Being in the Right Place: A Natural Field Experiment on the Causes of Position Effects in Individual Choice

Harris, Mark and Novarese, Marco and Wilson, Chris

Curtin University, Universit del Piemonte Orientale, Loughborough University

23 April 2019

Being in the Right Place: A Natural Field Experiment on the Causes of Position Effects in Individual Choice

Mark N. Harris, Marco Novarese, and Chris M. Wilson*

April 23, 2019

Abstract

This paper uses a natural field experiment to better understand why individuals tend to select items at the top of lists. After randomizing the order in which new economics research papers are presented in email alerts and measuring the subsequent downloads, we provide robust evidence of position effects. Moreover, our novel user-level data offers two key findings: i) most users exhibit both top and bottom position effects, and ii) distinct groups of users consider the listed items in different orders. These results allow us to conclude that the causes of top position effects are complex and heterogeneous across individuals, but are most consistent with a version of choice fatigue where users consider the listed items in a non-monotonic order.

Keywords: Position Effects; Order Effects; Choice Fatigue; Prominence; Lists

JEL Codes: D01, D83, L00

*Harris: Curtin Business School, Curtin University, Perth, Australia; mark.harris@curtin.edu.au. Novarese: Center for Cognitive Economics, Universit del Piemonte Orientale, Italy; marco.novarese@jp.unipmn.it. Wilson: School of Business and Economics, Loughborough University, UK; c.m.wilson@lboro.ac.uk. Corresponding author: Wilson. This paper is a substantially revised draft of an earlier, now defunct, version that only used aggregate data by Novarese and Wilson, entitled 'Being in the Right Place: A Natural Field Experiment on List Position and Consumer Choice'. Novarese is the director and general editor of New Economics Papers (NEP). We acknowledge the financial support of the Australian Research Council (Harris) and the Regione Piemonte, 'Bando per le Scienze Umane e Sociali', under the project 'Istituzioni, comportamento e mercati in contesti locali e globali' (Novarese). Thanks to the seminar participants and discussants at the Economic Science Association, the Internet Search and Innovation Conference at Northwestern, the Search and Switching Costs Workshop at Groningen, IIOS, EARIE, JEI, and UEA. We also thank RePEc, NEP, and the participating editors, as well as Chris Jagau and Cesare Tibaldeschi for their assistance.

1 Introduction

When investigating search results, choosing products from websites, considering job listings, or using a comparison site, individuals frequently make choices from lists. It is well known that when faced with such lists, individuals often show a disproportionate tendency to select the item in top position. This is evident from the large expenditures that firms pay for sponsored links, the recent antitrust cases into Google's alleged bias[1], and the findings from a broad range of rigorous academic studies. For example, as later reviewed, the literature has shown that demand increases markedly for firms at the top of search results, investors trade more frequently with stocks at the top of investment listings, consumers are more likely to select items at the top of menus, and voters are more inclined to choose candidates at the top of ballots.[2] However, the explanations for such choice-based 'top position effects' or 'primacy effects' remain far less clear. Are top-placed options more likely to be selected simply because higher quality options have been positioned at the top? If not, why might individuals show a systematic tendency to select top- positioned items? Insights into these questions would have a broad range of policy implications, and help understand several important market issues regarding firms and platforms (e.g. Armstrong 2017).

However, when analyzing top position effects, researchers face a methodological trade-off. Standard field data offers high external validity but often suffers from an identification problem because the order of items is normally endogenous, and potentially further complicated by the presence of price setting incentives. Alternatively, experimental data can enable easier identification but has lower external validity. In this paper, we take a middle ground by conducting a natural field experiment that facilitates an exogenous item order in a setting with no pricing incentives, while maintaining a moderate level of external validity.

In particular, this paper analyzes the causes of top position effects by using a natural

[1] See `https://www.nytimes.com/2019/03/20/business/google-fine-advertising.html`, accessed 23/04/19.

[2] Recent examples include Fedyk (2019), Ursu (2018), Feenberg *et al.* (2017) and Meredith and Salant (2013).

field experiment with a group of subjects that should be the least likely to depart from standard theory - economists. Economics research papers are often available on a well-known online database, Research Papers in Economics (RePEc). Many economists choose to be kept informed of recent additions to the database by subscribing to a free email alert service conducted by New Economic Papers (NEP). The service regularly compiles lists of new papers and arranges them into descending order of estimated popularity. This setting offers an excellent experimental environment because RePEc carefully records the subsequent download activity for each item on each list. Specifically, we measure the download activity on a control group of lists that retain their usual ordering of papers, and on a treatment group of lists that have their papers re-arranged into a random order.

Our findings document the presence of robust position effects even within the treatment group. Moreover, by exploiting some novel dis-aggregate features of our data, we go beyond the existing literature by showing i) how most users exhibit both top and bottom position effects at the individual level, and ii) how distinct groups of users consider the listed items in different orders. These findings allow us to conclude that the causes of top position effects are complex and heterogeneous across individuals, but are most consistent with a version of choice fatigue where users consider the listed items in a non-monotonic order.

After further detailing the experiment in Section 2, Section 3 outlines our three main explanatory hypotheses for top position effects (as defined where items in the top position are significantly more likely to be downloaded than items in other positions). Under Specific Item Order (H1), top position effects arise only because an item with a relatively high value happens to be in top position. Under Value Signals (H2), users cannot fully assess the quality of items but are more likely to select top-positioned items because they expect (perhaps incorrectly) that the items are arranged in descending order of value. Under Choice Fatigue (H3), top position effects exist because users find top-positioned items less costly to inspect or select. This is consistent with users who i) find it increasingly costly to consider individual items, and ii) evaluate top-positioned items relatively early in their decision process when they are most fresh. We specify some different forms of this hypothesis depending on the exact order in which users consider the listed items, H3a-H3d. For instance, H3a predicts top position effects by suggesting that users consider the items in a strict descending order from the top downwards.[3]

Section 4.1 presents our initial empirical results by investigating how list position influences users' download decisions. First, as expected within the control group, items in the top position are significantly more likely to be downloaded. However, highly significant top position effects are also evidenced within the treatment group despite the order of items having been randomized. Hence, top position effects cannot be solely explained by the specific order of items (H1). Nevertheless, as randomization significantly weakens the size of the estimated top position effects by approximately 30%, H1 does offer an important partial explanation.

Second, beyond top position effects, the data in both the control and treatment groups is also characterized by some relatively smaller but strongly significant positive effects for items in bottom position. Such 'bottom position effects' contradicts Value Signals (H2) and the simple form of Choice Fatigue (H3a) because they are not consistent with users who expect items to be arranged in descending value order or users who becoming increasingly tired as they work sequentially down a list. This finding also rules out the possibility that top position effects exist in the data purely because NEP typically sorts the items in descending value.

To go beyond these initial results, Sections 4.2 and 4.3 then utilize two novel user-level features of our dataset to reveal some key results. First, we exploit the user-level aspect of our data to conduct a set of random parameter estimates that allow the estimated position effects to vary across users. This shows that very few users exhibit position effects that monotonically increase or decrease with position. Moreover, it suggests that the documented top and bottom position effects are largely an individual-level phenomenon: 75% of users

[3]As later reviewed, choice fatigue was first introduced into this literature by Augenblick and Nicholson (2012).

display *both* top and bottom position effects. Second, we exploit another novel feature of our data which records the exact time that each download was made (to the nearest second). We use this information to recover the order in which users made their downloads in instances where they selected multiple items from a list. Crucially, we find evidence of two distinct large groups of users: 40-43% of users always download their items in a strict *descending* order from top down, while 52-58% of users show no systematic monotonic order. In contrast, only 0-2% of users always download their items in a strict *ascending* order, and only 2-3% of users vary between using a strict ascending or descending order.

Hence, after accounting for the important partial explanatory role of H1 (Specific Item Order), our novel user-level results suggest that the causes of top position effects are complex and heterogeneous across individuals. Nevertheless, our findings are most consistent with a version of choice fatigue where users consider the listed items in a non-monotonic order (H3d). This is akin to Feenberg *et al.*'s (2017) explanation of 'skimming' where individuals focus on items in the prominent top and bottom positions, and is also in line with the theoretical predictions of Fishman and Lubensky (2018). For instance, users may make their initial download decisions from top position downwards, but then potentially return up the list to reconsider some items that they did not download previously. Unlike our other considered hypotheses, this favored explanation is consistent with both our findings of i) a large group of users who download their items in a non-monotonic sequence, and ii) users who typically display both top and bottom position effects. While less strongly supported, our results also indicate that top position effects may be further accentuated by a second group of users who exhibit a different form of choice fatigue by always considering their items in a strict *descending* order, H3a.

Previous Literature: The existence of top and bottom position effects has been previously well documented in a variety of contexts, but our paper focuses on position effects in individual choice from visually presented lists.[4] Moreover, in addition to carefully documenting the

[4]Other contexts include how individuals i) form impressions or judgments (e.g. Asch 1946), ii) evaluate

existence of such position effects, we differ from much of the previous literature by testing between different explanations.

In reviewing the literature, we now classify studies into two settings: i) limited selection, and ii) unlimited selection. In limited selection, individuals may only select one item (or some other fixed number of items) from a list. This is the most common setting in previous research, but differs from our 'unlimited selection' setting where i) there is no inherent constraint on the number of items an individual can select, and ii) the items are sufficiently non-substitutable that individuals often wish to select multiple items. In addition to our download environment, other everyday examples of this setting include choosing articles to read from a news aggregator, browsing amongst different items on a website, or selecting items from a bestsellers list.

i) Limited Selection: This setting includes many 'market' studies that use data from online search results. As search results often place the most relevant items first, researchers must employ some method to rule out a simple explanation of Specific Item Order (H1). To do this, some studies use a variety of econometric techniques.[5] A notable exception is Ursu (2018) who randomizes the order of search results at an online travel agent. She quantifies the effects of rankings with a consumer search model and shows that position effects are significant but lower than typically estimated. Murphy *et al.* (2006) and Dayan and Bar-Hillel (2011) also use randomization in restaurant websites and menus, respectively. Unlike the papers on search results, they also test for, and provide evidence of, bottom position effects. However, contrary to our paper, none of these market studies focus on testing different explanations of position effects.

