
9. In the absence of well-defined data generating processes, one could adopt of prior distribution of ignorance (or whether continuous or discrete) and the resulting entropy solution would carry a posterior distribution that reflects the underlying prior. In such cases, (Bayesian) credible regions can be constructed for the resulting estimates.
10. It might also be argued that, as the population size approaches infinity, or as long as the sample size is much, much less than the population size, that the approximations produced by this spreadsheet model will be reasonable approximations of the underlying joint probabilities. These results may be true in theory, but in practice rarely occur. Hence, this analysis chooses not to make such assumptions as a general rule. See Friesner, Mittelhammer and Rosenman (2013) for more details.
11. When computing all percentages for the pharmaceutical and medicinal manufacturing sector (and to generate percentages that can be compared to those for the sample as a whole), we are creating conditional probabilities by taking each joint relative frequency, dividing that frequency by the total frequency for the pharmaceutical and chemical manufacturing sector, and converting the resulting proportion to a percentage.

REFERENCES

Brajcich, D., Friesner, D., & McPherson, M. (2013a). Key determinants of repatriated earnings by U.S. multinational enterprises. *Multinational Business Review*, 21(3), 269-289.

Brajcich, D., Friesner, D., & McPherson, M. (2013b). Trends in the Shifting of Resources by U.S.-based Multinational Companies. *Journal of Accounting and Finance*, 13(6), 92-106.

Eff, M., Gschwend, T., & Johnston, R. (2008). Ignoramus, ignorabimus? The principle of maximum entropy and ecological inference for RxC tables. *Political Analysis*, 16, 70-92.

Fiedler, B., Bebber, R., & Oetjen, R. (2013). A regulatory clinical trial primer: Distribution challenges and policy implications of modern vaccine development. *International Journal of Pharmaceutical and Healthcare Marketing*, 7(1), 45-57.

Friesner, D., Mittelhammer, R., & Rosenman, R. (2013). Inferring the incidence of industry inefficiency from DEA estimates. *European Journal of Operational Research,* 224(2), 414-424.

Golan, A., Judge, G., & Miller, D. (1996). *Maximum entropy econometrics: Robust estimation with limited data.* New York, NY: John Wiley and Sons.

Golan, A., Judge, G., & Perloff, J. (1996). A maximum entropy approach to recovering information from multinomial response data. *Journal of the American Statistical Association*, 91(434), 841-853.

Good, I. (1963). Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables. *The Annals of Mathematical Statistics,* 34(3), 911-934.

Government Accounting Office. (2013). *Effective Tax Rates Can Differ Significantly from the Statutory Rate* retrieved from http://www.gao.gov/assets/660/654957.pdf.

Greene, W. (2000). *Econometric analysis, 4th Edition*. Prentice Hall, Upper Saddle River, NJ.

Internal Revenue Service. (2013). *SOI Tax Stats - Controlled Foreign Corporations* retrieved from http://www.irs.gov/uac/SOI-Tax-Stats-Controlled-Foreign-Corporations.

Jaynes, E. (1957). Information theory and statistical mechanics. *Physics Review*, 106(4), 620-630.

Jaynes, E. (1982). On the rationale of maximum-entropy methods. *Proceedings of the IEEE,* 70(9): 939-952.

Masri, M., Ramirez, B., Popescu, C., & Reggie, E. (2012). Contract research organizations: An industry analysis. *International Journal of Pharmaceutical and Healthcare Marketing,* 6(4): 336-350.

McHugh, D., Berez, J., & Small, D. (2013). Hospitals with higher nurse staffing had lower odds of readmissions penalties than hospitals with lower staffing. *Health Affairs*, 32(10): 1740-1747.

Organization for Economic Cooperation and Development. (2013). *OECD Tax Database* retrieved from http://www.oecd.org/tax/tax-policy/tax-database.htm#C_CorporateCaptial.

Ruther, M., Maclaurin, G., Leyk, S., Buttenfield, B., & Nagle, N. (2013). Validation of spatially allocated small area estimates for 1880 Census demography. *Demographic Research, 29*(Article 22): 579-616.

Shannon, C., 1948, “A Mathematical Theory of Communication,” *Bell System Technical Journal*, 27, 379,423.

Yadav, S. (2013). Building regulatory capabilities for pharmaceutical firms’ internationalization. *International Journal of Pharmaceutical and Healthcare Marketing, 7*(1): 58-74.

