Segev, E. (2014). Visible and Invisible Countries: News-Flow Theory Revised. Journalism. Available online at: http://jou.sagepub.com/content/early/2014/03/03/1464884914521579

Visible and Invisible Countries: News-Flow Theory Revised

Elad Segev

Tel Aviv University

Abstract

In this paper the validity and limits of the news-flow theory is examined, utilizing a large digital corpus of 35 popular news sites in 10 different languages over a three-year period. Three key variables were identified: GDP, Foreign Population and Conflict Intensity, counting together for more than 70% of the variance of country prominence in the news. After offering a robust model, over-represented and under-represented countries are listed and divided based on their different characteristics and news trends.

Findings show that conflicts tend to be visible only if they are in the Middle East. Over-represented regional centers in West Europe and Asia overshadow the under-represented regional peripheries in East Europe and the Middle East. Finally, the US and Africa serve as a global center and a periphery respectively. The implications of these findings are discussed in order to invite scholar to further expand the theory and the explanatory model accordingly.

Keywords: news-flow theory, news sites, country prominence, over- and under-representation

Visible and Invisible Countries: News-Flow Theory Revised

What makes a country newsworthy? Over the years scholars have studied the relationship between the news prominence of a country and its physical, economic, political, social and cultural characteristics, such as its size or economic power. It has been widely accepted that the international news in different countries and languages is fundamentally different (Kayser, 1953; Gerbner and Marvanyi, 1977; Segev, 2010; Segev and Hills, 2013; Wilke, Heimprecht, and Cohen, 2013), reflecting local interests and agendas. It has also been argued that the presence or absence of countries in the news has further implications on decision-making (Hawkins, 2008), and on the general public opinion (Wanta, Golan, and Cheolhan, 2004).

While previous studies largely agree on the most important determinants of country news worthiness, this paper goes a step further by looking at the outliers, in other words, countries that are over-represented or under-represented in the news when explained by traditional factors documented by the literature. Advanced web-mining techniques are employed in order to analyse a very large corpus of news sites from around the world, identify and classify countries that do not fall into traditional categories, and suggest directions to expand the news-flow theory accordingly.

The current news flow theory

The different ways in which international information was reported and disseminated worldwide have attracted academic attention already in the 1950s (International Press Institute, 1953; UNESCO, 1954). Theories attempting to explain the news-worthiness of certain issues and actors soon emerged (Galtung and Ruge, 1965; Ostgaard, 1965). These theories further stimulated a growing body of empirical research in order to identify the main determinants of the international news flow (Robinson and Sparkes, 1976; Rosengren, 1970).

Today it is widely accepted that the prominence of a foreign country in the news is attributed to three groups of variables: (a) *national traits* (e.g., the size and power of the foreign country), (b) *relatedness*, namely proximity to that foreign country in terms of geography, demography, etc., and (c) *events* (e.g., disasters, wars, conflict, local protest) (Sheafer et al. 2013; Wu, 2000). While national trait variables can explain the similar focus of news from different countries on the core countries such as the US and Western European countries, relatedness variables explain their different regional focus. Event oriented variables can explain both the regional and global focus based on the extent of the event and its relevance to the reporting country.

National Traits: Variables of this group often measure the size of a country and its economic and political power. Following Wallerstein's (1974) World System Theory (WST), Chang (1998) classified states as core, semi-peripheral, or peripheral, and found that core-countries get much greater news attention than

semi-peripheral and peripheral countries. Particularly, the economic power of a country was found to be a strong indicator of its news prominence (Ishii, 1996; Kim and Barnett, 1996; Robinson and Sparkes, 1984; Wu, 2000, 2007). Military power is another important indicator for locating countries on that divide (Kariel and Rosenvall, 1984; Shenhav et al., 2012), followed by population size (Charles et al., 1979; Dupree, 1971; Rosengren, 1977).

Relatedness. The second group of variables measures the economic, political, social and cultural ties between a reporting and a reported country. It was found, for example, that bilateral trade is a strong predictor of mutual newsworthiness of two countries (Charles et al., 1979; Kariel and Rosenvall, 1984; Rosengren, 1977, Wu, 2000). Geographic proximity (Dupree, 1971; Galtung and Ruge, 1965) and cultural proximity, which often refers to ethnic similarity (Shoemaker et al., 1991), as well as immigration, travel, and shared languages (Chang et al. 1987; Kariel and Rosenvall, 1984), were also found to be determinants of international news coverage.

Events: The third group of variables measures the deviance of a country, such as involvement in conflicts (Golan and Wanta, 2003) as well as the political, economic and social changes it is experiencing (Chang et al. 1987; Shoemaker et al. 1986, 1991). The recent uprisings in the Middle East are a good example in this regard, demonstrating that smaller and less powerful countries such as Tunisia and Syria can get very high news attention worldwide, at least for a time,

due to their provision of outstanding news events. The tsunami in Japan and the earthquake in Haiti underscore the same point.

In order to explain the news prominence of countries based on variables from some of the groups mentioned above (hereafter: "news flow variables"), Wu (2000) studied the international news section of newspapers from 38 countries. He used multiple regression models for each country individually, and found that trade volume with the reporting country and the number international news agencies are the strongest predictors of countries' news prominence. Additionally, when combining the data from newspapers of all 38 countries he found that the US was by far the most prominent country, followed by France, the UK, Russia and China. In other words, it was found that the economic power of a country is a particularly strong predictor of its news prominence.

