

The Determinants of International Investment and Attention Allocation: Using internet search query data*

Jordi Mondria[†]
University of Toronto

Thomas Wu[‡]
UC Santa Cruz

Yi Zhang[§]
UC Santa Cruz

February 2009

Abstract

Few studies have successfully examined the empirical impact of limited information processing on real economic variables. The challenge relies, of course, on the difficulty of measuring an economic agent's degree of attention/inattention paid to different types of information. This paper overcomes such challenge exploring a unique dataset containing the "search/click-through" behavior of internet search engine users. We analyze the effect of attention allocation on international investment decisions by combining U.S. data on portfolio holdings of foreign securities with the attention allocated by 657,426 America Online customers in search queries towards these countries. We find evidence that: (i) agents tend to search more information about countries in which they hold more assets, and (ii) agents tend to invest more in countries where they process more information.

Keywords: Foreign Asset Holdings, Attention Allocation, Rational Inattention, Internet Search Query.

JEL Codes: F30, D82, G11.

*The authors would like to thank comments and suggestions from Ricardo Reis, and seminar participants at UC Davis, Getulio Vargas Foundation - Sao Paulo and the National Taiwan University.

[†]Economics Department, Sidney Smith Hall, Room 4085. E-mail: jordi.mondria@utoronto.ca.

[‡]Economics Department, 465 E2 Building, Santa Cruz, CA 95064. E-mail: thomaswu@ucsc.edu.

[§]School of Engineering, 565 E2 Building, Santa Cruz, CA 95064. E-mail: yiz@soe.ucsc.edu.

1 Introduction

Few studies have successfully examined the empirical impact of limited information processing on real economic variables. The challenge relies, of course, on the difficulty of measuring an economic agent's degree of attention/inattention paid to different types of information. This paper overcomes such challenge exploring a unique dataset containing the "search/click-through" behavior of internet search engine users. We analyze the effect of attention allocation on international investment decisions by combining U.S. data on portfolio holdings of foreign securities with the attention allocated by 657,426 America Online (henceforth AOL) customers in search queries towards these countries.

The idea that "*most people everyday encounter, or could very easily encounter, much more information that is in principle relevant to their economic behavior than they actually respond to*" (Sims, 2006) - known as rational inattention - has recently been incorporated into economic models. The basic implication of this assumption is that agents have to decide the relevant information for their decision-making.

A number of applications to macroeconomics and finance have been written in the past years. In macroeconomics, rational inattention helps explaining sticky prices in Mackowiak and Wiederholt (2007) and Mankiw and Reis (2002) and consumption dynamics in Luo (2007). Faulkinger (2007) develops a theoretical model in which firms have to compete first for consumers' attention before competing for their budgets. In finance, it has been used to better understand the equity-premium puzzle in Gabaix and Laibson (2002), contagion across emerging economies in Mondria (2007), portfolio under-diversification in Van Nieuwerburgh and Veldkamp (2007a) and home equity bias in Van Nieuwerburgh and Veldkamp (2007b) and Mondria and Wu (2007).

Due to limitations on datasets, most of these models rely on calibration and simulation. One major criticism to this technique is with respect to the parameters that refer to the limits in the information processing capacity. Other papers rely on regressions with variables that proxy attention allocation. As a measure of information flow, Portes and Rey (2005) uses volume of call traffic and Mondria and Quintana-Domeque (2006) uses the number of stories

about a country in the newspaper. Mondria and Wu (2007) uses the number of telephone mainlines divided by a country's GDP as a measure of information capacity. On an third different approach, Gabaix et al (2003) study limited attention allocation on an experimental lab. However, the degree of noise that separates these proxy variables to the real unobservable economic concepts may generate some skepticism towards the empirical results.

Measuring attention allocation is a complicated task. Agents process information using different media (i.e., books, magazines, newspapers, radio, TV, the internet). For most of these different type of media, it is virtually impossible to track the amount of effort spent by the agent on processing information about different topics. On August 3, 2006, AOL released a dataset that includes over 21 million web search queries¹ from 657,426 customers. The data includes all searches from those users for a three month period (March 1st to May 31st of 2006), as well as whether they clicked on a result and where it appeared on the result page. In this paper, we explore the AOL dataset to measure the attention allocated to different countries. We believe that a measure based on the internet search queries can proxy attention allocation because: first, the World Wide Web is becoming the predominant information media; second, search engines are the most popular tool to help users find reliable information on the "Web" since it minimizes the time required and the amount of information which must be consulted; third, by having the exact search query, we know what is the topic that the user was interested in.

Using this dataset, we constructed a click-through series that counts the number of times a user clicked through a search result from a particular country. The idea is to measure the attention allocated to a country by the number of times this country provided the answer to a search query. We combine the click-through series with data from the 2006 survey of U.S. portfolio holdings of foreign securities, sponsored by the Department of the Treasury, the Federal Reserve Bank of New York, and the Board of Governors of the Federal Reserve System. The survey measured U.S. holdings of foreign securities at year-end 2006 of approximately \$6.0 trillion. Figure 1 reveals a very close relationship between attention and investment allocation

¹A web search query is the exact phrase that a user types into the search engine to satisfy his or her information needs.

on foreign equities.

[Insert Figure 1 about here]

The relationship between U.S. foreign equity holdings as of 31 of December of 2006 and the AOL country attention allocation between March and May of 2006 can potentially be explained by three hypotheses:

- i. No causality, just correlation caused by the development level in each country;
- ii. Causality runs from asset holding to attention: other things equal, agents tend to pay more attention to the countries at which they hold more assets;
- iii. Causality runs from attention to holdings: other things equal, agents tend to invest more in countries from which they process more information.