Other papers show how voters tend to select the candidate placed at the top of a ballot.[6] As legislation often requires ballot orders to be (quasi-) random, these results cannot be

alternatives in contests or product sampling tests (e.g. Haan *et al.* 2005; Biswas *et al.* 2010), iii) respond in surveys (e.g. Schwarz *et al.* 1992) and iv) recall items in memory tasks (e.g. Tan and Ward 2000).

[5] For example, Ansari and Mela (2003), Narayanam and Kalyanam (2015), Baye *et al.* (2016a, 2016b), De los Santos and Koulayev (2016).

[6] For example, Miller and Krosnick (1998), Koppell and Steen (2004), Ho and Imai (2008), Meredith and Salant (2013).

explained by Specific Item Order (H1). Instead, most papers jump to an explanation of satisficing (Simon 1955) where individuals make costly inspections of items sequentially from the top downwards, and optimally stop to select an item that is sufficiently attractive. However, by exploiting some features of multi-winner elections, this explanation is rejected by Meredith and Salant (2013). Augenblick and Nicholson (2012) consider a different setting where voters have to vote on multiple different contests within the same ballot. Consistent with voters depleting their cognitive resources as they work down the ballot paper, they show that voters become more likely i) to abstain, ii) vote for the default option, or iii) display a bias towards candidates listed first. Augenblick and Nicholson refer to this as 'choice fatigue'. In contrast, we use some novel features of our data to analyze some specific forms of choice fatigue as explanations for top (and bottom) position effects within our alternative context.

ii) Unlimited Selection: Some evidence within this setting comes from finance. As lists of stocks are often presented in alphabetical order, Itzkowitz *et al.* (2016) and Jacobs and Hillert (2016) show that firms with earlier names have higher trading activity, while Fedyk (2019) documents how top placed news articles at Bloomberg can lead to substantially higher trading volumes and price changes for associated financial assets.

Other evidence comes from academia. Pinkowitz (2002) and Coupe *et al.* (2010) use clever strategies to show a partial role for Specific Item Order (H1). Pinkowitz (2002) uses data from the Journal of Finance website where individuals can download fully published papers as well as accepted papers that have yet to be allocated to an issue. Consistent with H1, papers that are later allocated a top position receive significantly more downloads before being assigned their position. However, consistent with other explanations, such papers also receive an additional download effect after being listed first. Alternatively, Coupe *et al.* (2010) show top position effects exist within issues of the European Economic Review even when the order of papers is determined alphabetically rather than by the editor. Closest to our research is the excellent paper by Feenberg *et al.* (2017) who use the random ordering of NBER paper alerts to show top and bottom position effects in individuals' download and citation activity. Among other results, they suggest that the most consistent explanation is 'skimming' where, similar to our H3d, time-constrained individuals focus on salient positions such as top and bottom. In contrast, while our NEP alerts have a lower readership than NBER alerts and are therefore less influential on citation activity, our data contains rich dis-aggregate information on download decisions at the *user-level*. This allows us to employ random parameter techniques and analyze the timing of downloads in order to i) further test between competing hypotheses, ii) provide different forms of evidence, and iii) better allow for heterogeneity in explanations across users.

Finally, our paper also makes a secondary contribution to a literature on search behavior. For instance, by using data on consumers that click on more than one online search result, Jeziorski and Segal (2015) demonstrate that less than half of such consumers make their clicks in a monotonic descending order. Alternatively, within a sophisticated search-theoretic laboratory experiment, Caplin *et al.* (2011) provide some persuasive support for satisficing while also showing that some subjects inspect items in an ascending rather than descending order. Within our different setting, we also document some related patterns of behavior, but use such evidence to analyze the causes of top position effects.[7]

2 Setting, Experiment, and Data

2.1 RePEc and NEP

Research Papers in Economics (RePEc) is a popular online database of economics research papers. As part of RePEc, New Economics Papers (NEP) offers a free email alert service to notify individuals about new papers that have been recently added to the RePEc database. Such alerts are often provided on a weekly basis and are generated for separate research subfields, such as health economics or monetary economics. Subscribers can select which

[7] Tests of standard search theory, such as De los Santos *et al.* (2012), are less relevant to our paper as they use settings without a pre-defined list or search order. However, as potentially consistent with our findings, they do find that individuals often go back to select a previously searched option.

subfields they wish to subscribe to and NEP has over 80000 total subscriptions.[8]

Each email alert has two sections of text. An extract from an example alert is provided in Section 1.1 of the Supplementary Appendix. The top section states how many papers are included in the alert and presents a brief list of the papers with their titles and authors. If a reader clicks on the title of any paper within the list, or scrolls down, she is taken to the bottom section of the alert. The bottom section repeats the same list of papers but with additional summary information including each paper's abstract, keywords, JEL classification codes, date (if these are available) and most importantly, a link to a full text version of each paper. By clicking on a paper's link, a new window is opened and the paper is downloaded.[9]

The alerts for each subfield are managed by an editor, who is a volunteer from academia or the public sector. Although never made explicit to subscribers, the list of papers within each alert is compiled as follows. First, NEP gathers a master list of all new papers that have been recently added to the RePEc database. An algorithm then uses past data together with information about each paper's title and abstract to arrange the papers into descending order of estimated popularity. This master list is then passed to the subfield editors for them to extract the papers that are relevant for their next subfield alert. After selecting their relevant papers, each editor is free to amend the order in which the papers are presented within their alert or leave them in the order suggested by the algorithm. Most editors amend the order of their lists with the intention of further improving upon the algorithm's attempts to put the more interesting and relevant papers towards the top.

As later discussed in more detail, papers can be selected to be in the alert of more than one subfield. Therefore, to avoid confusion, we will now make a distinction between 'papers' and 'items'. An item will refer to an entry on a specific alert, whereas a paper will refer to

[8] For more, see `http://nep.repec.org/`, accessed 23/04/19.

[9] Given the importance of bottom position effects within our later analysis, one may ask whether users are artificially drawn to the bottom item via the two-section design of the alerts. However, this is not the case. When inspecting the summary information of the top item in the lower section of the alert, the bottom item within the upper section of the alert is off-screen.

the underlying piece of research that can appear as an item in multiple subfield alerts. For ease of exposition, we will also refer to ‘alerts’ and ‘lists’ interchangeably.

RePEc measures the download activity for each item in an extremely precise manner. First, it measures downloads that occur specifically via the links contained within NEP alerts, not just those that occur through RePEc more generally. Second, in cases where a paper appears in multiple subfield alerts, RePEc records the downloads within each separate alert. Hence, the measurement of downloads is item-specific, not paper-specific, such that the relationship between list position and subsequent download activity can be analyzed in a meaningful manner. Finally, for each download, RePEc records the individual device (anonymized ip address) to which the download was made, and the time at which the download was initiated (to the nearest second). Thus, even when there are multiple devices being used within the same institution, RePEc carefully records download decisions and download timings at the level of each individual device.

2.2 Experimental Procedure

After requesting permission from NEP, we were granted access to the download data for the alerts released over a 5-month period across 29 subfields.[10] Moreover, we were given permission to manipulate the order in which the items were presented for a small proportion of alerts. To do this, we asked NEP and the relevant editors to continue collecting and ordering their alerts as they would do under normal circumstances. However, before the release of any given alert, we intervened and randomly allocated the alert into one of two groups. Within each subfield, around two-thirds of the alerts were allocated to a control group and the remaining alerts were allocated to a treatment group. Any alert within the control group was sent to subscribers with no alterations - the list of items was left completely unchanged. In contrast, any alert within the treatment group had its list of

[10]The range of subfields appear representative and cover a wide range of different areas of economics including areas like Cognitive and Behavioural, Time Series and Post Keynesian. A full list is provided in Section 1.2 of the Supplementary Appendix.

items rearranged into a new random order. An identical email with the revised order was sent to all subscribers. Beyond this, no changes were ever made to the content or presentation of the alerts. Importantly, the subscribers were left unaware of the experiment and that some lists were being randomized.

2.3 Data

Our analysis considers how download activity is related to four list positions within the email alerts: top, second, second-from-bottom, and bottom. As these positions are ill-defined in lists with less than four items, we drop the 43 such alerts from our initial sample. This leaves a final sample of 530 alerts, including 350 alerts within the control group and 180 alerts within the treatment group.

Some summary statistics are provided in Table 1 (all tables and figures are included in the main Appendix unless otherwise stated). Across the 530 alerts, the sample covers a total of 6624 items with an average of 12.5 items per alert. The 6624 listed items stem from 4942 different papers such that an average paper appears on 1.33 subfield alerts within our sample (or 3.90 subfield alerts across all of NEP). We later address this feature of the data within our estimation procedures.

Table 1 also uses NEP's item-specific download measures to record the aggregate number of downloads made from the release date of each item's subfield alert until a single cut-off date, almost two years later. This measurement period is sufficient to cover all relevant downloads as most downloads are made within a few weeks after the alert is released. However, the use of a single cut-off date does imply that alerts with different release dates are monitored for slightly different lengths of time. Our analysis later controls for this.

Within the sample, downloads were made from 9364 ip addresses. To ease exposition, we broadly refer to an ip address as a 'user'. After deleting a handful of duplicative cases whereby the same user downloaded the same item more than once, we end up with a total of 35004 downloads.

In subsequent sections, we often combine the download data with a range of alert-specific and item-specific control variables. These as summarized in Table 2. The alert-specific control variables include the total number of items within the alert and a measure of each alert's 'availability' - the number of days between the alert's release date and the final download cut-off date. The item-specific (or paper-specific) control variables are constructed from each item's summary information. They include variables related to an item's title language, title length, number of authors, abstract length, number of keywords, number of JEL codes, and the total number of lists (within the entire population of NEP) in which the item's underlying paper appeared.