As a first stage, this paper systematically surveys the most significant news flow variables. In line with theoretical arguments and empirical findings of previous studies (Galtung and Ruge, 1965; Ishii, 1996; Kariel and Rosenvall, 1984; Kim and Barnett, 1996; Robinson and Sparkes, 1984; Segev and Blondheim, 2013a; Wu, 2000, 2007) the first hypothesis is that:

> H1: Out of all news flow variables, the GDP of a country would have the strongest correlation with its news prominence.

Yet, previous studies suggest other important factors. In particular, apart from national traits, relatedness and event-oriented variables are also associated with

the news prominence of countries. The second stage is therefore to confirm the current news-flow theory by finding the best-fit model that explains most of the news prominence of countries. While looking at the best-fit model for each country separately, Wu (2000) found that his news-flow variables could explain between 17.2% (US news) and 36% (Iranian news) of the variance of country prominence. In a later study, comparing online and traditional media[1] the US alone, Wu (2007) could explain between 38% (online media) and 58.6 (traditional media) of the variance of country prominence. While Wu's studies examined the best fit-model for each country separately, this study examines the prominence of a country in the news of all other countries together. Additionally, Wu's study does not count in his model variables from the event group. It is therefore expected that the model in the present study would be more comprehensive and explain more of the news prominence variance.

H2: News flow variables would explain together at least 50% of the variance of countries' news prominence.

Once the best-fit model is found, the final and most important stage is to identify the outliers—countries that are over-represented and under-represented. It is believed that countries experiencing sudden incidents such as natural disasters or social protests would get much higher news attention than expected if solely explained by news flow variable. One of the methodological reasons for this is

that the translation of these unexpected events into measurable variables is not always possible (see Shoemaker et al. 1986, 1991 and Pritchard and Hughes, 2006 for some examples for the operationalization of deviance in the news). Hence, it is expected that:

> H3: The news prominence of countries experiencing unexpected events would not be explained by the news flow variables.

The outlying countries are classified based on the explanatory power of each group of variables. This enables finding patterns among the over- and under-represented countries in the news, and thus also shedding light on the limitations of the current news-flow literature and the possible directions to expand and develop it further.

Methodology

News Collection and Analysis

The data analyzed in this study were collected from a variety of news sites in 11 different countries (the US, the UK, Germany, France, Spain, Russia, China, Japan, Iran, Egypt and Israel). These countries were selected on the basis of several considerations. First, countries with a large number of online users were chosen, since they often act as cultural and media centers for the smaller countries in their peripheries (Tunstall, 2008). Applied to the internet, this factor must reflect the popularity of the language used in the country (i.e., the most popular online languages such as English, Chinese, Spanish, Japanese, French, German and Arabic were preferred).

Second, economically leading countries were chosen. As the previous studies surveyed above suggest, a combination of economic, political and cultural factors contribute to their news prominence. Hence, analysis of this kind should particularly include countries with high GDPs such as Japan, China, Germany, the UK and France. Finally, when addressing the question of news prominence, previous studies pointed on the particularly high attention given to some Middle Eastern countries (Segev and Blondheim, 2013b). News sites from Iran, Egypt and Israel, were therefore included, as they represent prominent geopolitical centers of that region.

In each of the countries selected for analysis, three popular news sites were chosen for tracking. Two of them were popular and well-established news

sources, such as the *New York Times* in the US, or the *BBC* in the UK. The third news source was *Google News* site of each country, a news aggregator of several hundreds and sometimes thousands of popular country specific news sources. The popularity of news sites was determined by cross referencing of several indicators and sources, including the recent statistics provided by the *World Association of Newspapers*,[2] the *State of the News Media*,[3] Nielsen online,[4] and *IVW* (Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V.). The list of popular news sites was further supported and validated by online tools such as *Alexa* and *Google Trends*.

In each news source five main topical categories were observed, including "top news," "world news," "business and economy," "technology," and "entertainment and culture." These categories were chosen because they were common to all news-sites and thus enabled a cross-national comparison. The data of each of the chosen news sites was collected based on its RSS feeds if available or by direct web-mining of the text in each news category every other day over a period of three years between February 1, 2009 and January 31, 2012 at 12:00 UTC. In total, 1,041,283 news items from 35 news sites[5] were collected and analyzed. It is important to note that these news items were not a random sample, but rather the entire collection of news items that appeared in the RSS feeds of the chosen news sites during the three-year period. Table 1 summarizes the list of countries and news sites used for the analysis as well as the number of news items provided by each.

[Table 1 here]

For each news item the date, title, content, category and source were automatically identified and documented. The title of each news item and its content were used to extract the countries mentioned in it. For this purpose, a database of 195 country names in 10 different languages was built.[6] Several native-speaker research assistants were employed to translate country names into these languages. For each country name, the research assistants were asked to provide all the common names and alternative names (e.g. “United States”, “USA”, and so on). Then they were asked to omit all alternative country names that might be ambiguous and therefore yield irrelevant search results.