While the first explanation is the most obvious one, the other two are by far more interesting. In this paper, we build a model with inattentive investors that rationalizes the two-way causal relationship between attention allocation and international investment. We find empirical evidence that the strong relationship between asset holdings and attention allocation is actually a result of all three factors. First, there is correlation caused by a country's development level. This effect is captured by the positive effect of GDP per capita on both the attention allocated to and also the amount of securities held in each country. Agents not only invest more but also pay more attention to more developed economies. Second, causality runs from asset holdings to attention allocation. Using instrumental variables that capture the level of development in financial markets, we find that agents do pay more attention to the countries at which they hold more assets. Third, causality also runs from attention allocation to asset holdings. Using instrumental variables that are related to a country's popularity among internet users, we find that international investors tend to hold more assets at more familiar countries. The rest of the paper is organized as follows. Section 2 develops the model. Section 3 describes the dataset. Section 4 shows the empirical results. Section 5 concludes.

2 Model

The objective of the model is to show the interaction between asset holdings and attention allocation. In particular, we are interested in finding the effect of asset holdings on attention allocation and the reverse, the effect of asset holdings on attention allocation.

2.1 Model Description

This is a partial equilibrium model with two countries. The economy consists of a risk free asset, which pays R units of the consumption good, and two independent risky assets. Even though this is a static model, four discrete events occur during the operation of the market. First, traders are endowed with an initial wealth W_{i0} and limited information processing resources κ . Second, investors allocate their limited information processing resources to analyze both stock markets. Third, each investor decides the optimal asset holdings $X_i = (x_{i,1}, x_{i,2})'$ given private information, which depends on the amount of information processed about each stock market. Fourth, trading ceases and investors consume the payoff realized from their portfolio.

Investors, with absolute risk tolerance parameter ρ , maximize a mean-variance objective function

$$U_i = E \left(E \left[W'_i \mid \tilde{Y}_i \right] - \frac{1}{2\rho} \text{Var} \left[W'_i \mid \tilde{Y}_i \right] \right) \quad (1)$$

subject to the following budget constraint

$$W_{i1} = W_{i0} + X'_i(\tilde{R} - RP) \quad (2)$$

where W_{i0} is the initial wealth of agent i , $X_i = (x_{i,1}, x_{i,2})'$ is the asset holdings vector of agent i , \tilde{R} is the vector of risky asset payoffs and \tilde{P} is the price vector of the risky assets.

Investors devote information capacity to process information about the vector of unknown and independent asset payoffs $\tilde{R} = (\tilde{r}_1, \tilde{r}_2)'$ with $\tilde{r}_j \sim N(\bar{r}_j, \sigma_{r,j}^2)$, for $j = 1, 2$. Agent i

receives a private signal about each risky asset given by

$$\tilde{y}_{i,j} = \tilde{r}_j + \tilde{\varepsilon}_{i,j} \text{ where } \varepsilon_{i,j} \sim N(0, \sigma_{i,j}^2) \quad (3)$$

Investors want to obtain information about the risky assets in order to reduce the uncertainty of their optimal portfolio. They face the following linear technology on processing information about the state of the economy

$$\frac{1}{\sigma_{i,j}^2} = \frac{\kappa_{i,j}}{\sigma_{r,j}^2} \quad (4)$$

The precision of a private signal is higher if more attention is allocated to that particular signal. However, investors have a limited capacity to process information about asset payoffs.

$$\kappa_{i,1} + \kappa_{i,2} \leq \kappa \text{ and } \kappa_{i,j} \geq 0 \text{ for } j = 1, 2 \quad (5)$$

where $\kappa_{i,1}$ and $\kappa_{i,2}$ are the attention allocated to each stock market. The information constraint imposes a limit in the reduction in agent i uncertainty about the future payoff of the risky asset j . The information processing resources have to be optimally divided between the two risky assets. After deciding the amount of information to be processed about each stock market, investors incorporate the information from the private signal, \tilde{Y}_i , into their beliefs through Bayesian updating. Then, investors derive their posterior beliefs about the asset payoffs and decide their optimal asset holdings.

2.2 Effect of Attention Allocation on Asset Holdings

The objective of this section is to show the effect of attention allocation on asset holdings. Suppose each agent chooses the optimal risky asset demand taking as given any attention allocation. After observing a given private signal about each risky asset, investors derive their posterior beliefs about the asset payoffs in order to choose their optimal asset holdings

$$X_i(\tilde{Y}_i) = \rho \text{Var}[\tilde{R} | \tilde{Y}_i]^{-1} E[\tilde{R} - RP | \tilde{Y}_i] \quad (6)$$

Asset holdings are increasing with attention allocation, $\frac{\partial x_{i,j}}{\partial \kappa_{i,j}} \geq 0$ when expected excess returns are positive since

$$\frac{\partial \text{Var}_{jj}[\tilde{R}|\tilde{Y}_i]^{-1}}{\partial \kappa_{i,j}} \geq 0 \quad \text{where} \quad \text{Var}[\tilde{R} | \tilde{Y}_i] = \begin{pmatrix} \frac{\sigma_{r1}^2}{1+\kappa_{i,1}} & 0 \\ 0 & \frac{\sigma_{r2}^2}{1+\kappa_{i,2}} \end{pmatrix} \quad (7)$$

Therefore, investors hold more assets in countries at which they pay more attention.

2.3 Effect of Asset Holdings on Attention Allocation

The objective of this section is to show the effect of asset holdings on attention allocation. Suppose that investors take as given their asset holdings when choosing where to allocate their information resources, then the objective function becomes

$$\min_{\kappa_{i,1}, \kappa_{i,2}} X_i' V \left(\tilde{R} | \tilde{Y}_i \right) X_i \quad (8)$$

which can be rewritten as

$$\min x_{i1}^2 \left(\frac{\sigma_{r1}^2}{1 + \kappa_{i,1}} \right) + x_{i2}^2 \left(\frac{\sigma_{r2}^2}{1 + \kappa_{i,2}} \right) \quad (9)$$

subject to the information processing constraint

$$\kappa_{i,1} + \kappa_{i,2} \leq \kappa \text{ and } \kappa_{i,j} \geq 0 \quad (10)$$

The optimal attention allocation is given by

$$\kappa_{i,1} = \begin{cases} \kappa & \text{if } x_{i1}\sigma_{r1} \geq (1 + \kappa) x_{i2}\sigma_{r2} \\ 0 & \text{if } x_{i2}\sigma_{r2} \geq (1 + \kappa) x_{i1}\sigma_{r1} \\ \frac{x_{i1}\sigma_{r1}(1+\kappa) - x_{i2}\sigma_{r2}}{x_{i1}\sigma_{r1} + x_{i2}\sigma_{r2}} & \text{otherwise} \end{cases} \quad (11)$$

and

$$\kappa_{i,2} = \kappa - \kappa_{i,1} \quad (12)$$

The attention allocated to a particular stock market is increasing with the asset holdings on this market since

$$\frac{\partial \kappa_{i,j}}{\partial x_{i,j}} \geq 0 \tag{13}$$

Therefore, investors optimally decide to pay more attention to countries in which they hold more assets.