3 Explanatory Hypotheses

We now outline our three main explanatory hypotheses for top position effects, H1-H3. As a backdrop, we conceptualize the decision environment as follows. For any given alert or 'list' l with $n_l \geq 4$ items, we define the position of item j as $p_j \in \{1, ..., n_l\}$, where $p_j = 1$ if item j is in top position, and $p_j = n_l$ if item j is in bottom position. We then categorize any given user i's value of downloading item j into three components. The first component refers to the 'observable value'. User i can quickly and freely assess this value component from item j's summary information, such as its title and authors. In contrast, the second component refers to the 'unobservable value'. This value component relates to the underlying quality of the item and cannot be assessed until after user i has downloaded and read the item more carefully. Finally, the third component refers to the 'inspection value'. This value component can be assessed before the item is downloaded but only if user i makes a costly inspection of item j's additional summary information from the lower section of the alert, such as its abstract.

Our three explanatory hypotheses, H1-H3, each revolve around one of the three discussed value components. In what follows, we define top position effects to exist when items in the top position are significantly more likely to be downloaded than items in other positions.

H1: Specific Item Order. Top position effects exist because items in the top position have a relatively large observable value.

This rather trivial explanation suggests that top-positioned items are more likely to be selected only because they happen to have a relatively large observable value.

H2: Value Signals. Top position effects exist because users believe (perhaps incorrectly) that the items have been arranged in descending order of unobservable value.

In contrast, H2 suggests that top-positioned items are more likely to be downloaded because users believe that some better informed agent has arranged the items in order of unobservable value.

H3: Choice Fatigue. Top position effects exist because users find that top-positioned items are less costly to inspect or select.

Under the assumption that users find it costly to inspect and select items, H3 is built on the ideas that users may only consider a subset of the listed items and that top-positioned items are more likely to be in this subset because they are less costly to consider. In particular, H3 is consistent with a user who i) exhibits total inspection/selection costs that are convex in the number of inspections/selections they make, and ii) considers top-positioned items earlier than other items within their decision process.[11]

3.1 Specific Forms of Choice Fatigue

For the later analysis, we now outline a number of specific forms of choice fatigue which depend upon the exact order in which users consider the listed items, H3a-H3d. In addition to top position effects, some of these forms also predict the simultaneous presence of bottom position effects (where items in the bottom position are also significantly more likely to be downloaded than items in other positions).

[11] In this sense, choice fatigue is related to the idea of convex search costs (e.g. Ellison and Wolitzky 2012).

H3a: Choice Fatigue with Descending Monotonic Direction. Top position effects exist because users consider the items in a strict descending order from top position downwards.

Assuming that users have increasing inspection/selection costs, H3a suggests that top position effects exist because users consider the items sequentially from top downwards. Hence, items in the top position are more likely to be downloaded because they are considered first when users are still fresh and when inspection/selection costs are at their lowest.

H3b: Choice Fatigue with Heterogeneous Monotonic Direction. Top position effects co-exist with bottom position effects because some users consider the items in a strict descending order from top downwards while others consider the items in a strict ascending from bottom upwards.

Under the assumption that users have increasing inspection/selection costs, H3b predicts the simultaneous existence of top and bottom position effects by recognizing the potential heterogeneity in users' behavior. It suggests that top position effects derive from a group of users who consider the items in a descending direction, while bottom position effects arise from a *different* group of users who consider the items in an ascending direction.[12]

H3c: Choice Fatigue with Varied Monotonic Direction. Top position effects co-exist with bottom position effects because each user varies between considering the items in a strict descending or ascending order.

Assuming that users have increasing inspection/selection costs, H3c predicts that an individual user may display both top and bottom position effects. It suggests that each user may vary the order in which they consider a list depending on the context - a user may employ a descending order in some lists, but an ascending order in others.

H3d: Choice Fatigue with Non-Monotonic Direction. Top position effects co-exist with bottom position effects because users consider the items in a non-monotonic order.

[12] The possibility of users considering items from bottom position upwards may seem odd. However, such behavior is later evidenced directly in Section 4.3.1, and is experimentally supported by Caplin *et al.* (2011).

H3d is consistent with Feenberg *et al.*'s (2017) explanation of 'skimming' where users focus on items in prominent positions, such as top and bottom. Given increasing inspection/selection costs, one example of a non-monotonic order may involve users i) making their initial download decisions from top position downwards, but then ii) potentially returning up the list to reconsider some items that they did not download previously. Such users may then display both top and bottom position effects because items towards the top are initially easier to consider, but once a user reaches the bottom, lower-placed items become relatively less costly to consider instead.[13][14]

4 Analysis

This section takes the presented explanatory hypotheses to the data. It is structured around four empirical tests, I-IV, and their associated analysis.

4.1 Empirical Tests I and II

We first consider Empirical Tests I and II in order to assess H1 (Specific Item Order), H2 (Value Signals) and H3a (Choice Fatigue with Descending Monotonic Direction).

Empirical Test I: Comparison of the Control and Treatment Groups. Reject H1 if significant top position effects remain within the treatment group.

Intuitively, top position effects under H1 (Specific Item Order) only exist because an item with a relatively large observable value has been placed in top position. Consequently, under H1, any such effects should only arise within the control group where the items have

[13] Fishman and Lubensky (2018) provide some related theoretical results. By building on Janssen and Parakhonyak (2014), they consider a limited selection setting where individuals face i) positive inspection costs, and ii) positive return costs to reconsider previously inspected options. They show that options at both the start and the end of a sequence are more likely to be selected.

[14] H3d can also be motivated from a bounded rationality perspective if users employ heuristics to economize on cognitive resources and such heuristics exhibit top and bottom position effects by making the items in the top and the bottom position appear more salient. Salant (2011) provides some related theoretical results within a limited selection setting to show that any heuristic that is procedurally simpler than rational choice displays top and bottom position effects.

been deliberately ordered, and not within the treatment group where the item order has been randomized. In contrast, any evidence of top positions within the treatment group cannot be used to rule out the other explanations as users might still continue to i) hold (now incorrect) beliefs that top-placed items have high value (H2), ii) find top-positioned less costly to consider (H3).

Empirical Test II: Analysis of Other Position Effects. Reject H2 and H3a if items in some position p, are significantly more likely to be downloaded than items in some preceding position, $p' < p$.

This test is based on analyzing a broader set of position effects beyond top position. In particular, under H2 (Value Signals) and H3a (Choice Fatigue with Descending Monotonic Direction), download activity is predicted to be decreasing in position because i) users expect item values to be decreasing from top position downwards, or ii) users consider the items in a descending direction. Hence, H2 and H3a can be rejected if there are significant instances where download activity is increasing in position.

4.1.1 Descriptive Results

For an initial descriptive analysis, we first consider aggregate downloads. In particular, Table 3 and Figure 1 show how the aggregate number of downloads per item varies with list position within the control and treatment groups. First, as expected, there are strong top position effects in the control group: top-positioned items receive 57% more aggregate downloads than an average item. Second, while the randomization of item order reduces this effect by around a quarter, top-positioned items still receive 42% more downloads than average within the treatment group. Hence, with the use of Empirical Test I, Specific Item Order (H1) offers an important partial explanation of top position effects, but it cannot be the main explanation. Third, if we consider a broader set of positions while ignoring bottom position (as often done in some parts of the existing literature), then downloads appear to be strictly decreasing in item position. However, contrary to this pattern, across both the control and treatment groups, bottom-positioned items attract 22-26% more downloads than items in the preceding, second-from-bottom, position. Moreover, this effect is so large that items in the bottom position actually receive 23% more downloads than average in the treatment group, and even receive 9% more downloads than average in the control group where the bottom item has the lowest predicted popularity. Using Empirical Test II, the presence of such bottom position effects contradicts Value Signals (H2) and Choice Fatigue with Descending Monotonic Direction (H3a), and also rules out the possibility that top position effects exist in the data purely because NEP typically sorts the items in descending value. Instead, such patterns are more consistent with the other forms of choice fatigue, H3b-H3d.

4.1.2 Random Effects Estimations

To consider Empirical Tests I and II more deeply, we now provide a more rigorous analysis of how list position affects download activity. Such an analysis could be done in several ways. For instance, one could continue to use the aggregate download data to estimate how list position affects the total number of downloads received by each item. Alternatively, one could investigate the data at a dis-aggregated 'user' level to estimate how an active user's decision to download an item is affected by its list position. To demonstrate the robustness of our results, we take both approaches. In the main text, we now focus on a dis-aggregate analysis in order to i) control for unobservable user-effects, and ii) provide a useful foundation for later sections. However, we also show that the main conclusions of this subsection remain robust under the alternative aggregate approach within Section 1.5 of the Supplementary Appendix.

To conduct the dis-aggregate analysis, we construct a dataset of active user download decisions. In particular, for any list l where user i has downloaded one or more items, we construct n_l user-item level observations where d_{ipl} equals one if user i downloaded the item in position p of list l, and zero if not. For example, if a user downloaded the first two papers from a list of four items, four observations would be created with values of d_{ipl} equal to $\{1, 1, 0, 0\}$ respectively. After applying this procedure over all users and all lists, we finish with a dataset of 288,788 user-item level observations.

To understand how users' download decisions are influenced by item position, we estimate a multi-dimensional random effects (RE) probit model. In particular, to consider user i's decision of whether to download the item in position p of list l we construct the following propensity variable where, as usual, $d_{ipl} = 1$ if and only if $d^*_{ipl} > 0$.

$$d^*_{ipl} = \beta_0 + \Pi'\beta_\Pi + z'_l\beta_z + q'_{pl}\beta_q + \psi_i + \mu_{pl} + \varepsilon_{ipl}. \tag{1}$$

The vector Π includes a set of position dummies for items in the top, second, second-from-bottom, and bottom positions. Any position effects will then be captured by the estimated values within $\beta_\Pi = \{\beta_{top}, \beta_{sec}, \beta_{secbot}, \beta_{bot}\}$. The vectors z_l and q_{pl} include the list-specific and item-specific control variables that were presented in Section 2.3. To control for user heterogeneity and the fact that some papers are included on more than one alert, we then include random effects at two levels. First, we include a 'user random effect', ψ_i, to capture the unobservable effects of an individual ip address. Second, we include a 'paper random effect', μ_{pl}, to control for the unobservable effects of the underlying paper in position p of list l.