Table 1: Imputed MNC Subsidiary Location by Region and Sub-Industry

Panel A: Imputed Relative Frequencies

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Raw Materials and Energy Production	0.004	0.004	0.013	0.001	0.007	0.002	<0.001	0.031
Pharmaceutical and Medicinal Manufacturing	0.001	0.001	0.003	<0.001	0.002	<0.001	<0.001	0.008
Non-Pharmaceutical/Medicinal Chemical Manufacturing	0.005	0.004	0.013	0.001	0.007	0.002	<0.001	0.031
Computer and Electronic Product Manufacturing	0.004	0.004	0.012	0.001	0.006	0.001	<0.001	0.028
Transportation Equipment Manufacturing	0.003	0.002	0.008	<0.001	0.004	0.001	<0.001	0.018
All Other Manufacturing Goods	0.026	0.021	0.071	0.003	0.039	<0.001	0.001	0.161
Distribution and Transportation of Goods	0.031	0.025	0.083	0.004	0.046	0.010	0.001	0.200
Information	0.009	0.008	0.025	0.001	0.014	0.003	<0.001	0.061
Finance, Insurance, Real Estate, and Rental/Leasing	0.000	0.018	0.059	0.003	0.033	0.007	0.001	0.122
Services	0.053	0.043	0.142	0.007	0.076	0.017	0.002	0.339
All Other Goods/Services	<0.001	<0.001	<0.001	<0.001	0.000	<0.001	<0.001	0.001
Total	0.137	0.130	0.429	0.021	0.235	0.042	0.007	1.000
Estimated Entropy	398585.711							

Panel B: Imputed Absolute Frequencies

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Raw Materials and Energy Production	363	328	1077	54	604	129	18	2572
Pharmaceutical and Medicinal Manufacturing	101	82	272	14	151	33	4	657
Non-Pharmaceutical/Medicinal Chemical Manufacturing	396	321	1062	53	589	126	17	2564
Computer and Electronic Product Manufacturing	365	296	978	49	543	116	16	2363
Transportation Equipment Manufacturing	237	193	635	32	352	76	12	1538
All Other Manufacturing Goods	2184	1782	5916	279	3257	0	86	13503
Distribution and Transportation of Goods	2580	2095	6956	346	3842	829	110	16758
Information	786	642	2096	105	1156	248	33	5066
Finance, Insurance, Real Estate, and Rental/Leasing	0	1503	4974	247	2792	598	78	10192
Services	4406	3595	11859	544	6376	1392	186	28359
All Other Goods/Services	11	8	32	1	17	3	0	72
Total	11429	10845	35857	1724	19679	3550	560	83642

Note: All frequencies were rounded, and values less than the last significant digit are reported in discrete terms.

Table 2: Imputed MNC Subsidiary Year-End Assets by Region and Sub-Industry

Panel A: Imputed Relative Frequencies

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Raw Materials and Energy Production	0.003	0.004	0.016	0.001	0.006	0.001	<0.001	0.032
Pharmaceutical and Medicinal Manufacturing	0.002	0.002	0.009	<0.001	0.004	<0.001	<0.001	0.017
Non-Pharmaceutical/Medicinal Chemical Manufacturing	0.002	0.003	0.006	0.001	0.003	0.001	<0.001	0.017
Computer and Electronic Product Manufacturing	0.002	0.003	0.006	<0.001	0.003	0.001	<0.001	0.016
Transportation Equipment Manufacturing	0.002	0.002	0.008	<0.001	0.003	0.001	<0.001	0.017
All Other Manufacturing Goods	0.002	0.026	0.041	<0.001	0.001	0.002	<0.001	0.073
Distribution and Transportation of Goods	0.001	0.032	0.020	<0.001	0.001	0.006	<0.001	0.061
Information	<0.001	0.007	0.011	0.001	0.002	0.002	<0.001	0.023
Finance, Insurance, Real Estate, and Rental/Leasing	0.027	0.047	0.299	0.002	0.080	0.009	0.001	0.464
Services	<0.001	0.066	0.207	<0.001	0.005	0.003	<0.001	0.281
All Other Goods/Services	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Total	0.042	0.192	0.625	0.005	0.107	0.027	0.002	1.000
Estimated Entropy	53,616,888,486,621							

Panel B: Imputed Absolute Frequencies in Tens of Billions of U.S. Dollars

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Raw Materials and Energy Production	4	6	23	1	9	2	< 1	46
Pharmaceutical and Medicinal Manufacturing	3	2	13	< 1	5	1	< 1	25
Non-Pharmaceutical/Medicinal Chemical Manufacturing	3	5	9	1	4	2	< 1	24
Computer and Electronic Product Manufacturing	3	5	9	1	4	2	< 1	24
Transportation Equipment Manufacturing	3	3	11	1	5	1	< 1	25
All Other Manufacturing Goods	3	38	60	< 1	1	3	1	106
Distribution and Transportation of Goods	2	46	30	< 1	2	8	1	88
Information	< 1	10	16	1	3	3	< 1	34
Finance,Iinsurance, Real Estate, and Rental/Leasing	40	68	435	2	117	13	1	675
Services	< 1	96	302	< 1	7	4	< 1	409
All Other Goods/Services	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Total	60	279	909	7	156	39	3	1454

Note: All frequencies were rounded, and values less than the last significant digit are reported in discrete terms.