3 Dataset

The model presented in the previous section developed a relationship between attention allocation and asset holdings. In order to test the predictions of our model, we need to find an appropriate measure of the attention that economic agents allocate to different countries. However, this is not an easy task. First, agents process information using different media (i.e., books, magazines, newspapers, radio, TV, the internet). Second, for most of these different type of media, it is virtually impossible to track the amount of effort spent by the agent on processing information about each individual country. In this section, we will show how data from internet search queries can be used to overcome these difficulties.

3.1 How to Measure Attention Allocation?

3.1.1 The World Wide Web and Internet search engines

The World Wide Web (commonly shortened to “the Web”) is becoming the predominant information system. The Web is a system of interlinked hypertext documents accessible via the Internet. With a web browser, a user may access the information contained in those documents by viewing web pages and exploring their texts, images, videos, and other multimedia. However, the exponential growth in the number of documents available, one of the main reasons for its increasing popularity, also causes a problem referred to as information overload. The high rate of new documents being added to the already available large stock, combined with a significant proportion of websites with questionable reliability, make it very difficult for

the user to identify what information is relevant. This is why Internet search engine tools are so popular. Search engines help minimize the time required to find reliable information and the amount of information that must be consulted.

Some numbers confirm this increasing popularity. According to the *World Development Indicators* database, the number of internet users in the US as a fraction of the total population has increased from 43.9% in 2000 to 63.0% in 2004. And according to Fallows (2005), 84% of adult internet users have used search engines to help them find information on the Web.² On an average day, about 53% of internet users will go online and more than half of them will use a search engine. As the author concludes, “searching is becoming a daily habit for about a third of all internet users.”

Fallows (2005) also finds that search engines are not only popular among Internet users, but also that search engine users are confident about their searching abilities, have positive experiences when using search engines and trust the search results they obtain. These conclusions were achieved after analyzing a daily tracking survey of Americans’ use of the Internet. The data was collected between May 14 and June 17 of 2004, among a sample of 2,200 adults, aged 18 and older. The study’s main findings show that:

- 84% of internet users have used search engines. On any given day, 56% of those online use search engines;

- 92% of those who use search engines say they are confident about their searching abilities, with over half of them, 52%, saying they’re “very confident”;

- 87% of searchers say that they have successful search experiences most of the time, including some 17% of users who say they always find the information they are looking for;

- 68% of users say that search engines are a fair and unbiased source of information; 19% say that they don’t place their trust in search engines.

²Among all internet activities, only the act of sending and receiving email is more popular than using search engines.

3.1.2 The AOL click-through series

The previous section shows that internet with the help of search engine tools is one of the most popular ways of finding information. However, we are still left with the problem of measuring the degree of attention allocated to each individual country.

Recently, AOL released the entire search/click-through record from 657,426 anonymous users collected over a three month period (01 March, 2006 - 31 May, 2006) for research purposes (Pass et al, 2007). This dataset consists of 36,389,567 lines of records, which includes 19,442,629 user click-through records and 16,946,938 records without user click-through. Examples of the records are shown below:

```
142  rentdirect.com 2006-03-01 07:17:12
142  www.prescriptionfortime.com 2006-03-12 12:31:06
142  staple.com 2006-03-17 21:19:29
142  westchester.gov 2006-03-20 03:55:57 1 http://www.westchestergov.com
1337 fiserv 2006-03-24 14:05:01 1 http://www.fiserv.com
1337 fiserv 2006-03-24 14:05:01 3 http://www.fiservlendingsolutions.com
1337 fiserv 2006-03-24 14:05:01 2 http://www.fiservinsurance.com
1337 fiserv 2006-03-24 14:05:01 3 http://www.fiservlendingsolutions.com
1337 integrated real estate 2006-03-27 14:52:29 1 http://www.integratedreal.com
1337 integrated real estate 2006-03-27 14:52:29 2 http://www.irisnet.net
```

A click-through record corresponds to a click-through event on a result in the search result list returned by the search engine for a query. Each record contains five pieces of information (u, q, t, r, c) : an anonymous user ID number (u), the query issued by the user (q), the time at which the query was submitted for search (t), the rank of the clicked result on the search result page (r), and the domain portion of the clicked result URL (c). If a user clicked multiple results for a single query, it is considered as multiple click-through events and saved as multiple records.

A record without user click-through corresponds to an event that a user issued a query, but did not click on any results returned back by the search engine. The record is a triplet

(u, q, t) consisting of the anonymous user ID number, the query issued by the user, and the time at which the query was submitted for search.

The AOL search/click-through data help us measure the attention allocated to different countries through the Internet not only because it contains the search query, but also because it tells us whether the user clicked in any result, and what is the domain portion of the clicked result URL. While having the exact search query helps to see what information the user was looking for, many search queries are too broad or too ambiguous, yielding many results that are irrelevant to the user. But since we know whether the user clicked through any of the results, we can identify the search results that were irrelevant. Moreover, having the web address of the results that were clicked through can help us extract additional information about the user's exact needs.