After estimating equation (1) on the control and treatment groups separately, we formally examine how the estimated position effects differ between the control and treatment groups by re-estimating the equation on the full sample with the following additional variables: $treat_l$ - a dummy variable that equals one only if list l is in the treatment group, and $\Pi' * treat_l$ - a vector of interacted position terms.

Table 4 presents the results (with some further technical explanation), while Figure 2 plots the estimated position effects for the control and treatment groups for an example specification. The main results can be summarized as follows.[15]

[15]The estimated effects of the control variables are presented and discussed separately in Section 1.3 of

First, as expected within the control group, items in the top position are significantly more likely to be downloaded. However, highly significant top position effects also remain within the treatment group despite the order of items having been randomized. Hence, echoing the descriptive results, this rules out Specific Item Order (H1) as a full explanation, but shows that H1 plays an important minor explanatory role. In particular, as randomization significantly weakens the size of the estimated top position effects by approximately 30%, H1 appears to account for a third of the documented position effects within the control group.

Second, there are smaller, yet significant, position effects for items placed in second position and bottom position in both the control and treatment groups. Indeed, randomization has *no* significant effect in reducing the size of the bottom position effect. Thus, the estimated position effects are not strictly decreasing in size from top to bottom. Some formal tests at the bottom of Table 4 confirm that the position effects are strictly decreasing from top to second, and from second to second-from-bottom, but not from second-from-bottom to bottom. Hence, via Empirical Test II, this rules out H2 (Value Signals) and H3a (Choice Fatigue with Descending Monotonic Direction) as full explanations - top position effects cannot exist just because users expect items to be arranged in descending value order (H2), or because users consider the items in descending order.

Finally, while the estimated 'paper random effects' are only significant with the control group, the reported 'user random effects' are heavily significant across all cases. This implies substantial user heterogeneity - as further explored in the remaining sections.[16]

4.2 Empirical Test III

Next, we introduce and evaluate Empirical Test III in order to test explanatory hypothesis H3b (Choice Fatigue with Heterogeneous Direction). To recall, H3b predicts the simultaneous existence of top and bottom position effects by suggesting that top position effects

the Supplementary Appendix.

[16] Extended estimations show how the documented top position effects become more pronounced in longer lists. However, we place less weight on this part of the analysis. For details, see Section 1.4 of the Supplementary Appendix.

derive from a group of users who consider the items in a descending direction, while bottom position effects arise from a different group of users who consider the items in an ascending direction.

Empirical Test III: Other Position Effects and User Heterogeneity. Reject H3b if download activity is not decreasing from top position downwards for some users and increasing from bottom position upwards for others.

Under H3b, download activity should be decreasing in position for those users who always employ a descending order, but increasing in position for those users who employ an ascending order. In contrast, no such fixed monotonic patterns are required at the user-level under Choice Fatigue with Varied Direction (H3c) where any given user may vary the order in which they consider a list, or under Choice Fatigue with Non-Monotonic Direction (H3d) where users may consider the items in a non-linear order.

4.2.1 Random Parameter Estimations

Clearly, Empirical Test III hinges on how the estimated position effects vary across different users. Hence, rather than using our previous random effects model, (1), we now further exploit our user-level data to estimate a random parameters model. As illustrated in (2), this allows the set of estimated position effects, $\beta_{i,\Pi}$, to vary across each user i:[17]

$$d^*_{ipl} = \beta_0 + \Pi'\beta_{i,\Pi} + z'_l\beta_z + q'_{pl}\beta_q + \varepsilon_{ipl} \tag{2}$$

In particular, the vector of position effects, $\beta_{i,\Pi}$ is now specified as $\beta_{i,\Pi} = \bar{\beta}_\Pi + \alpha_{i,\Pi}$ where $\bar{\beta}_\Pi$ is a vector of average coefficients, and each element in the vector, $\alpha_{i,\Pi}$, follows a normal distribution with a zero mean and variance to be estimated.

Table 5 presents the key results by summarizing some user-level features of the estimated

[17] We focus only on user heterogeneity for these estimations. Adding additional paper-level heterogeneity to the random parameters makes little difference to the results and only complicates interpretation.

random parameters, $\beta_{i,\Pi} = \{\beta_{i,top}, \beta_{i,sec}, \beta_{i,secbot}, \beta_{i,bot}\}$.[18] First, within the treatment group, as initially consistent with Choice Fatigue with Heterogeneous Direction (H3b), 64% of users are estimated to have their largest position effect in top position, $max\{\beta_{i,\Pi}\} = \beta_{i,top}$, while 16% of users are estimated to have their largest position effect in bottom position, $max\{\beta_{i,\Pi}\} = \beta_{i,bot}$. Second, however, in contrast to H3b, only one percent of users are estimated to have position effects that monotonically decrease with position, $\beta_{i,top} > \beta_{i,sec} > \beta_{i,secbot} > \beta_{i,bot}$, and even fewer users are estimated to have position effect that monotonically increase with position, $\beta_{i,top} < \beta_{i,sec} < \beta_{i,secbot} < \beta_{i,bot}$. This suggests that H3b has a negligible role in explaining the documented position effects. It also adds to the evidence against Value Signals (H2) or Choice Fatigue with Descending Monotonic Direction (H3a) where we would expect all users' to exhibit decreasing position effects. Finally, instead, the results indicate that 75% of users exhibit *both* top and bottom position effects such that $\beta_{i,top} > \beta_{i,sec}$ and $\beta_{i,secbot} < \beta_{i,bot}$. More specifically, 72% of users display a specific non-monotonic pattern where $\beta_{i,top} > \beta_{i,sec} > \beta_{i,secbot} < \beta_{i,bot}$. These key results suggest that top and bottom position effects are largely an individual phenomenon as more more consistent with Choice Fatigue with Varied Direction (H3c) or Choice Fatigue with Non-Monotonic Direction (H3d).

4.3 Empirical Test IV

Finally, we consider Empirical Test IV to get further traction on the choice fatigue hypotheses. Empirical Test IV is based upon the order in which individual users make download decisions.

Empirical Test IV: Ordering of Multiple Downloads. Under H3b, some users always make their downloads in a monotonic descending order, while others use a monotonic ascending order. Whereas, each user will vary between a monotonic descending or ascending order under H3c, or use a non-monotonic order under H3d.

[18]For the full estimation results, see Section 1.5 of the Supplementary Appendix.

4.3.1 Download Timing

To implement Empirical Test IV, we make use of our unlimited selection setting and NEP's accurate data on download timing to recover the order in which each user made their downloads in instances where they downloaded more than one item from a given list. We first assess the download ordering patterns in the aggregate user population, before further considering the extent to which individual users show systematic download ordering patterns across different lists.

As an initial analysis, Table 6 summarizes the download ordering patterns for all instances where a user downloads k items from an individual list. First, let $k \geq 2$. Here, there are 6370 instances where a user downloads at least two items from a list. The results show that users download their top-most selected item first in 76% of the instances, and download their bottom-most selected item first in 18% of instances. This gives clear evidence that not all users select their items from the top down, and that some users start their selections from the bottom.

However, to study whether users download their items in a monotonic order, it is better to focus on instances where users download more than two items from a list. While this reduces the sample size, it avoids including artificial instances of monotonic behavior when a user downloads exactly two items. Hence, the right-hand side of Table 6 presents some results for instances where users download at least three items ($k \geq 3$) or four items ($k \geq 4$) per list. This shows that items are downloaded in a monotonic order 70-73% of the time: 67% of instances exhibit a monotonic descending order, while 3-6% exhibit a monotonic ascending order. While this is particularly in line with Choice Fatigue with Heterogeneous Direction (H3b) or Choice Fatigue with Varied Direction (H3c), it also implies that the remaining 27-30% of instances exhibit non-monotonic download behavior as more consistent with Choice Fatigue with Non-Monotonic Direction (H3d). Hence, this evidence is consistent with the presence of all three explanations, H3b-H3d.

To go deeper, we now repeat this style of analysis but consider the extent to which users show systematic download ordering patterns across different lists. To do this, we analyze the 992 users who download at least $k = 2$ items in more than one instance. On average, we observe such users' multiple download behavior across 4 different lists. Table 7 presents the results. Overall, we see that 52% of such users 'always' download their top-most selected item first, while 3% of such users 'always' download their bottom-most selected item first.

To examine the extent to which these 'multiple download users' show systematic *monotonic* ordering behaviors across different lists, we now further restrict attention to users who download at least three or four items per list (with $k \geq 3$ or $k \geq 4$) in more than one instance. Table 7 indicates the following key results.

First, only 2-3% of users 'always' download their items in a monotonic order with *varied* directions. This suggests that Choice Fatigue with Varied Direction (H3c) has a negligible role.

Second, only 0-2% of users always download their items in a strict *ascending* order. Thus, the evidence for Choice Fatigue with Heterogeneous Direction (H3b) is also weak as the proportion of users that employ an ascending order is insignificant from zero and seemingly too small to account for the size of the documented bottom position effects.

Instead, the download timing data provides evidence of two large groups of users that behave in line with H3a and H3d. Consistent with Choice Fatigue with Monotonic Descending Direction (H3a), 40-43% of users always download their items in a strict *descending* order. Further, consistent Choice Fatigue with Non-Monotonic Direction (H3d), 52-58% of users show no systematic monotonic behavior. Instead, such users typically download their selected items in a non-monotonic order 49-54% of the time, with 14-19% of users always downloading their items in a non-monotonic order.