Table 3: Imputed MNC Subsidiary Current Earnings and Profits after Taxes (CEPAT) by Region and Sub-Industry

Panel A: Imputed Relative Frequencies

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Raw Materials and Energy Production	0.004	0.007	0.058	0.001	0.011	0.002	<0.001	0.084
Pharmaceutical and Medicinal Manufacturing	0.003	0.002	0.030	<0.001	0.007	0.001	<0.001	0.042
Non-Pharmaceutical/Medicinal Chemical Manufacturing	0.002	0.004	0.008	0.001	0.003	0.001	<0.001	0.020
Computer and Electronic Product Manufacturing	0.003	0.006	0.018	0.001	0.006	0.002	<0.001	0.036
Transportation Equipment Manufacturing	0.002	0.002	0.005	<0.001	0.003	0.001	<0.001	0.014
All Other Manufacturing Goods	0.002	0.040	0.068	<0.001	0.001	0.002	0.002	0.115
Distribution and Transportation of Goods	0.001	0.054	0.029	<0.001	0.001	0.008	0.002	0.095
Information	<0.001	0.010	0.017	0.001	0.002	0.003	0.001	0.034
Finance, Insurance, Real Estate, and Rental/Leasing	0.025	0.050	0.027	0.003	0.062	0.012	0.003	0.181
Services	0.015	0.077	0.278	<0.001	0.005	0.003	<0.001	0.378
All Other Goods/Services	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Total	0.057	0.252	0.539	0.008	0.100	0.035	0.009	1.000

Estimated Entropy 3,113,699,024,409

Panel B: Imputed Absolute Frequencies in Billions of U.S. Dollars

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Raw Materials and Energy Production	3	5	42	1	8	1	< 1	61
Pharmaceutical and Medicinal Manufacturing	2	2	22	< 1	5	< 1	< 1	31
Non-Pharmaceutical/Medicinal Chemical Manufacturing	1	3	6	1	3	1	< 1	15
Computer and Electronic Product Manufacturing	2	5	13	1	4	1	< 1	26
Transportation Equipment Manufacturing	1	2	4	< 1	2	1	< 1	10
All Other Manufacturing Goods	1	30	50	< 1	< 1	2	1	85
Distribution and Transportation of Goods	1	39	21	< 1	1	6	1	70
Information	0	7	13	1	1	2	1	25
Finance, Insurance, Real Estate, and Rental/Leasing	18	37	20	3	46	9	2	133
Services	11	56	204	< 1	3	2	< 1	278
All Other Goods/Services	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Total	42	185	396	6	74	25	7	734

Note: All frequencies were rounded, and values less than the last significant digit are reported in discrete terms.

Table 4: Imputed MNC Subsidiary Dividends Paid to the Parent Corporation by Region and Sub-Industry

Panel A: Imputed Relative Frequencies

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Raw Materials and Energy Production	0.007	0.011	0.111	0.002	0.024	0.002	<0.001	0.158
Pharmaceutical and Medicinal Manufacturing	0.003	0.002	0.010	<0.001	0.008	0.001	<0.001	0.025
Non-Pharmaceutical/Medicinal Chemical Manufacturing	0.003	0.006	0.012	0.001	0.005	0.002	<0.001	0.030
Computer and Electronic Product Manufacturing	0.004	0.008	0.018	0.001	0.009	0.002	<0.001	0.043
Transportation Equipment Manufacturing	0.003	0.003	0.010	0.001	0.005	0.001	<0.001	0.023
All Other Manufacturing Goods	0.002	0.047	0.029	0.001	0.001	0.002	<0.001	0.082
Distribution and Transportation of Goods	0.002	0.036	0.019	<0.001	0.001	0.009	<0.001	0.067
Information	<0.001	0.004	<0.001	0.003	0.002	0.003	<0.001	0.012
Finance, Insurance, Real Estate, and Rental/Leasing	0.017	0.019	0.009	0.014	0.019	0.005	<0.001	0.084
Services	0.039	0.243	0.183	<0.001	0.007	0.004	<0.001	0.476
All Other Goods/Services	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.000
Total	0.081	0.380	0.400	0.023	0.083	0.031	0.002	1.000
Estimated Entropy	398,490,505,095							