On the basis of this list, the software could automatically identify what countries are mentioned in each news item. A validation process was carried out by randomly choosing 100 news items and manually coding the mentioned countries. There was 78% of agreement between the human coder and the software. This was mainly since the software counts only nouns (such as “the US” or “USA”), while human coders also include adjectives (such as “American”). This decision to strictly focus on country names and not include nationalities was based on the fact that it would complicate the translation process due to divergent grammatical rules in different languages. Previous studies (Segev, 2010; Segev and Blondheim, 2013a; Segev, Sheafer, and Shenhav, 2013), however, show that focusing on country names in a high volume of news items provides a very good proxy of the actual attention a country gives to another in its news outlets.

The main variable "News Prominence" of a country was calculated based on the percentage of news items that mentioned a county out of the total news items that mentioned country names. For this purpose self-reporting of country names was omitted, thus American news items that mentioned the US or Chinese news items that mentioned China were not included. The combined news prominence of a country by all other countries was then tested for correlation with the various independent news-flow variables listed below in order to find the best-fit model and to identify the outlying countries. Additionally, the monthly news prominence of the outlying countries was studied in order to examine their trends over the three-year period and divide them into distinctive groups.

A possible caveat is that new sites within each country may vary greatly in terms of their international scope. This could be specially the case when comparing *Google News*, an aggregator that automatically selects news items, with other news sources that employ human editors. However, a Spearman test shows a very high correlation between the country rankings in *Google News* and that of other news sites of the same county (ranging from $r = .836$, $p < .01$ in Russia to $r = .907$, $p < .01$ in the US and France). The country ranking correlations between news sites and *Google News* from different countries were significantly lower (A Fisher z-transformation indicates significant differences between the two with $p < .01$). In other words, *Google News* presents a significantly similar international scope to that of other news sites from the same country.

News flow variables

A series of political, economic, social and geographic indicators was gathered based on the three groups of variables (i.e., national traits, relatedness and event-oriented variables) suggested by the literature surveyed above on international news flow. The initial list included some 30 variables, but the final 14 variables that were found to be the most significantly correlated with the news prominence of countries are reported here:

National Traits.

CINC (political) – the Composite Index of National Capability developed by Singer et al. (1987) is a comprehensive index for the general power of a country based on its population, urban population, iron and steel production, energy consumption, military personnel, and military expenditure. The latest CINC scores are from 2007 (http://correlatesofwar.org).

Military Capability (political) – based on a qualitative assessment of the Economist Intelligence Unit's analysts, this variable measures the level of military sophistication and extent of research and development in each country on a scale of 1 to 5. The 2010 scores are used in this study (http://www.visionofhumanity.org/gpi-data/#/2010/mcap).

GDP (economic) – calculated in US dollars and is based on the World Economic Outlook Database of the International Monetary Fund. The 2010 data are used in this study (http://www.imf.org).

Population (social) – based on the World Population Prospects of the UN Department of Economic and Social Affairs Population Division, the 2010 data are used in this study (http://esa.un.org/unpd/wpp/unpp/panel_population.htm).

Area (geographic) – measures the size of a country in square kilometers based on the Demographic Yearbook of the UN Statistics Division (http://unstats.un.org/unsd/demographic/products/dyb/dyb2008/Table03.pdf).

Relatedness.[7]

Bilateral Trade (economic) – measures the total value of imports and exports of a country in US dollars. The 2010 data is based on the Direction of Trade Statistics Yearbook of the International Monetary Fund (http://www.imf.org/external/data.htm).

Foreign Population (social) – the foreign-born population in a country is based on national censuses conducted between 2000 and 2011 as reported by the UN Statistics Division (http://data.un.org/Data.aspx?d=POP&f=tableCode%3a44).

Region (geographic) – the number of countries sharing the same region with a country is based on the UN classification of regions (http://unstats.un.org/unsd/methods/m49/m49regin.htm).

Border (geographic) – the number of countries sharing the same border with a country is based on the CIA World Factbook (https://www.cia.gov/library/publications/the-world-factbook).

Event-oriented.

Conflict Intensity (political) – the extent of national and international conflicts of a country on a scale of 0 to 2 between 2009 and 2010 is based on PRIO Armed Conflict Dataset v.4-2011.

GPI (political) – the Global Peace Index is composed by the Institute for Economics and Peace, and measures the level of country's peacefulness. It is scaled from 1 to 5, the lower the score the more peaceful the country. The 2010 data is available from http://www.visionofhumanity.org

Unemployment (economic) – the unemployment rate measures the percentage of economically active population without work. The 2010 data is based on the CIA World Factbook (https://www.cia.gov/library/publications/the-world-factbook).

GDP Change (economic) – the change in percentages between the GDP level of 2009 and that of 2010. See also the 'GDP' variable.

Death Disaster (social) – is the total death toll from natural disasters occurred between 2009 and 2010 based on the International Disaster Database (EM-DAT) (http://www.emdat.be).

Results

Best-fit model based on the literature

Table 2 presents the Pearson correlation between news prominence of countries and their characteristics. In case variables were not normally distributed a logarithmic transformation was used to allow a normal distribution of the variables.[8] It shows that all variables from the national trait group are significantly correlated with news prominence. In other words, the size and power of a country is a particularly good predictor of its news prominence. Among all, the GDP has the strongest correlation ($r = .791$, $p < .01$) with news prominence. A Fisher z-transformation shows that this correlation coefficient is significantly different from that of the population variable ($z = 2.79$, $p < .01$), which has the second highest correlation with news prominence in the national trait group.