4.4 Summary

We now bring our findings together to summarize the paper's overall evidence on the causes of top position effects. As approximately 30% of the observed top position effects are eliminated once the item order is randomized, H1 (Specific Item Order) does offer an important partial explanatory role. However, accounting for the remaining, more fundamental, causes is more complex and our results suggest that a number of explanations are at play for different groups of users. Nevertheless, we find that H3d (Choice Fatigue with Non-Monotonic Direction) best fits the evidence due its consistency with both i) the random parameter results where 75% of users exhibit *both* top and bottom position effects, and ii) the download timing results which suggest that a large group of users download their items in a non-monotonic sequence.

The other considered hypotheses have less explanatory power. Our download timing data indicates that very few users behave in line with Choice Fatigue with Heterogeneous Direction (H3b) or Choice Fatigue with Varied Direction (H3c), while H3b is also inconsistent with our result that top and bottom position effects are largely an individual-level phenomenon. H2 (Value Signals) and H3a (Choice Fatigue with Descending Monotonic Direction) lack support due to the presence of bottom position effects and the non-monotonic patterns within the random parameter results. However, the finding that a large group of users regularly download their items in a descending order suggests that H3a may play a role in accentuating the magnitude of top position effects.

5 Conclusion

Top position effects in individuals' choices from lists are an important phenomenon in many market, financial, and political settings. However, their cause has remained ill-understood. To provide an opportunity to cleanly measure and assess the sources of top position effects, this paper has used a natural field experiment capable of navigating several confounding factors.

Among other results, we have found that i) significant top position effects remain even when the order of items is randomized, and ii) top position effects co-exist with smaller, but highly significant, bottom position effects. Moreover, our novel user-level data has allowed us to show that iii) most users exhibit *both* top and bottom position effects, rather than just one or the other, and iv) distinct groups of users behave differently with regard to the order in which they select multiple items. We conclude that the causes of top position effects are complex and heterogeneous across individuals. However, our findings are most consistent with a version of choice fatigue where users consider the listed items in a non-monotonic order. In addition, there may be an additional secondary explanation resulting from a different group of users who also exhibit choice fatigue but consider the items in a strict descending order.

While we must be careful not to transfer these findings too far outside their context, our insights should prompt future research in several broad regards. First, it would be valuable to further understand what determines a user's inspection order and how this varies in different settings. Second, our insights may help analyze the potential for policy to nudge individuals into selecting beneficial options, such as more suitable savings and insurance plans or healthier foods (e.g. Dayan and Bar-Hillel 2011). Finally, our results may also assist in understanding a number of issues that are currently highly relevant in industrial economics and antitrust in relation to search order and prominence. For instance, as recently reviewed by Armstrong (2017), the causes of position effects are key when analysing the extent to which firms can manipulate consumers' choices through the presentation of their product ranges (e.g. Petrikaitė 2018), and a variety of issues regarding the effects, design, and regulation of list-based platforms (e.g. Athey and Ellison 2011, McDevitt 2014, de Cornière and Taylor 2014).

References

[1] Ansari A. and Mela C.F. (2003) "E-customization" Journal of Marketing Research vol.40 p.131-145

[2] Armstrong M. (2017) “Ordered Consumer Search” Journal of the European Economic Association vol.15 p.989-1024

[3] Asch S.E. (1946) “Forming Impressions of Personality” Journal of Abnormal and Social Psychology vol.41 p.258-290

[4] Athey S. and Ellison G. (2011) “Position Auctions with Consumer Search” Quarterly Journal of Economics vol.126 p.1213-1270

[5] Augenblick N. and Nicholson S. (2016) “Ballot Position, Choice Fatigue, and Voter Behavior” Review of Economics Studies vol.83 p.460-480

[6] Baye M.R., De los Santos B. and Wildenbeest M.R. (2016a) “Search Engine Optimization: What Drives Organic Traffic to Retail Sites?” Journal of Economics and Management Strategy vol.25 p.6-31

[7] Baye M.R., De los Santos B. and Wildenbeest M.R. (2016b) “What’s in a Name? Measuring Prominence, and Its Impact on Organic Traffic from Search Engines” Information Economics and Policy vol.34 p.44-57

[8] Biswas D., Grewal D. and Roggeveen A. (2010) “How the Order of Sampled Experiential Products Affects Choice” Journal of Marketing Research vol.47 p.508-519

[9] Caplin A., Dean M., and Martin, D. (2011) “Search and Satisficing” American Economic Review vol.101 p.2899-2922

[10] Coupe T., Ginsburgh V. and Noury A. (2010) “Are Leading Papers of Better Quality? Evidence from a Natural Field Experiment” Oxford Economic Papers vol.61 p.1-11

[11] Dayan E. and Bar-Hillel M. (2011) “Nudge to Nobesity II: Menu Positions Influence Food Orders” Judgement and Decision Making vol.6 p.333-342

[12] de Cornière A. and Taylor G. (2014) “Integration and Search Engine Bias ” RAND Journal of Economics vol.45 p.576-597

[13] De los Santos B., Hortaçsu A. and Wildenbeest M.R. (2012) "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior" American Economic Review vol.102 p.2955-2980

[14] De los Santos B. and Koulayev S. (2017) "Optimizing Click-Through in Online Rankings with Endogenous Search Refinement" Marketing Science vol.36 p.542-564

[15] Ellison G. and Wolitzky A. (2012) 'A Search Cost Model of Obfuscation' RAND Journal of Economics vol.43 p.417-441

[16] Fedyk A. (2019) "Front Page News: The Effect of News Positioning on Financial Markets", working paper

[17] Feenberg D., Ganguli I., Gaule P., and Gruber, J. (2017) "It's Good to be First: Order Bias in Reading and Citing NBER Working Papers" Review of Economics and Statistics vol.99 p.32-39

[18] Fishman A. and Lubensky D. (2018) "Search Prominence and Return Costs" International Journal of Industrial Organization, vol.58 p.136-161

[19] Greene W. (2010) "Testing Hypotheses about Interaction Effects in Nonlinear Models" Economics Letters vol.107 p.291-296

[20] Haan M.A., Dijkstra S.G. and Dijkstra P.T. (2005) "Expert Judgement Versus Public Opinion - Evidence from the Eurovision Song Contest" Journal of Cultural Economics vol.29 p.59-78

[21] Ho D.E. and Imai K. (2008) "Estimating Causal Effects of Ballot Order from a Randomized Natural Experiment: California Alphabet Lottery 1978-2002" Public Opinion Quarterly vol.72 p.216-240

[22] Itzkowitz J., Itzkowitz J., and Rothbort S. (2016) "ABCs of Trading: Behavioral Biases Affect Stock Turnover and Value " Review of Finance vol.20 p.663-692

[23] Jacobs H. and Hillert A. (2015) "Alphabetic Bias, Investor Recognition, and Trading Behavior" Review of Finance vol.20 p.693-723

[24] Janssen M.C. and Parakhonyak A. (2014) "Consumer Search Markets with Costly Revisits" Journal of Economic Theory vol.55 p.481-514

[25] Jeziorski P. and Segal I. (2015) "What Makes Them Click: Empirical Analysis of Consumer Demand for Search Advertising" American Economic Journal: Microeconomics vol.7 p.24-53

[26] Koppell K. G. S. and Steen J.A. (2004) "The Effects of Ballot Position on Election Outcomes" Journal of Politics vol.66 p.267-281

[27] McDevitt, R.C. (2014) "'A' Business by Any Other Name: Firm Name Choice as a Signal of Firm Quality" Journal of Political Economy vol.122 p.909-944

[28] Meredith M. and Salant Y. (2013) "On the Causes and Consequences of Ballot Order Effects" Political Behavior vol.35 p.175-197

[29] Miller J.M. and Krosnick J.A. (1998) "The Impact of Candidate Name Order on Election Outcomes" Public Opinion Quarterly vol.62 p.291-330

[30] Murphy J., Hofacker C. and Mizerski R. (2006) "Primacy and Recency Effects on Clicking Behavior" Journal of Computer-Mediated Communication vol.11 p.522-535

[31] Narayanan S. and Kalyanam K. (2015) "Position Effects in Search Advertising and their Moderators: A Regression Discontinuity Approach" Marketing Science vol.34 p.388-407

[32] Petrikaitė V. (2018) "Consumer Obfuscation by a Multiproduct Firm" RAND Journal of Economics vol.49 p.206-223

[33] Pinkowitz L. (2002) "Research Dissemination and Impact: Evidence from Website Downloads" Journal of Finance vol.57 p.485-499

[34] Salant Y. (2011) "Procedural Analysis of Choice Rules with Applications to Bounded Rationality" American Economic Review vol.101 p.724-48

[35] Schwarz N., Hippler H.-J. and Noelle-Neumann E. (1992) "A Cognitive Model of Response Order Effects in Survey Measurement" in "Context Effects in Social and Psychological Research" (Eds) Schwarz N. and Sudman S., New York: Springer

[36] Simon H. A. (1955) "A Behavioral Model of Rational Choice" Quarterly Journal of Economics vol.69 p.99-118

[37] Tan L. and Ward G. (2000) "A Recency-Based Account of Primacy Effects in Free Recall" Journal of Experimental Psychology: Learning, Memory and Cognition vol.26 p.1589-1625

[38] Train K.E. (2003) "Discrete Choice Methods with Simulation", New York: Cambridge University Press

[39] Ursu R. M. (2018) "The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions" Marketing Science vol. 37 p.530-552

Appendix: Tables and Figures

Table 1: Descriptive Statistics

	All	Control	Treatment
Number of alerts	530	350	180
Total number of items	6624	4269	2355
Average number of items per alert	12.50	12.64	12.20
Total number of downloads across items	35002	22856	12146
Average number of downloads per item	5.28	5.35	5.16
Total number of users that downloaded at least one item	9367	7024	4065
Average number of items downloaded per active user per alert	1.73	1.72	1.75
Average number of days between download and alert release	14.69	14.16	15.67
Total number of papers	4942	-	-
Average number of alert appearances per paper (within sample)	1.34	-	-
Average number of alert appearances per paper (within NEP)	3.90	-	-

Table 2: Alert- and Item-Specific Control Variables

Name	Description	Mean	St. Dev.	Min	Max
n	Number of items in alert (divided by 10)	1.25	0.83	0.40	1.18
ln(av)	Number of days alert was available (log)	6.64	0.06	6.54	6.73
engtitle	=1 if item has English title	0.99	0.10	0	1
title	Number of characters in item title (divided by 100)	0.75	0.28	0.00	2.43
title2	Title variable squared (divided by 10)	0.06	0.05	0.00	0.59
zeroab	=1 if item has no abstract	0.02	0.15	0	1
abstract	Number of characters in abstract (divided by 1000)	0.97	0.55	0	14.82
authors	Number of item authors	2.16	1.11	1	15
zerokey	=1 if item has no keywords	0.20	0.40	0	1
keywords	Number of item keywords (divided by 10)	0.37	0.28	0	3.20
keywords2	Keywords variable squared	0.21	0.38	0	10.24
zerojel	=1 if item has no JEL codes	0.42	0.49	0	1
jel	Number of item JEL codes	1.84	1.90	0	13
repstotal	Number of lists within NEP in which paper appears	3.90	1.39	2	12

Note: The descriptive statistics are calculated at the relevant alert- or paper-level.

