Queen Mary
University of London

Centre for Globalization Research
School of Business and Management

Does Online Search Predict Sales? Evidence from Big Data for Car Markets in Germany and the UK

CGR Working Paper 71

Georg v Graevenitz, Christian Helmers, Valentine Millot and Oliver Turnbull

Abstract

We use online search data to predict car sales in the German and UK automobile industries. Search data subsume several distinct search motives, which are not separately observable. We develop a model linking search motives to observable search data and sales. The model shows that predictions of sales relying on observable search data as a proxy for prepurchase search will be biased. We show how to remove the biases and estimate the effect of prepurchase search on sales. To assist identification of this effect, we use the introduction of scrappage subsidies for cars in 2008/2009 as a quasi-natural experiment. We also show that online search data are (i) highly persistent over time, (ii) potentially subject to permanent shocks, and (iii) correlated across products, but to different extent. We address these challenges to estimation and inference by using recent econometric methods for large N, large T panels.

Keywords: Online search, Google Trends, Serial correlation, Non-stationarity, Common Correlated Effects, Large Panels

JEL codes: D83, L62, M31

CGR Working Paper Series

http://www.busman.qmul.ac.uk/cgr

Does Online Search Predict Sales?
Evidence from Big Data for Car Markets in Germany and the UK [*]

Georg von Graevenitz[a][†]   Christian Helmers[b]   Valentine Millot[c][‡]

Oliver Turnbull[d]

[a] *Queen Mary University, CCP and CREATe* [b] *Santa Clara University* [c] *OECD* [d] *Bristol University*

August 11, 2016

ABSTRACT

We use online search data to predict car sales in the German and UK automobile industries. Search data subsume several distinct search motives, which are not separately observable. We develop a model linking search motives to observable search data and sales. The model shows that predictions of sales relying on observable search data as a proxy for pre-purchase search will be biased. We show how to remove the biases and estimate the effect of pre-purchase search on sales. To assist identification of this effect, we use the introduction of scrappage subsidies for cars in 2008/2009 as a quasi-natural experiment. We also show that online search data are (i) highly persistent over time, (ii) potentially subject to permanent shocks, and (iii) correlated across products, but to different extent. We address these challenges to estimation and inference by using recent econometric methods for large N, large T panels.

KEYWORDS: Online search, Google Trends, Serial correlation, Non-stationarity, Common Correlated Effects, Large Panels

---

1   Introduction

The widespread use of online search engines, most prominently Google, is producing big and fast growing sets of micro-level data that can be used for empirical research. Research into online search behavior is currently also a fast growing field across a range of disciplines.

Empirical research on online search started with attempts to trace the onset of epidemics of influenza and similar diseases (Polgreen et al., 2008; Ginsberg et al., 2009).[1] Data that aggregate search intensity for specific topics, such as the flu, have proven to be an important innovation. For instance, data from Google Trends (GT) have been used to forecast consumption (Vosen and Schmidt, 2011; Choi and Varian, 2012), unemployment (Pavlicek and Kristoufek, 2014), real estate transactions (Dietzel et al., 2014), stock markets (Preis et al., 2013), and a number of political and governance related topics (Mellon, 2013, 2014; Mathews and Tucker, 2014). The field is growing rapidly: Nuti et al. (2014) report a seven-fold increase in publications using GT data in health care research between 2009 and 2013.

The breadth of the cited work shows that online search data are useful in many different contexts, due to the diversity of online activities and related search motives. This 'catch-all' characteristic of online search represents an opportunity for researchers in many fields. But it also creates a challenge: the observable aggregate data on search behavior result from a mix of distinct motives, such as search to inform a purchase, search for information post sale and search motivated by fandom. Identifying pre-purchase search intensity from aggregate search by using specific search terms is currently the only method available when predicting product sales or other outcomes from aggregate sales data. This approach can fail as research on prediction of flu outbreaks has shown.

This paper shows that a model of the three search motives outlined above can be used to derive an empirical specification identifying the effect of pre-purchase search on sales using aggregate search data. The model also predicts biases that arise if only product specific search terms are used to collect data. These biases are avoided by adding lagged sales and specific non-purchase search measures identified by the model to the empirical specification. The modeling approach we propose does not require collection of product characteristics, such as prices. With one exception, Hu et al. (2014), the importance of structural assumptions for identification has gone unnoticed in the literature on use of aggregate search data.

We apply our model to data on online search for products and product-level sales in the German and UK car markets. The objective is to demonstrate how aggregate online search data can be used to predict product-level car sales. We collect data on product searches and complement this with data on brand search and product level sales. Additionally, we control for the exogenous introduction of scrappage subsidies for cars in the aftermath of the 2007-2008 Global Financial Crisis that affected search behavior. We use the fact that the scrappage subsidies affected both sales and search. That is, controlling for the direct effect on sales, the scrappage event shifted the 'intention to buy' component of search, helping us to identify this component

---

[1]The use of Google Trends data in this context has recently been criticized by several authors pointing to poor documentation of the underlying flu trends data provided by Google (Ormerod et al., 2014; Santillana et al., 2014; Lazer et al., 2014).

in the aggregate search data.