Panel B: Imputed Absolute Frequencies in Hundreds of Millions of U.S. Dollars

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Raw Materials and Energy Production	7	11	107	1	24	2	< 1	153
Pharmaceutical and Medicinal Manufacturing	3	2	9	< 1	8	1	< 1	24
Non-Pharmaceutical/Medicinal Chemical Manufacturing	3	6	12	1	5	2	< 1	29
Computer and Electronic Product Manufacturing	4	8	17	1	9	2	< 1	41
Transportation Equipment Manufacturing	3	3	10	1	5	1	< 1	23
All Other Manufacturing Goods	2	45	28	1	1	2	< 1	79
Distribution and Transportation of Goods	2	35	18	< 1	1	9	< 1	65
Information	< 1	4	< 1	3	2	2	< 1	12
Finance, Insurance, Real Estate, and Rental/Leasing	17	18	8	14	19	5	< 1	81
Services	37	235	178	< 1	7	4	< 1	461
All Other Goods/Services	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Total	78	367	387	22	80	30	2	967

Note: All frequencies were rounded, and values less than the last significant digit are reported in discrete terms.

Table 5: Imputed Characteristics per MNC Subsidiary by Region and Industry

Panel A: Imputed Absolute Frequencies, Assets per MNC

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Pharmaceutical and Medicine Manufacturing	275,053,005	274,321,660	494,601,917	181,844,159	338,824,759	207,300,240	174,164,567	374,508,941
All Other Industries	50,846,130	257,133,826	251,652,516	41,351,863	77,239,161	110,158,886	60,400,669	172,303,678
Total	52,827,484	257,263,784	253,495,454	42,492,752	79,246,348	111,061,890	61,213,268	173,891,981

Panel B: Imputed Absolute Frequencies, CEPAT per MNC

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Pharmaceutical and Medicine Manufacturing	20,298,223	19,132,413	81,133,476	10,793,465	31,791,454	11,888,424	11,909,226	47,304,196
All Other Industries	3,538,459	17,026,481	10,497,315	3,504,337	3,528,580	7,093,328	11,633,426	8,473,825
Total	3,686,568	17,042,404	11,033,139	3,563,530	3,745,445	7,137,902	11,635,396	8,778,834

Panel C: Imputed Absolute Frequencies, Dividends per MNC

Industry	Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions	Total
Pharmaceutical and Medicine Manufacturing	3,267,137	2,689,211	3,383,434	1,719,149	5,416,758	1,610,125	1,522,866	3,610,372
All Other Industries	661,331	3,393,956	1,062,437	1,296,506	368,149	830,107	389,299	1,137,190
Total	684,359	3,388,628	1,080,043	1,299,938	406,888	837,358	397,396	1,156,616

Table 6: Countries by Region

Latin America	Other Western Hemisphere	Europe	Africa	Asia	Oceania	Puerto Rico and U.S. Possessions
Mexico	Bahamas	Austria	Egypt	Kazakhstan	Australia	Puerto Rico
Costa Rica	Barbados	Belgium	Morocco	Turkey	New Zealand	Virgin Islands, U.S.
Dominican Republic	Bermuda	Bulgaria	Other North African countries	Other Central, Northern, and Southwestern Asian	Other countries of Oceania	Other U.S. Possessions
El Salvador	British Virgin Islands	Cyprus	Mauritius	Israel		
Guatemala	Canada	Czech Republic	Other East African countries	Saudi Arabia		
Honduras	Cayman Islands	Denmark	Liberia	United Arab Emirates		
Panama (including Canal Zone)	Jamaica	Estonia	Nigeria	Other Middle East countries		
Other Central American and Caribbean countries	Netherland Antilles	Finland	Other West and Central African countries	India		
Argentina	Trinidad and Tobago	France (including Corsica, Guadeloupe, Martinique, and Reunion)	South Africa	Indonesia		
Bolivia	Other Western Hemisphere countries	Germany	Other Southern African countries	Malaysia		
Brazil		Greece		Pakistan		
Chile		Hungary		Philippines		
Colombia		Ireland		Singapore		
Ecuador		Italy		Thailand		
Peru		Latvia		Vietnam		
Uruguay		Lithuania		Other Southern and Southeastern Asian countries		
Venezuela		Luxembourg		China		
Other South American countries		Malta		Hong Kong		
		Netherlands		Japan (including Okinawa and Ryukyu Islands)		
		Poland		South Korea		
		Portugal (including Azores)		Taiwan		
		Romania		Other Eastern Asian countries		
		Slovakia				
		Slovenia				
		Spain (including Canary Isles)				
		Sweden				
		United Kingdom and Northern Ireland (including Gibraltar)				
		Croatia				
		Guernsey				
		Jersey				
		Norway				
		Russia				
		Serbia				
		Switzerland				
		Ukraine				
		Other European countries				