Table 3: Aggregate Downloads by Position

	All	Control	Treatment
Number of Alerts	530	350	180
Average downloads per item across all positions	5.28	5.35	5.16
Average downloads per item in top position	8.05	8.42	7.33
Average downloads per item in second position	6.49	6.88	5.85
Average downloads per item in second-from-bottom position	4.86	4.78	5.02
Average downloads per item in bottom position	5.99	5.81	6.35

Figure 1: Aggregate Downloads by Position

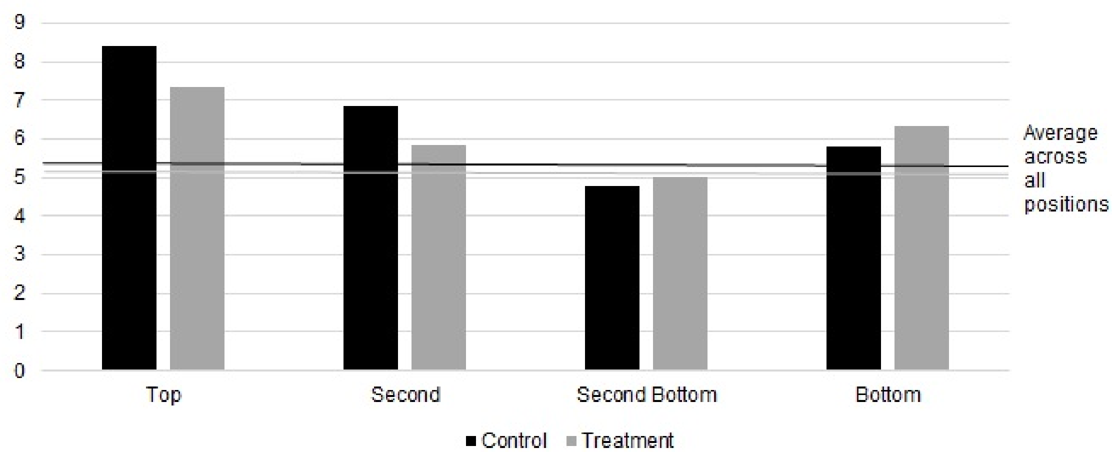


Table 4: Estimated Position Effects from Random Effects Estimations

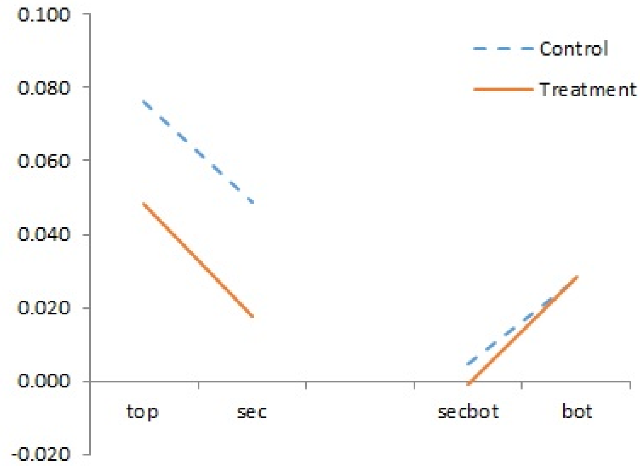
	Control		Treatment		All	
	i)	ii)	i)	ii)	i)	ii)
top	0.080	0.076	0.051	0.049	0.079	0.075
	(0.003)***	(0.003)***	(0.004)***	(0.004)***	(0.003)***	(0.003)***
top*treat	-	-	-	-	-0.023	-0.023
	-	-	-	-	(0.004)***	(0.004)***
sec	0.053	0.049	0.021	0.018	0.052	0.048
	(0.003)***	(0.003)***	(0.004)***	(0.004)***	(0.003)***	(0.003)***
sec*treat	-	-	-	-	-0.03	-0.03
	-	-	-	-	(0.005)***	(0.005)***
secbot	0.008	0.005	0.002	-0.001	0.007	0.003
	(0.003)**	(0.003)	(0.004)	(0.004)	(0.003)*	(0.003)
secbot*treat	-	-	-	-	0.001	0.000
	-	-	-	-	(0.005)	(0.005)
bot	0.031	0.028	0.032	0.029	0.030	0.026
	(0.003)***	(0.003)***	(0.004)***	(0.004)***	(0.003)***	(0.003)***
bot*treat	-	-	-	-	0.007	0.006
	-	-	-	-	(0.005)	(0.004)
treat	-	-	-	-	0.004	0.004
	-	-	-	-	(0.002)*	(0.002)*
Controls	Yes	Yes	Yes	Yes	Yes	Yes
User RE	-	0.340	-	0.324	-	0.358
	-	(0.023)***	-	(0.028)***	-	(0.024)***
Paper RE	-	0.093	-	0.000	-	0.000
	-	(0.021)***	-	(0.000)	-	(0.000)
Obs	189313	189313	99475	99475	288788	288788
Lists	350	350	180	180	530	530
LogLik	-67533	-65343	-36004	-35012	-103615	-100081
BIC	135297	135321	72227	72250	207531	207556
LR Tests:						
All 4 pos equal	462.7***	498.9***	99.9***	111.0***	-	-
Top=Sec	66.9***	71.8***	37.4***	42.7***	-	-
Sec=Secbot	153.3***	165.9***	13.1***	13.9***	-	-
Secbot=Bot	(-) 39.8***	(-) 43.4***	(-) 32.9***	(-) 35.8***	-	-

Notes for Table 4: Marginal effects are reported with the standard deviations of the random effects. (Robust) standard errors are in parentheses. Significance is denoted by * at 5%, ** at 1%, and *** at 0.1%. For comparison, we present two specifications involving

i) no random effects, and ii) both the user and paper random effects. The bottom of the table reports a series of likelihood ratio (LR) tests to assess i) the overall equality of the estimated position effects, $\beta_{top} = \beta_{sec} = \beta_{secbot} = \beta_{bot}$, and ii) the equality of 'adjacent' position effects; $\beta_{top} = \beta_{sec}$, $\beta_{sec} = \beta_{secbot}$, and $\beta_{secbot} = \beta_{bot}$.

Technical Notes for Table 4: Interpreting the marginal effects from interacted terms is difficult within discrete choice models (see Greene 2010, for example). However, i) our interest is more focused on their sign and significance, rather than their exact size, and ii) our resulting conclusions are robust if we only consider the estimated coefficients instead (details available on request). Similar arguments also apply to the results of the estimations in Table 6 and in the Supplementary Appendix. Although the underlying model is a simple binary probit, the estimation of equation (1) is clearly burdened by the presence of (potentially several) stochastic elements which need to be integrated out of the likelihood function. To this end, we have assumed that the elements follow independent normal distributions with zero means, and with variances that can be freely estimated. The estimations are then undertaken using simulated maximum likelihood techniques with 500 Halton draws (Train 2003). The results are robust to further increases in the number of draws.

Figure 2: Illustration of Estimated Position Effects from RE Estimations



Note: These position effects are derived from the estimated marginal effects in Table 4 for specification (ii).

Table 5: Summary of Users' Random Parameter (RP) Patterns

	Control	Treatment
Proportion of Users with Highest RP = top	0.81	0.64
Proportion of Users with Strictly Decreasing RPs	0.01	0.01
Proportion of Users with Highest RP = bot	0.03	0.16
Proportion of Users with Strictly Increasing RPs	0.00	0.00
Proportion of Users with Strictly Decreasing RPs (excluding bot)	0.70	0.72

Table 6: Summary of Download Ordering in Instances of Multiple Downloads

	$k \geq 2$			$k = 2$	$k \geq 3$	$k \geq 4$
Proportion of Instances	All	Control	Treatment	All	All	All
Top-Most Item Downloaded First	0.760	0.759	0.761	0.756	0.764	0.785
Bottom-Most Item Downloaded First	0.181	0.183	0.179	0.244	0.105	0.064
Downloaded in Mono' Order	0.880	0.881	0.879	1.000	0.734	0.698
Downloaded in Mono' Descending Order	0.718	0.717	0.720	0.756	0.672	0.670
Downloaded in Mono' Ascending Order	0.162	0.163	0.158	0.244	0.062	0.028
Items Downloaded in Non-Mono' Order	0.120	0.119	0.121	0.000	0.266	0.302
Number of Instances	6370	4096	2274	3494	2876	1562

Note: Specifically, these refer to instances where a user downloads k items from an individual list.

Table 7: Summary of Download Ordering for Multiple Download Users

Proportion of Such Users	$k \geq 2$	$k \geq 3$	$k \geq 4$
Always Download Top-Most Item First	0.52	0.53	0.56
Always Download Bottom-Most Item First	0.03	0.04	0.00
Always Download Items in a Mono' Order	0.67	0.48	0.42
Always Download Items in a Mono' Order with Same Direction	0.49	0.45	0.40
Always Download Items in a Mono' Descending Order	0.46	0.43	0.40
Always Download Items in a Mono' Ascending Order	0.03	0.02	0.00
Always Download Items in a Non-Mono' Order	0.03	0.14	0.19
Number of Such Users	992	441	233

Supplementary Appendix

As explained in the main paper, this supplementary appendix provides additional details on i) the subfields within the sample, ii) an example email alert, iii) the estimated effects of the control variables, iv) the estimated effects of list length, v) the robustness analysis using aggregate data, and vi) the random parameter estimations. All associated Tables and Figures are provided at the end of this appendix.