From the relatedness group, bilateral trade (imports and exports) is a very strong predictor of countries salience ($r = .727$, $p < .01$), but also the number of foreigners in a country ($r = .654$, $p < .01$). A Fisher z-transformation indicates that the difference between the correlation coefficient of the two variables with news prominence is not significant. Hence, the first hypothesis was partly confirmed. Indeed, the economic power of a country is the strongest predictor of its news prominence. However, country that has many foreigners has also higher chances to be mentioned in the news than countries with few foreigners.

Finally, from the event group, the intensity of a conflict has a weak yet significant correlation with news prominence. Similarly, the death toll from disaster is significantly correlated with news prominence. A Fisher z-transformation indicates no significant difference between the correlation coefficient of these two variables. In contrast, economic-related variables that reflect deviance or instability such as GDP change or unemployment rate are not significantly correlated with news prominence. In short, economic strength on the one hand, and political and social instabilities on the other, have the strongest association with news prominence.

[Table 2 here]

All inner correlations between the news flow variables were studied in order to reach the best fit multiple regression model that explains the maximum variance of countries news prominence. When conducting multicollinearity tests, it was found that out of all potential variables, three variables—GDP, Foreign Population and Conflict Intensity—account for more than 70% of the news prominence variance ($R^2 = .714$). This provides a strong support for the second hypothesis. Adding other variables, however, does not increase the explanatory power of the model, and provokes a multicollinearity problem. These three variables have also the strongest correlation with news prominence and are taken from the three groups of variables suggested by the literature—national traits, relatedness and events. Interestingly, they represent three different spheres of explanation for the news prominence—economy, society and politics

respectively. Table 3 presents the correlation matrix between these variables and the results of the multiple linear regression model. It is worth to mention here that GDP is also strongly correlated with Foreign Population and very weakly correlated with Conflict Intensity. Yet, the high tolerance and low VIF levels indicate that these inner correlations do not reduce the predictive power of the model or the individual predictors.

[Table 3 here]

The limits of the model

In order to identify the over-represented and under-represented countries in the news, Table 4 lists the outlying countries that are one standard deviation or more away from the expected value if predicted by the model. Over-represented countries are countries that got much higher attention by the news than would be expected if explained by their GDP, foreign population and involvement in conflicts. When looking carefully at the values of these variables as well as at the news monthly trends of these countries they could be divided into four groups:

Ad-Hoc Events. The first group of countries includes Haiti, Iceland, Egypt, Honduras and Madagascar. They are located in different parts of the world, they all have relatively low GDP, and foreign population and they are not involved in major international conflicts. What made them newsworthy was the occurrence of an unexpected major event during the data-gathering period such as a natural disaster (in the case of Haiti and Iceland), or a sudden uprising (in the case of Egypt, Honduras and Madagascar). This could not be reflected by the

independent variables and thus rendered these countries as residuals with more than one standard deviation than the expected value.

In order to quantify the level of change in news attention for the over-represented countries a Surprise Index was developed. It is defines as the maximum monthly news coverage divided by the average news coverage of each country. Table 4 shows that the Surprise Index for all countries in the first group is higher than five, suggesting that the news attention for these countries was unusual during one month, namely at least five times (or in the case of Haiti's earthquake 13 times) more than the average. While these findings provide an initial support for the third hypothesis, the rest of the outlying countries discussed below fall into different categories and thus shed new light on the limits of the current model.

Ongoing Political Events. The second group includes countries that are involved in conflicts. They all have medium to low level of GDP and foreign population. Yet, news around the world covers them more than expected. Interestingly, this group includes: Afghanistan, Iran, Iraq, Lebanon, Jordan, Israel, Syria, Sudan and Georgia, most of which are Middle Eastern countries. Unlike the first group of ad-hoc events, countries with ongoing political events have a relatively low Surprise Index since their involvement in conflicts is covered by the world continuously.

Ongoing Economic Events. The third group includes the European countries Greece and Ireland with medium GDP and foreign population and with

no military conflicts. The ongoing news coverage on these countries is related to their recent economic crises and relation to the broader EU.

Regional Center. Finally, the fourth group includes both European and Asian countries: the UK, France, China and South Korea. These countries have relatively high level of GDP and foreign population, but they were not involved in major conflicts during the data-gathering period. Still, the attention given to these countries by world news is much higher than would be expected if explained by these variables.

[Table 4 here]

Similarly, Table 4 lists the under-represented countries. These countries usually do not get more than 0.1% of news attention, and can be again divided into three groups based on the independent variables. In most cases, it would be meaningless to look at changes in their coverage over time and compute the Surprise Index due to their very low average of news coverage.

Invisible conflicts. The first group includes countries that are involved in major conflict. Most of them also have medium levels of GDP and foreign population, but still they were not found to be newsworthy by news sites around the world. These countries are mainly from Africa and Asia: Angola, Ethiopia, Central African Republic, Algeria, and the Philippines.

Regional Peripheries. The second group includes countries that were not involved in major conflicts. Their news prominence was, however, much lower than would be expected based on their medium GDP or foreign population levels.

This group includes mainly East European and Middle Eastern peripheries: Slovakia, Slovenia, Lithuania, Albania, Oman, Kuwait, but also the Dominican Republic and Bangladesh.