3.1.3 The Attention Allocation Variable

Using the AOL search/click-through data, we constructed the *attention* variable, which measures the number of times a user clicked through a search result from a particular country from March 1 until May 31 of 2006. In other words, we are measuring the attention allocated to a country by the number of times this countries provided the answer to a search query. This measure is constructed by taking the host name from each click-through record, and mapping it to the country where the host address has been assigned. This procedure involves finding first the IP address associated to the domain portion of the website name. This process is known as DNS lookup. Then, since IP addresses are allocated by regional organisations, it is relatively simple to find the country associated with the IP address.

3.2 Other Data

We measure *holdings* using the annual survey of U.S. residents' holdings of foreign assets³ jointly conducted by the U.S. Treasury, the Federal Reserve Bank of New York, and the Board of Governors of the Federal Reserve System. Although the asset holding values come from surveys, "*reporting to the surveys was mandatory, with penalties of noncompliance, and*

³The survey has been annual since 2003. Previous years include 1994, 1997, 1999 and 2001.

the data received were subjected to extensive analysis and editing” (Ahearne, Griever and Warnock, 2004), which guarantees the accuracy of the information. The survey measured the value of U.S. holdings of foreign securities as of December 31, 2006, of approximately US\$ 6 trillion, being the largest for the United Kingdom (US\$ 1,076 billion), followed by Japan (US\$ 596 billion) and Canada (US\$ 478 billion).

The pre-determined variables are divided into three groups: the first group contains variables that should only influence the attention allocated to each country; the second group includes those variables that should only affect the holdings of foreign assets; and in the last group has variables that could potentially affect both the attention allocation and the holdings of foreign securities.

According to Pass, Chowdhury and Torgeson (2006), two of the Top Five most frequent search query topics are related to “porn” (#3) and “places” (#5). In light of these result, the variables we chose that should only affect attention allocation are:

- *female models*: a ranking based on how popular a country’s female models are among the American society. More specifically, we construct this ranking using three portals about men’s lifestyle: AskMen, Maxim and FHM. These portals have a target reader of 26-year-old men, with a core target of 15-34 year old, and each of them compute an annual list of the 100 most beautiful models, according to the preferences of its readers. Using the 2006 rankings, we assign $101 - \#rank^j$ points to each female model where $\#rank^j$ is her position in the ranking of magazine j (i.e. 100 points to the first place, 99 to the runner up, 98 to the third place and so on). Then we map each female model to her country of citizenship and finally we compute the total number of points for each country based on her models’ nationality. The higher a country’s number of points, the more popular are its *female models* and hence the more internet searches should the country receive;

- *protected areas*: locations which receive protection because of their environmental, cultural or similar value, as a percentage of a country’s total land area. The data is obtained from the International Union for Conservation of Nature (IUCN) and UNEP World Conservation Monitoring Centre *World Database on Protected Areas*.⁴ The larger a country’s share

⁴The official definition of a protected area is “*an area of land and/or sea especially dedicated to the protection*

of protected areas, the more likely it is that it has places of natural beauty or cultural interest that are worth searching about.

In the second group, we include two variables that should be directly related to the holdings of foreign assets. Both variables measure how developed is a country's stock market. Other things equal, investors tend to prefer stock markets that have better financial transaction technologies. The variables are:⁵

- financial *market sophistication*: a qualitative measure of the level of sophistication of financial markets in each country, that ranges from 1 (poor by international standards) to 7 (excellent by international standards). the World Economic Forum's annual *Global Competitiveness Report*;

- *market capitalization*: the stock market capitalization as a percentage of GDP is a quantitative measure that proxies a country's stock market development level by its relative size. The data is collected from the World Bank's *World Development Indicator* database. Since more recent data is missing for many countries, we use for each country the average stock market capitalization between 2001 and 2005.

The last group contains the log of each country's 2006 *GDP per capita* as a measure of its economic development. Economic development captures the effect of many variables that are associated to the quality of a country's institutions. Examples of such qualities could be corruption, the confidence in the judicial system or in courts to uphold property rights, and the stability of the political system. Other things equal, we expect Americans to hold more assets in countries with more favorable investor-protection mechanisms or more stable political systems. However, economic development may also affect the level of attention allocated towards a country. More developed economies could offer websites or webservices that are either more interesting or more sophisticated, and hence attract more attention.

Table 1 presents each country's statistics and table 2 shows the summary statistics.

[Insert Tables 1 and 2 about here]

and maintenance of biological diversity, and of natural and associated cultural resources, and managed through legal or other effective means."

⁵Both variables were used in Portes and Rey (2005) to explain bilateral gross cross-border equity flows.

Table 3 has the correlation matrix. We can notice that the highest correlation is between *attention* and *holdings*, 81%. It is also interesting to see that *female models* and *protected areas* have low correlation with *GDP per capita*. This is an indication that beauty patterns or locations with cultural value are unrelated to the development level.

[Insert Tables 3 about here]

4 Empirical Evidence

In this section we present the empirical evidence of the two-way causal relationship between attention allocation and foreign asset holdings:

$$\begin{aligned} attention_i &= \beta_0^a + \beta_1^a holdings_i + \beta_2^a female\ models_i + \dots \\ &\dots + \beta_3^a protected\ areas_i + \beta_4^a GDP\ per\ capita_i + \varepsilon_i^a \end{aligned} \quad (14)$$

$$\begin{aligned} holdings_i &= \beta_0^h + \beta_1^h attention_i + \beta_2^h market\ sophistication_i + \dots \\ &\dots + \beta_3^h market\ capitalization + \beta_4^h GDP\ per\ capita_i + \varepsilon_i^h \end{aligned} \quad (15)$$

where, from equations (7) and (13), we have $\beta_1^a = \frac{\partial \kappa_{i,j}}{\partial x_{i,j}} \geq 0$ and $\beta_1^h = \frac{\partial x_{i,j}}{\partial \kappa_{i,j}} \geq 0$.

We will present the estimation output using two different types of estimation procedures, starting with equation-by-equation 2SLS and then moving to simultaneous equation 3SLS. Although the 3SLS estimator is more efficient, it requires that the system is perfectly specified. A wrong instrument in one equation would bias the coefficients of the other equation in the system. In that sense, the 2SLS estimator is more robust.