1.1 An Example Email Alert

nep-cbe
New Economics Papers on Cognitive and Behavioural Economics
Issue of 2017-06-18
six papers chosen by
Marco Novarese
Università degli Studi del Piemonte Orientale
http://econpapers.repec.org/pno2

——————————————————————————————————————

1. Nudging in education: A survey
Mette Trier Damgaard; Helena Skyt Nielsen

2. Digestible information: The impact of Multiple Traffic Light nutritional labeling in a developing country
Defago, Daniel; Geng, José F.; Molina, Oswaldo; Santa María, Diego

3. Facing Yourself: A Note on Self-image
Armin Falk

4. Essays on behavioral finance
Terzi, Ayse

5. Revealing the Economic Consequences of Group Cohesion
Simon Gaechter; Chris Starmer; Fabio Tufano

6. The Merit Primacy Effect

Alexander Cappelen; Karl Ove Moene; Siv-Elisabeth Skjelbred; Bertil Tungodden

——

1. Nudging in education: A survey

Date: 2017-06-08

By: Mette Trier Damgaard (Department of Economics and Business Economics, Aarhus University, Denmark) ; Helena Skyt Nielsen (Department of Economics and Business Economics, Aarhus University, Denmark)

Can we nudge children, youths and their parents to make better educational decisions? Educational decisions involve immediate costs and potential future benefits. Research suggests that in such settings behavioral barriers (such as lack of self-control, limited attention and social norms) are likely to influence choices. This raises the question whether low cost ”nudges” can improve people's educational choices. While interventions targeting cognitive or attentional limitations seem to be effective, it is too soon to provide a roadmap for introducing nudges in the education sector.

Keywords: Behavioural bias, boost policies, education choice, human capital investment
JEL: D03 D04 I20
URL: http://d.repec.org/n?u=RePEc:aah:aarhec:2017-05&r=cbe

The remaining items 2-6 are then presented in a similar format.

1.2 List of Subfields Within the Sample

The 29 subfields within the sample are: Africa, Ageing, Agricultural, Cognitive and Behavioural, Collective Decision Making, Computational Economics, Dynamic General Equilibrium, Education, Efficiency and Productivity, Time Series, Experimental, Forecasting, Happiness, Health, History and Philosophy, Human Capital, International Trade, Intellectual Property, Knowledge Management, Microfinance, Microeconomics, Migration, Marketing, Monetary, Post Keynesian, Project and Portfolio Management, Risk Management, Sports, and Transition.

1.3 The Effects of the Control Variables

Here we report the estimated effects of the control variables from the random effects estimations within Section 4.1.2. Table 10 below presents the estimated marginal effects of the control variables, corresponding to the estimations from Table 4. The results are best interpreted within the treatment group. Per-item downloads decrease in lists which have a larger number of items - see Section 1.4 below for more on the effects of list length. Items with an English title are more likely to be downloaded, and the probability of download is U-shaped in an item's length of title. The length of abstract provides no effect, but items with no abstract have a higher download probability. Items with a higher number of authors are less likely to be downloaded. Items with more keywords have a slightly higher download probability, and items without any JEL codes are less likely to be downloaded. Lastly, the probability of download is mildly increasing in the number of lists in which the item's underlying paper appears, perhaps reflecting the paper's general appeal.

1.4 The Effects of List Length

This section reports the full results of the estimated effects of list length referred to in Section 4.1.2. Specifically, we briefly study how our random effects estimatations vary with the number of items contained within an alert or 'list'. To proceed, we re-estimate equation (1) with an additional set of interaction terms, $\Pi' * n_l$, to measure how each position effect varies with list length, n_l.

The results are reported in Table 11. Within the control group, the four position effects are all significantly decreasing in list length. Intuitively, as the number of items increases, position effects become weaker because users' download activity is spread over more positions. However, within the treatment group, while we continue to observe a similar pattern for most positions, the estimated top position effects do not significantly decrease. Instead, with a weak level of significance, the top position effects actually increase and become relatively more pronounced. This pattern is even stronger in our analysis of the aggregate data, and related findings have also been documented by Ho and Imai (2008) and Feenberg *et al.* (2017).

While this is an interesting result with important implications, we are careful to not place too much emphasis on it for two reasons. First, variations in list length are unlikely to be fully exogenous. For instance, in our setting, list length varies due to differences in the supply of academic papers over time and across subfields, and may be correlated with variations in the quality of papers. Second, a pattern of increasing top position effects does not help distinguish between the explanations.

1.5 Robustness Analysis with Aggregate Data

This subsection shows how the dis-aggregated user-level results from Section 4.1.2 are robust under an alternative aggregate approach which considers how an item's list position affects the total number of downloads it receives.

Mirroring the user-level estimation in (1), the total downloads received by the item in position p of list l, d_{pl}, is modeled as a function of the position dummies, Π, the list-specific control variables, z_l, and the item-specific control variables, q_{pl}:

$$\beta_0 + \Pi'\beta_\Pi + z_l'\beta_z + q_{pl}'\beta_q \tag{3}$$

Any such estimation needs to take account of two features of the aggregate data. First, item downloads can only take the form of a non-negative integer, $d_{pl} \in \{0, 1, 2...\}$. Rather than using a negative binomial model, which is argued to be less robust, we address this issue by using a quasi-maximum likelihood estimator based on the Poisson distribution (Poisson QMLE).[19] Second, to account for the fact that some papers are included on the lists of more than one subfield, we cluster the standard errors by paper. This allows the error terms of observations with the same underlying paper to have a correlated error structure, while maintaining the assumption of independent errors for observations with different underlying papers. Similar to before, after estimating (3) on the control and treatment groups separately, we also estimate (3) on the full sample with the addition of $treat_l$ and the interacted position effects, $\Pi' * treat_l$.

Table 8 (below) presents the results with similar conclusions to the main user-level analysis. First, items in the top position within the treatment group still receive significantly more downloads than average despite the randomization of item order. In particular, items in the top position receive 36-52% more downloads than average. Second, smaller, yet significant, effects still exist for items in second position and bottom position, even after randomization. Third, bottom position effects are still significantly larger than the effects from the preceding, second-from-bottom, position.[20]

1.6 Full Random Parameter Results

This section provides the full random parameter estimation results from Section 4.2.1. Below

[19]See Wooldridge (1999) for more details on the Poisson QMLE and its relative advantages. Unless otherwise stated, all our main results can also be replicated using the negative binomial model.

[20]The Wald tests reported at the bottom of the table assess i) the overall equality of the estimated position effects, $\beta_{top} = \beta_{sec} = \beta_{secbot} = \beta_{bot}$, and ii) the equality of 'adjacent' position effects; $\beta_{top} = \beta_{sec}$, $\beta_{sec} = \beta_{secbot}$, and $\beta_{secbot} = \beta_{bot}$.

in Table 12 we present the results. For each estimation, we report the marginal effects of the main variables, together with the estimated standard deviations of the random parameters. All (robust) standard errors are presented in parentheses. While the overall results are consistent with the previous random effects estimations, the random parameter results document a substantial heterogeneity in position effects across users. This is illustrated further below in Figure 3 where the estimated random parameters are recovered following the method by Train (2009) and presented graphically.

Table 8: Estimated Position Effects with Aggregate Data

	Control		Treatment		All	
	i)	ii)	i)	ii)	i)	ii)
top	3.768	2.976	2.724	1.915	3.731	2.872
	(0.378)***	(0.337)***	(0.507)***	(0.444)***	(0.374)***	(0.332)***
top*treat	-	-	-	-	-0.589	-0.490
	-	-	-	-	(0.378)	(0.364)
sec	2.143	1.658	1.195	0.537	2.120	1.588
	(0.350)***	(0.319)***	(0.461)**	(0.382)	(0.346)***	(0.314)***
sec*treat	-	-	-	-	-0.666	-0.690
	-	-	-	-	(0.413)	(0.378)
secbot	-0.035	-0.418	0.319	-0.240	-0.035	-0.451
	(0.287)	(0.256)	(0.392)	(0.331)	(0.284)	(0.253)
secbot*treat	-	-	-	-	0.364	0.372
	-	-	-	-	(0.508)	(0.478)
bot	1.076	0.683	1.716	0.989	1.064	0.627
	(0.398)**	(0.351)	(0.427)***	(0.358)**	(0.393)**	(0.344)
bot*treat	-	-	-	-	0.586	0.505
	-	-	-	-	(0.526)	(0.479)
treat	-	-	-	-	-0.098	0.092
	-	-	-	-	(0.157)	(0.152)
Controls	No	Yes	No	Yes	No	Yes
Observations	4268	4268	2355	2355	6623	6623
Lists	350	350	180	180	530	530
Clusters	3317	3317	1895	1895	4942	4942
LogLik	-14800	-14100	-7910	-7525	-22700	-21700
BIC	29663	28430	15859	15197	45529	43679
$\hat{\sigma}^2$	4.69	4.07	4.29	3.68	4.55	3.96
Wald Tests:						
All 4 pos equal	79.9***	79.3***	16.5***	18.8***	-	-
Top=Sec	11.4***	9.31**	5.45*	6.30*	-	-
Sec=Secbot	27.1***	30.5	2.37	2.73	-	-
Secbot=Bot	(-) 5.71*	(-) 7.25**	(-) 6.31*	(-) 7.05*	-	-

Note: Marginal effects are reported with (robust) standard deviations in parentheses. Significance is denoted by * at 5%, ** at 1%, and *** at 0.1%. We present two specifications with and without the list-specific and item-specific controls. The marginal effects of the controls are omitted for brevity.