Global Peripheries. Finally, the third group includes countries with very low GDP and foreign population levels. They were also not involved in major conflicts. Still, their news prominence was much lower than would be expected. This group includes mostly African countries: Swaziland, Botswana, Zambia, and Lesotho, but also Suriname and Barbados.

Discussion

This paper provides a clear empirical support for the general news-flow theory, but utilizing advanced web-mining tools and a large corpus of news data over a three year period, it indicates some of its current limitations, and identifies possible directions for developing it a step further. Following previous studies (Ishii, 1996; Kariel and Rosenvall, 1984; Kim and Barnett, 1996; Robinson and Sparkes, 1984; Segev and Blondheim, 2013a; Segev and Hills, 2013; Wu, 2000, 2007), the first hypothesis was that country's GDP would have the strongest association with news prominence. It is worth mentioning here that most previous studies were based on human observations and therefore also limited in their scope. The present study confirms previous observations in support of the first hypothesis. The economic power of a country was also the most significant variable associated with its prominence in online news around the world.

Similarly, bilateral trade was a very prominent factor among the relatedness variables. Still, in line with some previous observations (Chang et al. 1987; Kariel and Rosenvall, 1984), foreign population was found to be another important factor of country news prominence among the relatedness variables. Yet, when it comes to events, the political and social factors (such as conflicts and death toll from disasters) rather than the economic factors were much more significantly associated with news prominence.

In search for the best-fit model, the findings show that a combination of three factors: GDP, foreign population and conflict intensity account for more than 70% of the variance of country news prominence. Strikingly, these variables also represent three theoretical explanations—national traits, relatedness and events, taken from three different spheres—economy, society and politics respectively. The model developed in this paper outperformed the previous model suggested by Wu (2000, 2007), mainly since it was based on much larger data set and much longer period of observation. Additionally, the operationalization of variables from the three different theoretical groups contributed to locating the most powerful combination of factors. This provides an empirical support for the second hypothesis. The rest 30% of the news prominence variance could be partly revealed when looking at the over- and under-represented countries and grouping them based on their characteristics and news trends.

The third hypothesis assumed that over-represented countries would be those experiencing unexpected events that could not be measured by the current

variables. Indeed, previous studies suggested ways to operationalize the deviance of countries (Shoemaker et al. 1986, 1991; Pritchard and Hughes, 2006), and subsequently this study included a list of event-oriented variables such as conflict intensity or death toll from disaster. Yet, among the over-represented countries, there was a group of countries such as Haiti, Iceland and Egypt, in which the present measurements could not reflect the magnitude of their deviance. This, however, was only one group of countries found to be over-represented, providing only a partial support for the third hypothesis. While it might be challenging to operationalize variables that predict ad-hoc events, there are still many ways to improve our understanding of the national trait and relatedness groups of variables as suggested bellow.

Visible and Invisible Conflicts

The results suggest that conflict is indeed an important factor of newsworthiness of countries. Yet, there are some conflicts that are more newsworthy than others. In particular, findings indicate that Middle Eastern conflicts are more newsworthy than African or Asian conflicts. Over-represented countries in relation to ongoing political events and conflicts were mainly from the Middle East. Under-represented countries in relation to conflicts were solely from Africa and Asia.

There might be several reasons for this finding. First, the Middle East is physically closer to the west, and particularly to the strongest and wealthiest

countries in Europe that get more news attention anyway (see the discussion on regional centers). This could also be related to the growing number of immigrants coming from the Middle East to Europe. Second, all countries that were found to be over-represented in relation to ongoing political events involve religious clashes between Muslims and Jews/Christians. These findings provide initial indication to the possible inclusion of religious parameters when looking at the conflict component in relation with the news flow theory (See also Segev and Blondheim, 2013b).

Apart from political conflicts, two other over-represented countries—Ireland and Greece—were identified. Being an integral part of EU it is clear why a western dominated international news system would give priority to ongoing economic changes in Europe. Yet, in order to include these middle-term events, scholar should find creative ways to operationalize them into measurable variables. Similar to ongoing political events, in the case of ongoing economic events, variables should also interact with the general national trait variables, since findings clearly show that middle-term events are reported only when they are related to the economically leading countries.

Regional Centers and Peripheries

Findings indicate that over-represented countries with high GDP and foreign population but with no major military conflicts are mostly the regional centers of West Europe and Asia. While previous studies already pointed on the

importance of regional centers (Tunstall, 2008), and particularly the European ones when it comes to the news (Kim and Barnett, 1996; Wu, 2000, 2004), the present findings indicate that they cannot be attributed solely to their economic power. Perhaps the gap between their observed news prominence and their expected values should be also understood in the regional historical and cultural context (particular when it come to the UK, France and China). There is possibly also a certain "rich gets richer" effect in news prominence. The most prominent country in a region—For example, the leading regional economy, the leading regional military power, or the leading regional conflict—gets specially more news attention because of its being number one. These new considerations should be taken into account in future studies attempting to operationalize and provide a broader explanation for the news flow theory.

On the other hand, under-represented countries with medium GDP and foreign population but with no major conflicts are mostly regional peripheries in East Europe and the Middle East. Countries such as Slovakia, Slovenia, Lithuania, Oman or Kuwait are expected to have a greater news prominence based on their economies, but are forgotten, perhaps being in the shadow of the bigger regional players. This observation was not taken into account in the previous studies surveyed above.