4.1 Equation-by-equation 2SLS

Table 4 presents the estimation output of the first-stage regressions from the 2SLS procedure. At a first glance, the fact that *female models* is significant at the *holdings* equation might raise some concerns about its validity as an instrument in the *attention* equation. However, we believe that this coefficient is significant because it predicts *attention*, which affects *holdings*. One way to see this is through Table 5, that presents OLS regressions in which all variables

are included. Note that once *attention* is included as a regressor on the *holdings* equation, *female models* is no longer significant.

[Insert Tables 4 and 5 about here]

The estimation output from the second-stage regressions of the 2SLS procedure are on Table 6. The first column of results refers to the *attention* equation. In this equation, we used *market sophistication* and *market capitalization* as instruments for the *holdings* variable. We can see that the overidentification test does not reject the validity of the chosen instruments, the Hansen-Sargan statistic is 0.338, with a p-value of 56.1%. The model also has a good in-sample fit, the R^2 is 74.3%, and all estimated coefficients are significant. First, agents do pay more attention to countries at which they hold more assets. A 1% increase in the *holdings* of a particular country's asset is also associated with a 0.23% increase in the number of that country's click-throughs. Second, *attention* is affected by the popularity of a country's *female models*: an increase of 1 point in the *female models* popularity rank increases the number of click-throughs received by a country by 0.3%. Third, a 1% increase in a country's share of *protected areas* increases *attention* by 0.23%. Finally, the more developed a country, the higher the attention allocated to it: a 1% increase in *GDP per capita* increases the number of click-throughs by 0.57%.

[Insert Table 6 about here]

The second column of Table 6 presents the estimated coefficients for the *holdings* equation, in which we used *female models* and *protected areas* as instrumental variables for the *attention* allocated to each country. The overidentification test does not reject the validity of these three variables as instruments: the Hansen-Sargan statistic is 0.671, with a p-value of 41.3%. The in-sample fit is also good, with an R^2 of 70.7%. The coefficient associated to *attention* is significant at the 1% significance level: countries that receive 1% more click-throughs also have an increase of 0.72% on its asset *holdings*. Second, investors tend to invest more in more sophisticated financial markets: a 1% increase in the *market sophistication* increases asset *holdings* by 4.56%, with the coefficient being significant at 5%. Finally, even though the

coefficients associated with *market capitalization* and *GDP per capita* have the predicted sign, they are both not significant.

4.2 Multiple-equation 3SLS

The results of the 3SLS estimation are reported on Table 7. We can notice that all coefficients have the same sign, and very similar magnitudes compared to the ones estimated by 2SLS. The only noticeable differences are with respect to the significance level of a couple of variables. The coefficient associated to *holdings* on the *attention* equation has its significance reduced from the 1% to the 10% level. On the *holdings* equation, while *attention* has its significance level decreased from 1% to 5%, *market sophistication* has its significance increased from the 5% to the 1% level.

[Insert Table 7 about here]

Since the system is over-identified, we can test the null hypothesis that all the instruments are valid. The Hansen-Sargan statistic obtained was 0.930, with a p-value of 62.8%. This implies that we cannot reject the orthogonality conditions assumed when estimating the system by 3SLS.

5 Conclusion

This paper introduces a model and finds empirical evidence on the interaction between international asset holdings and attention allocation. We analyze a new a database that includes over 21 million web queries from 657,426 America Online users. The data includes all searches and results from those users for a three month period as well as whether they clicked or did not click on a search result and where it appeared on the result page. In this paper, we explore the AOL dataset to help us measure the attention allocated to different countries. We believe that a measure based on the internet search queries can proxy attention allocation because: first, the World Wide Web is becoming the predominant information media; second, search engines are the most popular tool to help users find reliable information on the “Web” since it

minimizes the time required and the amount of information which must be consulted; third, by having the exact search query we can know what was the topic that the user was interested at.

Using the AOL dataset, we constructed an attention allocation variable, which measures the number of times a user clicked through a search result from a particular country. We combine the click-through series with data from the latest survey of U.S. portfolio holdings of foreign securities. We find empirical evidence using both 2SLS and 3SLS that the strong relationship between asset holdings and attention allocation presented by Figure 1 is actually a result of three factors. First, correlation caused by the level of development, which is captured by the positive effect of GDP per capita in both equations. Investors hold more assets in countries with more favorable investor-protection mechanisms and more stable political systems, and they also search more information from developed economies due to their more sophisticated websites or webservicees. Second, causality runs from asset holdings to attention allocation. Using instrumental variables that capture the level of development in financial markets, the economic proximity and the importance of the U.S. as a trade partner for that country, we find that agents pay more attention to countries at which they hold more assets. Third, causality also runs from attention allocation to asset holdings. Using instrumental variables that are related to a country's popularity among internet users, we find that international investors tend to hold more assets in countries from which they process more information.

References

- [1] Ahearne, A., W. Grier, F. Warnock (2004). Information costs and home bias: An analysis of U.S. holdings of foreign equities. *Journal of International Economics* 62, pp. 313-336.
- [2] Chan, K., V. Covrig, L. Ng (2005). What determines the domestic bias and foreign bias?. *Journal of Finance* 60, pp. 1495-1534.
- [3] Falkinger, J. (2007a). Attention Economies. *Journal of Economic Theory* 133, pp. 266-294.