Table 9: Estimated Effect of List Length on Position Effects with Aggregate Data

	Control		Treatment	
	i)	ii)	i)	ii)
top	2.029	1.892	-0.028	0.133
	(0.482)***	(0.463)***	(0.647)	(0.616)
top*n	0.060	0.061	0.135	0.116
	(0.020)**	(0.020)**	(0.041)**	(0.036)**
sec	0.652	0.743	0.562	0.679
	(0.469)	(0.461)	(1.086)	(1.037)
sec*n	0.064	0.060	0.002	-0.013
	(0.029)*	(0.028)*	(0.063)	(0.059)
secbot	-0.470	-0.329	0.475	0.375
	(0.475)	(0.478)	(0.938)	(0.861)
secbot*n	0.001	-0.012	-0.055	-0.053
	(0.032)	(0.031)	(0.059)	(0.056)
bot	0.119	0.280	1.234	1.054
	(0.460)	(0.441)	(0.813)	(0.782)
bot*n	0.036	0.028	-0.012	-0.007
	(0.029)	(0.028)	(0.044)	(0.045)
n	-0.055	-0.050	-0.098	-0.072
	(0.004)***	(0.004)***	(0.014)***	(0.013)***
Controls	No	Yes	No	Yes
Obs	4268	4268	2355	2355
Lists	350	350	180	180
Clusters	3317	3317	1895	1895
LogLik	-14600	-14100	-7758	-7505
BIC	29256	28412	15594	15189
$\hat{\sigma}^2$	4.52	4.06	4.04	3.67

Table 10: Estimated Effects of Control Variables from Random Effects Estimations

	Control		Treatment		All	
	i)	ii)	i)	ii)	i)	ii)
n	-0.021	-0.019	-0.036	-0.034	-0.024	-0.022
	(0.001)***	(0.002)***	(0.001)***	(0.003)***	(0.001)***	(0.002)***
ln(av)	-0.015	0.005	-0.143	-0.117	0.003	-0.006
	(0.011)	(0.018)	(0.040)***	(0.052)*	(0.009)	(0.015)
engtitle	0.077	0.075	0.102	0.096	0.083	0.082
	(0.009)***	(0.011)***	(0.016)***	(0.017)***	(0.008)***	(0.009)***
title	-0.052	-0.066	-0.093	-0.094	-0.068	-0.077
	(0.009)***	(0.010)***	(0.014)***	(0.014)***	(0.008)***	(0.008)***
title2	0.098	0.170	0.327	0.318	0.192	0.231
	(0.057)	(0.059)**	(0.081)***	(0.080)***	(0.045)***	(0.046)***
zeroab	0.037	0.039	0.036	0.034	0.036	0.037
	(0.005)***	(0.005)***	(0.007)***	(0.007)***	(0.004)***	(0.004)***
abstract	-0.001	-0.002	0.000	0.000	-0.001	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
authors	-0.004	-0.004	-0.005	-0.005	-0.004	-0.004
	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
zerokey	0.017	0.018	0.009	0.012	0.014	0.016
	(0.003)***	(0.003)***	(0.004)*	(0.005)*	(0.003)***	(0.003)***
keywords	0.023	0.020	-0.015	-0.013	0.011	0.009
	(0.008)**	(0.008)*	(0.010)	(0.011)	(0.006)	(0.007)
keywords2	-0.006	-0.005	0.016	0.014	0.000	0.000
	(0.004)	(0.004)	(0.006)*	(0.006)*	(0.003)	(0.004)
zerojel	-0.001	-0.005	-0.006	-0.010	-0.002	-0.007
	(0.003)	(0.003)	(0.004)	(0.004)*	0.002	(0.002)**
jel	-0.002	-0.002	-0.001	-0.002	-0.001	-0.002
	(0.001)*	(0.001)**	(0.001)	(0.001)	(0.001)*	(0.001)**
repstotal	-0.001	-0.001	0.001	0.002	0.000	0.001
	(0.001)*	(0.001)	(0.001)	(0.001)**	(0.000)	(0.001)

Note: Marginal effects are reported for the main variables, with the standard deviations of the random effects. (Robust) standard errors are in parentheses. Significance is denoted by * at 5%, ** at 1%, and *** at 0.1%.

Table 11: Estimated Effect of List Length on Position Effects

	Control		Treatment	
	i)	ii)	i)	ii)
top	0.137	0.122	0.056	0.044
	(0.007)***	(0.008)***	(0.009)***	(0.008)***
top*n	-0.035	-0.026	0.001	0.007
	(0.005)***	(0.005)***	(0.006)	(0.005)
sec	0.094	0.080	0.063	0.049
	(0.006)***	(0.007)***	(0.010)***	(0.009)***
sec*n	-0.027	-0.019	-0.028	-0.021
	(0.004)***	(0.005)***	(0.007)***	(0.007)**
secbot	0.075	0.063	0.064	0.048
	(0.010)***	(0.010)***	(0.010)***	(0.010)***
secbot*n	-0.046	-0.039	-0.043	-0.033
	(0.008)***	(0.008)***	(0.008)***	(0.007)***
bot	0.095	0.082	0.084	0.069
	(0.007)***	(0.006)***	(0.010)***	(0.009)***
bot*n	-0.042	-0.035	-0.036	-0.027
	(0.005)***	(0.004)***	(0.007)***	(0.006)***
n	-0.018	-0.016	-0.032	-0.032
	(0.001)***	(0.002)***	(0.002)***	(0.003)***
Controls	Yes	Yes	Yes	Yes
User-level RE	No	Yes	No	Yes
Paper-level RE	No	Yes	No	Yes
Obs	189313	189313	99475	99475
Lists	350	350	180	180
LogLik	-68514.8	-66423.5	-36216	-35198
BIC	137309.3	137334	72698	72721

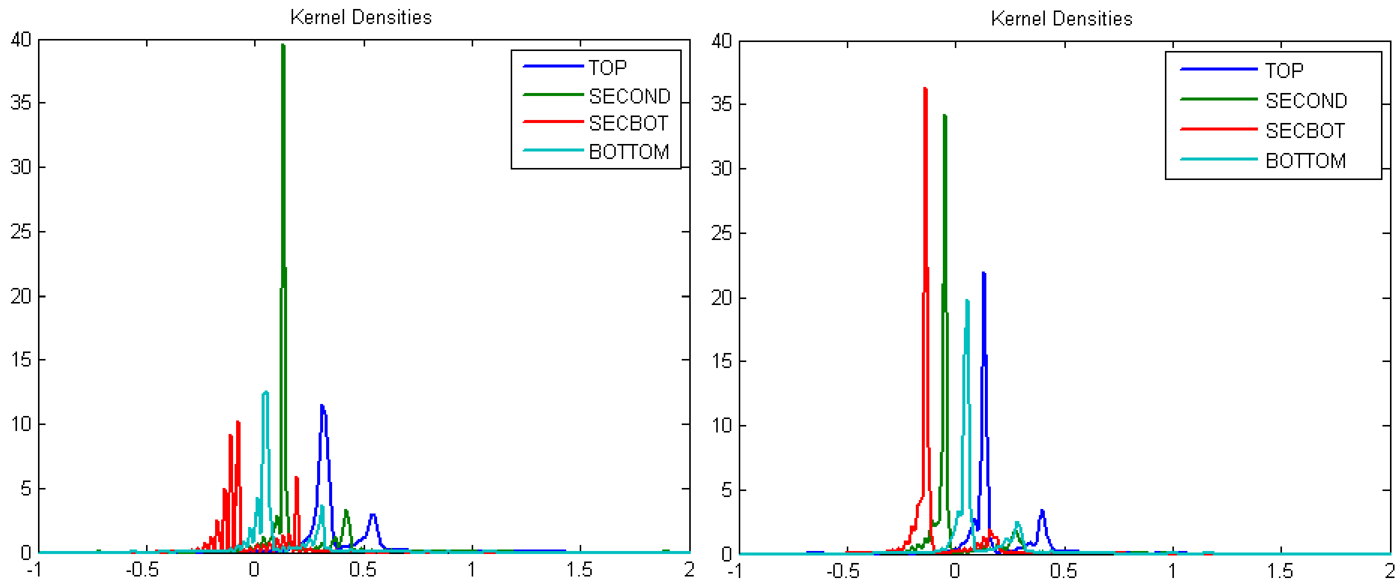
Note: Marginal effects are reported with (robust) standard deviations in parentheses. Test significance is denoted by * at 5%, ** at 1%, and *** at 0.1%.

Table 12: Estimated Position Effects from Random Parameter (RP) Estimations

	Control	Treatment
top	0.068	0.035
	(0.003)***	(0.006)***
sec	0.034	-0.001
	(0.004)***	(0.008)
secbot	-0.011	-0.019
	(0.005)*	(0.007)*
bot	0.017	0.017
	(0.004)***	(0.006)**
Full Controls	Y	Y
User-level RPs:		
top	0.377	0.410
	(0.038)***	(0.054)***
sec	0.420	0.451
	(0.038)***	(0.072)***
secbot	0.406	0.433
	(0.047)***	(0.062)***
bot	0.388	0.379
	(0.043)***	(0.057)***
Observations	189313	99475
Lists	350	180
LogLik	-67343	-35917
BIC	134954	72087

Note: Marginal effects are reported for the main variables, together with the estimated standard deviations of the (user-level) random parameters. All (robust) standard errors are presented in parentheses. Significance is denoted by * at 5%, ** at 1%, and *** at 0.1%. Like the random effects estimations, these augmented probit equations are estimated using simulated maximum likelihood techniques and Halton draws.

Figure 3: Estimated Random Parameter Distributions



Note: These distributions correspond to the random parameter estimations for each position effect in the previous table, Table 12. The graph on the left corresponds to the control group and the graph on the right corresponds to the treatment group.

References

[1] Wooldridge J.M. (1999) "Quasi-Likelihood Methods for Count Data" Ch.8 in 'Handbook of Applied Econometrics Volume II: Microeconomics', Pesaran M.H and Schmidt P. (Eds), Blackwell