Global Centers and Peripheries

At the global level, under-represented countries were mostly from Africa. Despite the relatively low values of their independent variables, they were expected to generate some level of news coverage. Yet, it is possible to say that these countries are literally invisible to the world. While this finding is not new (see for example Golan, 2008), it certainly requires operationalization in a new model of news-flow. Perhaps, future studies should first identify the core economic and political regional centers, and only then re-calculate the traditional relatedness and event-oriented variables based on their economic and political interactions with the core countries.

On the other hand, a proper candidate for an over-represented country in the global level would be the US. Indeed, the data in this study confirm that the US was just about one standard deviation away from its expected value. It was, however, not included among the top outliers. The US is a particular case since its GDP and foreign population level is the highest by far and it certainly involves in various international conflicts. Its news prominence around the world reflects this trend closely with 18.61% of all news items mentioning it. The US would be therefore a very good example for a world hegemon, demonstrating that the "rich gets richer" effect works also on the global level. The US gets much higher news attention than the rest of the world not only because of its high economic, social or political power, but also because of its being number one (see also Segev and Blondheim, 2013a for the broader study on the US news dominance).

It is important to acknowledge, however, the boundaries of this study. First, the present analysis is based on news sites from countries with the largest number of online users. An ideal study should include more countries representing more regions. Specifically, there are much fewer news-flow studies in developing countries. Although they get most of their news from international news agencies, their news flow might as well be more regional in focus and thus require further modifications in the current theory.

Second, one of the assumptions of many news-flow related studies is that "external" variables such as economic power or political events in a country can explain its news prominence around the world. In fact, there are also various "internal" factors such as the structure of international telecommunications, the presence of news agencies (Wu, 2000), as well as the editorial practices and traditions that affect the news prominence of a country. Yet, the approach taken here and in many similar studies is that the external factors have certain affect also on internal factors, particularly when looking at a long period of news coverage from various news sources. For example, although one outlet in China may take a more regional approach than another, they will both find it difficult to ignore the economic crisis in the US, which directly affects the Chinese economy as well. It might be worth, however, to examine more thoroughly this specific relationship between external variables and those related to the editorial decision-making and the journalistic tradition in each outlet separately.

It is out of the scope of this study to draw a comprehensive picture of the news coverage in the entire world, or to explore the journalistic practices behind it. The main goal of this paper was rather to develop a strong predictive model of news flow based on the existing theory, and to unveil its limitations. To this end, the web-mining tools employed here allowed exploring a much larger corpus of data and many more languages than done before. Findings clearly point on the visible and the invisible countries in world news based on their relation to the local, regional and global economy, politics and society. Most importantly, they invite scholars to expand the news-flow theory accordingly.

References

Charles J, Shore L and Todd R (1979) The New York Times coverage of equatorial and lower Africa. *Journal of Communication* 29: 148–155.

Chang TK (1998) All countries not created equal to be news: World system and international communication. *Communication Research* 25: 528–563.

Chang TK, Shoemaker P and Brendlinger N (1987) Determinants of International News Coverage in the US Media. *Communication Research* 14: 396–414.

Dupree JD (1971) International communication: View from 'a window on the world'. *The International Communication Gazzette* 17: 224–235.

Galtung J and Ruge MH (1965) The Structure of Foreign News. *Journal of Peace Research* 2(1): 64–91.

Gerbner G and Marvanyi G (1977) The many worlds of the world's press. *Journal of Communication* 27(1): 52–66.

Golan GJ (2008) Were in the world is Africa? Predicting coverage of Africa by US television networks. *The International Communication Gazette* 70: 41–57.

Golan G and Wanta W (2003) International Elections on the US Network News: An Examination of Factors Affecting Newsworthiness. *Gazette* 65(1): 25–40.

Hawkins V (2008) *Stealth Conflicts: How the World's Worst Violence is Ignored*. Hampshire, UK: Ashgate.

International Press Institute (1953) *The Flow of News*. Zurich: International Press Institute.

Ishii K (1996) Is the U.S. over-reported in the Japanese press? Factors accounting for international news in the Asahi. *Gazette* 57: 134–144.

Kariel HG and Rosenvall LA (1984) Factors Influencing International News Flow. *Journalism Quarterly* 60: 434–436.

Kayser J (1953) *One Week's News: Comparative Study of 17 Major Dailies for a Seven-Day Period*. Paris: UNESCO.

Kim K and Barnett GA (1996) The Determinants of International News Flow: A Network Analysis. *Communication Research* 23(3): 323–352.

Ostgaard E (1965) Factors influencing the Flow of News. *Journal of Peace Research* 2(1): 39–64.

Pritchard D and Hughes KD (2006) Patterns of Deviance in Crime News. *Journal of Communication* 47(3): 49–67.

Robinson GJ and Sparkes VM (1976) International News in the Canadian and American Press: Comparative News Flow Study. *International Communication Gazette* 22: 203–218.

Rosengren KE (1970) International News: Intra and Extra Media Data. *Acta Sociologica* 8(2): 96–109.

Rosengren KE (1977) Four types of tables. *Journal of Communication* 27: 67–75.