- [4] Falkinger, J. (2007b). Limited Attention as a Scarce Resource in Information-Rich Economies. *Economic Journal*, forthcoming
- [5] Fallows, D. (2005). *Search Engine Users*. Washington, DC: Pew Internet & American Life Project.
- [6] Gabaix, X. and D. Laibson (2002). The 6D Bias and the Equity Premium Puzzle. In B. Bernanke, K. Rogoff (Eds.), *NBER Macroeconomic Annual*, pp. 257-312.
- [7] Gabaix, X., D. Laibson and G. Moloche (2003). The Allocation of Attention: Theory and Evidence. Working Paper.
- [8] Joachims, T. (2002). Optimizing Search Engines Using Clickthrough Data, *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD)*, ACM.
- [9] Luo, Y. (2007). Consumption Dynamics under Information Processing Constraints. Working paper.
- [10] Mackowiak, B., M. Wiederholt (2007). Optimal Sticky Prices under Rational Inattention. Working paper.
- [11] Mankiw, G., R. Reis (2002). Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics* 117, pp. 1295-1328.
- [12] Mondria, J. (2007) "Financial Contagion and Attention Allocation", Working paper.
- [13] Mondria, J., C. Quintana-Domeque (2006). Financial Contagion through Attention Reallocation: An Empirical Analysis. Working paper.
- [14] Mondria, J., T. Wu (2007). The Puzzling Evolution of the Home Bias, Information Processing and Financial Openness. Working paper.
- [15] Pass, G., A. Chowdhury, C. Torgeson (2006). A Picture of Search. *The First International Conference on Scalable Information Systems*, Hong Kong, June.
- [16] Portes, R. and H. Rey (2005). The Determinants of Cross Border Equity Transaction Flows. *Journal of International Economics* 65, pp. 269-296.
- [17] Sims, C. (2006). Rational Inattention: A Research Agenda. *The American Economic Review*, 96, pp. 158-163.
- [18] *The Global Competitiveness Report (2008)*. World Economic Forum.
- [19] *The Travel & Tourism Competitiveness Report (2008)*. World Economic Forum.

- [20] Van Nieuwerburgh, S., L. Veldkamp (2007a). Information Acquisition and Portfolio Under-Diversification. Working paper.
- [21] Van Nieuwerburgh, S., L. Veldkamp (2007b). Information Immobility and the Home Bias Puzzle. Working paper.

Figure 1: AOL Attention Allocation and International Investment

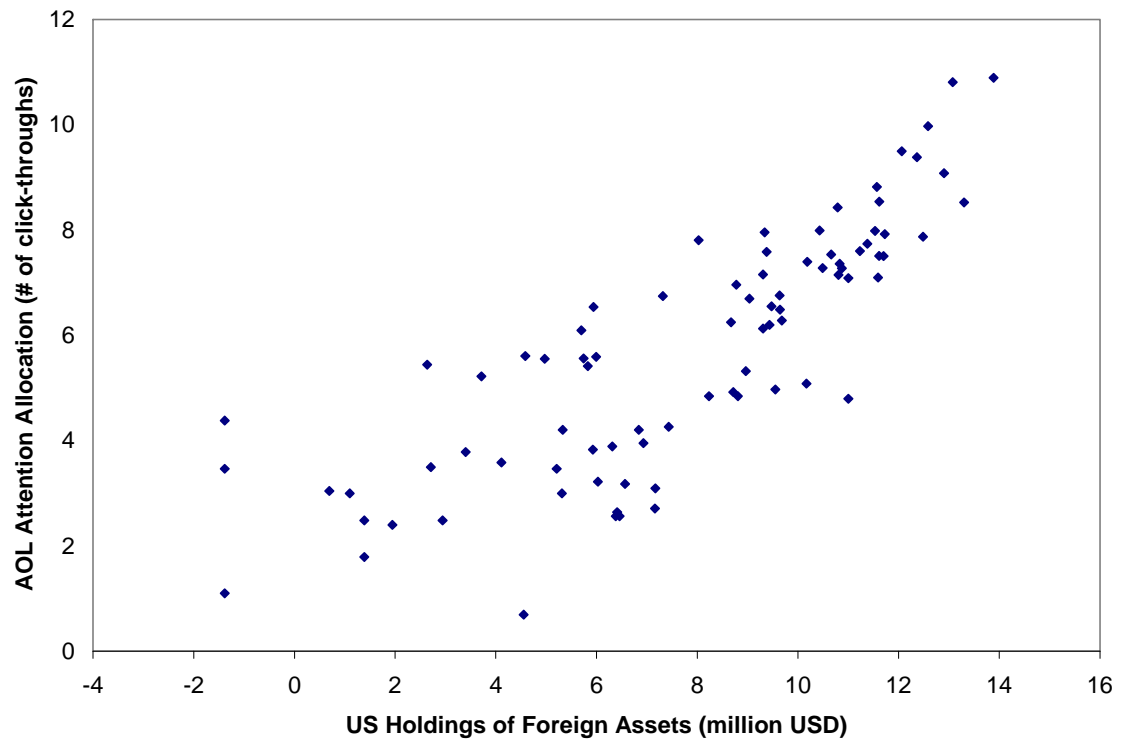


Table 1: Country Statistics

Country Name	AOL Attention	US Holdings of	Female Model	Protected Areas
	Allocation	Foreign Assets	Popularity Rank	(as % of total area)
	Mar-May 2006	Dec 2006	2006	2005
Argentina	1,281	11,019	0	6.2
Australia	13,331	173,153	228	16.7
Austria	1,632	26,568	0	28.0
Bangladesh	12	4	0	1.3
Belgium	1,447	35,968	0	3.3
Bolivia	32	0	0	19.8
Botswana	6	4	0	30.2
Brazil	1,824	110,294	412	18.7
Bulgaria	260	312	0	9.5
Canada	49,480	477,888	1,197	6.8
Chile	492	12,539	0	20.8
China	1,998	75,314	37	14.8
Colombia	517	5,827	190	31.6
Costa Rica	225	340	0	23.3
Croatia	268	400	0	6.5
Czech Republic	2,462	3,070	232	15.8
Denmark	2,957	33,978	0	7.1
Ecuador	49	552	0	19.4
Egypt, Arab Rep.	127	6,685	0	13.3
El Salvador	67	935	0	0.9
Estonia	272	98	0	31.0
Finland	1,196	59,934	0	7.8
France	8,765	401,388	81	11.7
Georgia	44	30	0	4.0
Germany	21,432	292,103	197	30.0
Greece	535	16,012	0	2.8
Guatemala	67	206	0	30.8
Hong Kong, China	2,291	87,518	0	51.5
Hungary	810	8,409	0	8.9
Iceland	204	7,829	0	5.6
India	1,271	49,231	54	5.4
Indonesia	144	14,072	0	9.1
Ireland	1,816	120,513	0	0.7
Italy	6,759	105,893	95	12.5
Jamaica	13	591	0	13.5
Japan	5,031	596,239	26	8.6
Jordan	36	61	0	10.9
Kazakhstan	15	1,288	0	2.9
Kenya	33	15	0	12.7
Korea, Rep.	2,762	123,876	22	3.9