Segev E and Hills T (2013) When News and Memory Come Apart: A Cross-National Comparison of Countries' Mentions. *International*

Communication Gazette. Available at: http://gaz.sagepub.com/content/early/2013/10/18/1748048513504159

Segev E, Sheafer T and Shenhav S (2013) Is the World Getting Flatter? A New Method for Examining Structural Trends in the News. *Journal of the American Society for Information Science and Technology (JASIST)* 64(12): 2537–2547.

Sheafer T, Ben Nun P, Shenhav S and Segev E (2013) The Conditional Nature of Value-Based Proximity between Countries: Strategic Implications for Mediated Public Diplomacy. *American Behavioral Scientist (ABS)* 57(9): 1256–1276.

Segev E and Blondheim M (2013a) America's Global Standing According to Popular News Sites from Around the World. *Political Communication* 30(1): 139–161.

Segev E and Blondheim M (2013b) Online News about Israel and Palestine: A Cross-National Comparison of Prominence and Trends. *Digital Journalism* 1(3): 1–13.

Segev E (2010) Mapping the International: Global and Local Salience and News-Links Between Countries in Popular News Sites Worldwide. *International Journal of Internet Science* 5(1): 48–71.

Shenhav SR, Rahat G and Sheafer T (2012) Testing the language-power assumption of critical discourse analysis: The case of Israel's legislative discourse. *Canadian Journal of Political Science* 45(1): 207–222.

Shoemaker PJ, Chang TK and Brendlinger N (1986) Deviance as a Predictor of Newsworthiness: Coverage of International Events in the US Media. *Communication Yearbook* 10: 348–365.

Shoemaker PJ, Danielian LH and Brendlinger N (1991) Deviant Acts, Risky Business and US Interests: The Newsworthiness of World Events. *Journalism Quarterly* 68: 781–795.

Singer JD, Bremer S and Stuckey J (1972) Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In: Russett B (ed) *Peace, War, and Numbers*. Beverly Hills: Sage, pp. 19–48.

Tunstall J (2008) *The Media Were American: US Mass Media in Decline*. Oxford: Oxford University Press.

UNESCO (1954) *How Nations See Each Other?* Paris: UNESCO Publications.

Wallerstein I (1974) *The modern world-system*. New York: Academic Press.

Wanta W, Golan G and Cheolhan L (2004) Agenda setting and international news: Media influence on public perceptions of foreign nations. *Journalism and Mass Communication* 81(2): 364–377.

Wilke J, Heimprecht C and Cohen A (2012) The geography of foreign news on television: A comparative study of 17 countries. *International Communication Gazette* 74(4): 301–322.

Wu DH (2000) Systematic Determinants of International News Coverage. *Journal of Communication* 50(2): 113–30.

Wu DH (2004) The world's windows to the world: an overview of 44 nations' international news coverage. In: Paterson C and Sreberny A (eds) *International News in the Twenty-First Century*. UK: John Libbey Publishing, pp. 95–108.

Table 1.

Countries, news sites, and the number of news items collected for analysis

Country	News Site	News Items	Country	News Site	News Items
USA	Google News	21542	Russia	Google News	22914
USA	CNN	15240	Russia	Gazeta	117125
USA	NY Times	36269	Russia	Pravda	30989
UK	Google News	36995	Arabic	Google News	22361
UK	BBC	57922	Arabic	Al Jazeera	35181
UK	Guardian	42221	Egypt	Al Ahram	3340
Israel	Google News	22536	Egypt	Al Masry Alyoum	9700
Israel	Ynet	24447	Iran	PressTV	2441
Israel	Haaretz	8497	Iran	Tabnak	71446
Germany	Google News	21541	Iran	Aftab	16768
Germany	Bild	29300	China	Google News	30903
Germany	Spiegel	29294	China	Sina	26633
Germany	Welt	4453	China	People Daily	95697
France	Google News	21304	Japan	NHK	38435
France	Le Monde	26770	Japan	Yomiuri	2650
France	Le Figaro	22669	Japan	Yahoo	15518
Spain	Google News	21574			
Spain	El Mundo	25752			
Spain	El Pais	30856	**Total News Items**		**1041283**

Table 2.

Pearson correlation of countries' news prominence with news flow variables

		News Prominence	Fisher z-transformation
National Traits	GDP	.791**	z = 2.79 ($p < .01$)
		(180)	
	Population	.654**	
		(191)	
	CINC	.649**	
		(188)	
	Area	.491**	
		(187)	
	Military Capability	.470**	
		(150)	
Relatedness	Bilateral Trade	.727**	z = 1.35 (n.s.)
		(195)	
	Foreign Population	.654**	z = 3.35 ($p < .01$)
		(182)	
	Border	.409**	
		(195)	
	Region	.002	
		(195)	
Events	Conflict Intensity	.328**	z = 0.35 (n.s.)
		(191)	
	Death Disaster	.293**	z = 1.9 ($p = .05$)
		(153)	
	GPI	.080	
		(147)	
	GDP Change	.052	
		(177)	
	Unemployment	-.030	
		(142)	

Note: Numbers in parentheses indicate the total number of countries (n). The data was log-transformed to allow a normal distribution. Variables are sorted by their correlation values to news prominence within each group. A Fisher z-transformation shows the difference between the correlation coefficient of a variable with that of the variable below. ** indicates $p < 0.01$, * indicates $p < 0.05$.