(continued)

Table 1: Country Statistics (continued)

Country Name	AOL Attention	US Holdings of	Female Model	Protected Areas
	Allocation	Foreign Assets	Popularity Rank	(as % of total area)
	Mar-May 2006	Dec 2006	2006	2005
Latvia	231	14	0	13.9
Lithuania	185	41	0	10.6
Luxembourg	121	60,101	0	17.1
Malaysia	658	15,404	0	17.3
Mauritania	3	0	0	1.7
Mauritius	13	639	0	0.9
Mexico	1,208	108,450	208	8.7
Moldova	80	0	0	1.4
Morocco	25	415	0	1.2
Namibia	21	2	0	14.6
Netherlands	11,868	234,066	0	14.3
New Zealand	2,853	11,372	71	19.6
Norway	1,571	50,573	31	4.8
Pakistan	52	1,026	0	9.1
Panama	161	26,069	0	24.6
Peru	127	3,763	0	13.3
Philippines	458	10,989	0	6.5
Poland	1,966	11,816	40	27.1
Portugal	1,054	6,485	0	5.0
Romania	691	380	0	5.1
Russian Federation	4,571	48,441	194	8.8
Singapore	1,443	52,731	0	2.2
Slovak Republic	443	299	0	25.2
Slovenia	258	144	33	7.3
South Africa	1,870	42,686	298	6.1
Spain	5,107	110,957	178	7.7
Sri Lanka	32	183	0	17.2
Sweden	2,934	102,066	83	9.2
Switzerland	2,625	264,243	0	28.7
Thailand	700	13,047	0	19.0
Trinidad and Tobago	24	711	0	1.8
Tunisia	14	607	0	1.3
Turkey	859	15,284	0	3.9
Ukraine	851	1,506	54	3.4
United Kingdom	53,856	1,075,579	1,195	12.5
Uruguay	71	1,694	0	0.4
Venezuela, RB	137	6,097	0	63.0
Zambia	12	19	0	41.5
Zimbabwe	20	203	0	14.7

(continued)

Table 1: Country Statistics (continued)

Country Name	Financial Market Sophistication	Market Capitalization	GDP Per Capita
	(1 = poor, 7 = excellent)	(% of GDP)	(current USD)
	2007	2001-2005 average	2006
Argentina	3.7	53.39	5,472
Australia	6.4	104.70	37,434
Austria	6.1	23.95	39,131
Bangladesh	2.8	3.80	429
Belgium	6.1	96.25	37,384
Bolivia	2.8	19.90	1,195
Botswana	4.2	25.01	5,875
Brazil	6.0	40.84	5,660
Bulgaria	3.3	9.46	4,089
Canada	6.4	107.39	38,440
Chile	5.8	101.28	8,865
China	3.8	36.17	2,034
Colombia	4.2	21.82	2,982
Costa Rica	4.1	12.31	5,047
Croatia	4.3	23.72	9,612
Czech Republic	4.7	23.00	13,877
Denmark	6.3	57.34	50,702
Ecuador	3.3	7.59	3,042
Egypt, Arab Rep.	3.5	44.99	1,426
El Salvador	4.9	14.46	2,618
Estonia	5.8	35.07	12,237
Finland	6.0	114.58	39,856
France	6.2	79.98	36,547
Georgia	3.5	4.10	1,702
Germany	6.2	44.44	35,270
Greece	4.6	61.74	22,042
Guatemala	4.1	1.11	2,735
Hong Kong, China	6.6	424.63	27,072
Hungary	4.4	23.31	11,227
Iceland	5.7	104.47	53,029
India	5.3	43.93	817
Indonesia	4.1	21.97	1,634
Ireland	6.2	58.72	52,893
Italy	4.2	43.70	31,496
Jamaica	5.0	104.72	3,954
Japan	4.7	73.22	34,023
Jordan	4.5	141.85	2,538
Kazakhstan	3.7	9.26	5,045
Kenya	4.3	21.00	603
Korea, Rep.	5.2	59.85	18,341

(continued)

Table 1: Country Statistics (continued)

Country Name	Financial Market Sophistication	Market Capitalization	GDP Per Capita
	(1 = poor, 7 = excellent)	(% of GDP)	(current USD)
	2007	2001-2005 average	2006
Latvia	4.5	10.81	8,797
Lithuania	4.3	19.95	8,770
Luxembourg	6.5	129.14	89,564
Malaysia	5.5	145.49	5,780
Mauritania	2.5	97.19	844
Mauritius	4.9	34.14	5,145
Mexico	4.5	22.33	8,052
Moldova	3.2	23.73	850
Morocco	4.3	36.69	1,879
Namibia	4.9	6.30	3,107
Netherlands	6.4	103.21	40,167
New Zealand	5.7	38.45	25,179
Norway	6.0	47.74	66,964
Pakistan	4.0	22.44	810
Panama	5.7	25.32	5,205
Peru	4.6	28.88	3,288
Philippines	4.5	31.55	1,382
Poland	4.0	20.89	8,888
Portugal	5.7	37.82	18,185
Romania	3.7	12.21	5,645
Russian Federation	3.7	46.27	6,932
Singapore	6.3	149.84	30,082
Slovak Republic	5.0	8.67	10,223
Slovenia	4.6	22.62	18,674
South Africa	6.3	177.83	5,381
Spain	5.8	80.60	28,108
Sri Lanka	4.6	15.20	1,364
Sweden	6.4	99.25	42,553
Switzerland	6.8	232.26	51,033
Thailand	5.1	58.59	3,187
Trinidad and Tobago	4.4	94.57	15,214
Tunisia	4.3	10.18	2,990
Turkey	5.0	31.26	5,521
Ukraine	3.6	13.34	2,278
United Kingdom	6.3	135.35	38,850
Uruguay	3.5	1.72	5,828
Venezuela, RB	3.9	4.56	6,730
Zambia	3.6	8.29	919
Zimbabwe	3.7	65.57	383

Table 2: Summary Statistics

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
AOL Attention Allocation	2,926	492	8,576	3	53,856
US Holdings of Foreign Assets	65,041	7,829	157,452	0	1,075,579
Female Model Popularity Rank	65	0	200	0	1,197
Protected Areas	13.3	9.5	11.7	0.4	63.0
Financial Market Sophistication	4.8	4.6	1.1	2.5	6.8
Market Capitalization	56.32	36.17	62.78	1.11	424.63
GDP Per Capita	15,862	5,875	18,572	383	89,564

Table 3: Correlation Matrix

	<i>attention</i>	<i>holdings</i>	<i>female models</i>	<i>protected areas</i>	<i>market sophistication</i>	<i>market capitalization</i>	<i>GDP per capita</i>
<i>attention</i>	100%						
<i>holdings</i>	81%	100%					
<i>female models</i>	50%	36%	100%				
<i>protected areas</i>	21%	11%	8%	100%			
<i>market sophistication</i>	62%	69%	28%	19%	100%		
<i>market capitalization</i>	46%	52%	24%	7%	60%	100%	
<i>GDP per capita</i>	68%	65%	24%	9%	74%	48%	100%

Note: all variables in natural logs

Table 4: Two Stages Least Squares Estimation Output – First Stage

	(1)	(2)
Estimation method:	OLS	OLS
Dependent variable:	<i>attention</i>	<i>holdings</i>
Regressors:		
<i>female models</i>	0.004*** (0.001)	0.003*** (0.001)
<i>protected areas</i>	0.238 (0.148)	0.016 (0.248)
<i>market sophistication</i>	0.927 (1.377)	5.598*** (2.542)
<i>market capitalization</i>	0.165 (0.175)	0.397 (0.371)
<i>GDP per capita</i>	0.762*** (0.199)	0.779** (0.314)
Number of observations	79	79
R^2	61.3%	57.3%

Note: Each equation was estimated individually by ordinary least squares. White's robust standard errors are given in parenthesis under the coefficients. The symbols ***, **, and * denote that the individual coefficient is significant at the 1%, 5%, and 10% significance level respectively.

Table 5: Ordinary Least Squares with All Variables Estimation Output

	(1)	(2)
Estimation method:	OLS	OLS
Dependent variable:	<i>attention</i>	<i>holdings</i>
Regressors:		
<i>attention</i>	-	1.089*** (0.169)
<i>holdings</i>	0.353*** (0.062)	-
<i>female models</i>	0.0026*** (0.0004)	-0.001 (0.001)
<i>protected areas</i>	0.236* (0.134)	-0.252 (0.239)
<i>market sophistication</i>	-0.979 (0.818)	4.397** (1.824)
<i>market capitalization</i>	0.020 (0.135)	0.231 (0.278)
<i>GDP per capita</i>	0.485*** (0.149)	0.043 (0.263)
Number of observations	79	79
R^2	76.1%	72.9%

Note: Each equation was estimated individually by ordinary least squares. White's robust standard errors are given in parenthesis under the coefficients. The symbols ***, **, and * denote that the individual coefficient is significant at the 1%, 5%, and 10% significance level respectively.

Table 6: Two Stages Least Squares Estimation Output – Second Stage

		(1)	(2)
Estimation method:		2SLS	2SLS
Dependent variable:		<i>attention</i>	<i>holdings</i>
Regressors:	<i>attention</i>	-	0.721***
		-	(0.138)
	<i>holdings</i>	0.232***	-
		(0.117)	-
	<i>female models</i>	0.003***	-
		(0.001)	-
	<i>protected areas</i>	0.230*	
		(0.129)	
	<i>market sophistication</i>	-	4.560**
		-	(1.989)
	<i>market capitalization</i>	-	0.308
		-	(0.298)
	<i>GDP per capita</i>	0.565***	0.253
		(0.215)	(0.234)
Overidentification test			
Hansen-Sargan statistic		0.338	0.671
		(0.561)	(0.413)
Number of observations		79	79
R^2		74.3%	70.7%

Note: Each equation was estimated individually by two stages least squares. White's robust standard errors are given in parenthesis under the coefficients and p-values are given in parenthesis under the Hansen-Sargan statistic. The symbols ***, **, and * denote that the individual coefficient is significant at the 1%, 5%, and 10% significance level respectively.

Table 7: Three Stages Least Squares Estimation Output

Estimation method:		3SLS	
Dependent variable:		<i>attention</i>	<i>holdings</i>
Regressors:	<i>attention</i>	-	0.722**
		-	(0.294)
	<i>holdings</i>	0.233*	-
		(0.119)	-
	<i>female models</i>	0.003***	-
		(0.001)	-
	<i>protected areas</i>	0.217*	
		(0.119)	
	<i>market sophistication</i>	-	4.480***
		-	(1.640)
	<i>market capitalization</i>		0.325
			(0.256)
	<i>GDP per capita</i>	0.563***	0.256
		(0.212)	(0.332)
Overidentification test			
Hansen-Sargan statistic		0.930	
		(0.628)	
Number of observations		79	
R^2		74.7%	70.7%

Note: Both equations were estimated simultaneously by three stages least squares. Standard errors are given in parenthesis under the coefficients, and the p-value is given in parenthesis under the Hansen-Sargan statistic. The symbols ***, **, and * denote that the individual coefficient is significant at the 1%, 5%, and 10% significance level respectively.