Table 3.
Correlation matrix and best-fit multiple linear regression of news prominence

	GDP	Foreign Population	Conflict Intensity
News Prominence	.791**	.654**	.328**
	(180)	(182)	(191)
GDP		.727**	.149*
		(170)	(181)
Foreign Population			0.121
			(180)

R Square = .714 Durbin-Watson = 1.847

	Beta	T	Sig.	Collinearity Statistics	
				Tolerance	VIF
(Constant)		-19.020	0.000		
GDP2010	0.652	10.634	0.000	0.463	2.159
Foreign Population	0.184	3.007	0.003	0.468	2.137
Conflict Intensity	0.192	4.559	0.000	0.979	1.021

Note: Numbers in parentheses indicate the total number of countries. ** indicates $p < 0.01$.

Table 4.
List of over- and under-represented countries in the news as predicted by the best-fit model

Country	News Prominence (%)	Standardized Residuals	Surprise Index	GDP	FP	CI	Group
Over-Represented Countries							
Haiti	1.06	2.99	13.72	134	48	0	
Iceland	0.36	2.04	9.70	115	162	0	**Ad-Hoc Events**
Egypt	1.28	1.52	5.68	40	78	0	**(Low GDP, FP, No CI)**
Honduras	0.38	1.37	9.13	109	60	0	Different parts of the world
Madagascar	0.22	1.29	7.44	129	126	0	
Afghanistan	2.84	2.40	1.76	108	102	2	
Lebanon	0.96	2.07	2.04	81	80	0	
Georgia	0.56	2.06	3.64	120	85	0	
Israel	3.85	1.82	1.68	41	24	1	**Ongoing Political Events**
Jordan	0.54	1.75	2.09	91	113	0	**(Med/Low GDP, FP, High CI)**
Iran	3.90	1.66	2.07	26	45	1	Mostly Middle Eastern Countries
Syria	0.66	1.52	2.35	66	104	0	
Sudan	0.93	1.26	3.21	64	135	1	
Iraq	2.52	1.25	1.69	62	65	2	
Greece	1.58	1.35	3.24	32	22	0	**Ongoing Economic Events**
Ireland	1.06	1.09	2.92	42	25	0	European Countries
China	5.53	1.41	1.51	2	28	0	
South Korea	1.97	1.37	1.75	15	98	0	**Regional Center**
UK	4.34	1.27	1.44	6	4	0	**(High GDP, FP, No CI)**
France	3.94	1.12	1.47	5	5	0	Europe and Asia

Table 4 - Continue -

Country	News Prominence (%)	Standardized Residuals	GDP	FP	CI	Group
Under-Represented Countries						
Angola	0.07	-2.19	61	93	1	**Invisible Conflicts (Medium GDP, FP, High CI)** Africa & Asia
Ethiopia	0.09	-1.50	89	90	1	
Central African Republic	0.02	-1.45	159	163	1	
Algeria	0.23	-1.41	49	26	1	
Philippines	0.32	-1.26	45	13	1	
Sri Lanka	0.24	-1.14	73	57	2	
Dominican Republic	0.05	-1.78	72	35	0	**Regional Peripheries (Med.-Low GDP, FP, No CI)** Mainly East European & Middle Eastern Peripheries
Slovakia	0.06	-1.72	60	59	0	
Oman	0.07	-1.43	68	36	0	
Lithuania	0.06	-1.37	84	71	0	
Albania	0.04	-1.36	118	50	0	
Bangladesh	0.14	-1.23	57	10	0	
Slovenia	0.07	-1.08	74	108	0	
Kuwait	0.09	-1.06	54	143	0	
Suriname	0.01	-3.01	151	99	0	**Global Peripheries (Low GDP, FP, No CI)** Mainly Africa
Swaziland	0.01	-1.79	150	154	0	
Botswana	0.02	-1.40	110	159	0	
Zambia	0.04	-1.25	106	119	0	
Barbados	0.02	-1.16	147	128	0	
Lesotho	0.02	-1.07	155	117	0	

Note: GDP and Foreign Population (FP) are represented here in their ranks out of the 200 countries rather than by the absolute values. Conflict Intensity (CI) is based on PRIO Armed Conflict Dataset. Surprise Index is calculated as the maximum news prominence divided by the average news prominence over three years. The list is sorted by the standardized residuals within each group of countries.

Notes

[1] Wu studied the countries mentioned in websites of the *New York Times* and the *CNN*, comparing with the countries mentioned in the newspaper and the TV channel respectively.

[2] http://www.wan-press.org

[3] http://www.stateofthenewsmedia.org

[4] http://www.nielsen-online.com

[5] Google News did not include special Egyptian or Iranian editions in the course of this study, but its Arabic edition as well as Al Jazeera were used as more general and popular news sources in the Arab world.

[6] Based on the most complete list of country names available from ISO (International Organization for Standardization). This list was translated into the following languages: English, French, German, Spanish, Russian, Chinese (Mandarin), Japanese, Persian, Arabic and Hebrew.

[7] Since the dependent variable—countries' news prominence—is an aggregation of news from all countries, the relatedness-variables are also built as an aggregation of the total relations of a country with the outside world. For example, the bilateral trade variable represents the total imports and exports of each country with all other countries.

[8] The logarithmic transformation also ensured that there were no outliers to significantly divert the correlation results. Another validation was a Spearman correlation test that yielded identical results.