Apart from identification, search data also pose specific challenges to estimation and inference. These problems arise from specific data properties whereby searches are (i) highly persistent over time (serial correlation), (ii) potentially subject to permanent shocks (non-stationarity) and (iii) potentially correlated across searches but to different extent (cross-section dependence). In addition, search data typically offer many potential measures of the same or similar underlying construct. These data properties are commonly referred to as tall (large T) and fat (large N), characteristics well-known in the analysis of 'Big Data' (Varian, 2014). Online search data pose both challenges: daily, weekly or monthly data on searches are available for many different combinations of terms used to search for the same product. Each combination of terms can lead to a slightly different aggregate series. In addition, the data for multiple products can easily be aggregated to form panels, creating large T, large N panels.

Explicitly modeling the different search motives that give rise to aggregate search data allows us to address the fatness of the data. The resulting model of search behavior restricts attention to specific search terms to identify each search motive. This approach, therefore, dispenses with the need to use complex and time-varying queries that attempt to identify specific components of the aggregate search data.

Tallness creates challenges for estimation and inference: online search data are non-stationary, serially correlated, and subject to cross-sectional dependence. If these challenges are ignored models using such data will be biased and inconsistent. We find these biases to be important in our applications. All of the challenges noted can be addressed with estimators developed for large N, large T panel data (Pesaran, 2004, 2006; Banerjee et al., 2010; Pesaran and Tosetti, 2011; Eberhardt et al., 2013; Hsiao, 2014; Chudik and Pesaran, 2015). To address the heterogeneous relationship between search and sales across products, we apply Common Correlated Effects Mean Group (CCE MG) estimators to our empirical model and show that the estimation of the model using OLS or fixed effects estimators leads to misleading interpretations of the underlying relationships in the data.

We use data from two separate car markets, the UK and Germany, to show that the problems we identify are not the result of the idiosyncrasies of a particular market. There are a number of differences in the data for the two markets which make for an interesting comparison. Data from the UK is available at weekly frequency, yielding taller data than the German data, which are available at monthly frequency. Also, the UK car market has a very strong bi-annual cycle, which is not evident in the German data. Nonetheless, the methodological problems we identify apply to both markets. We also find that the added tallness of the UK data makes empirical modeling more challenging. This confirms that less aggregated data contain richer dynamics (Rossana and Seater, 1995).

Our results confirm that omitted variable bias affects regressions of search on sales if search motives are not properly modeled. These effects are particularly significant where the scrappage subsidy is concerned. We also find that a failure to control for common correlated effects that we cannot directly observe, such as changes to Google's algorithms, changes to firms' uses of online marketing or changes to demand for specific

types of vehicles can significantly bias results. Overall we find that product level search is a highly significant and important predictor of product level sales.

Data capturing consumers' internet searches have been related to demand in a number of recent papers. Hu et al. (2014) study the interaction between search and advertising using aggregate search engine data to analyze a long panel ($T > 100$) of product level sales and advertising data. They also model search, but do not consider the empirical challenges presented by the tall nature of their data. De los Santos et al. (2012) throw light on *how* consumers search using a panel of individual level data. More generally, a better understanding of consumer exposure to advertising and consumer search behavior is emerging as an important factor in studies of product demand (Sovinsky Goeree, 2008; Moraga-González et al., 2015; Helmers et al., 2015). Our analysis contributes to this literature by showing how different motives buried in aggregate search data can be isolated and used to predict economic outcomes such as product-level sales. We also contribute to the literature by pointing to three data challenges that are likely to be present in any high-frequency online data, such as Google search, Twitter or Facebook data, but have gone largely unnoticed in the literature. We test for the presence of these problems and use recently developed panel estimators for large N and T to address them.

The next section of this paper provides a brief dicussion of online search behavior and the data it generates. Our model of search motives is discussed in Section 3, where we derive an empirical model. The data used are presented in Section 4. Section 5 provides results. Section 6 contains our conclusions. The appendices provide details on data collection and construction.

2   Search Data

The use of the internet has become widespread; while only between 30% and 40% of the population in the U.S. used the internet in 2000, the share reached around 87% in 2014 (Purcell et al., 2012; Lebo, 2014). Use of search engines and email are the most common reasons for people to access the internet by a wide margin. For example, in the U.S. 54% of users currently report using search engines at least once a day.[2] The vast majority (83%) of all U.S. internet users reported in 2012 using Google as the main search engine. Google had a market share of search of around 95% in Germany[3] and 88% in the UK[4] in 2015.

The Google search engine is used to search for information ranging from entertainment to work. A significant share of searches is for information on goods and services that people are interested in buying on- and offline. In Germany 90% of internet users search for information on goods and services on the internet and 75% have bought such goods and services over the internet within the last year (Statistisches Bundesamt, 2015). In the UK 85% of internet users compare goods and services online, 87% report buying online (Dutton et al., 2013).

---

[2]For the UK there is some evidence that there is a growing group of users who report mainly going to specific websites, but this group currently stands at 21% of all internet users, indicating that the overwhelming majority of users rely on search engines very regularly.

[3]Compare http://www.seo-united.de/suchmaschinen.html, accessed on the 6.8.2015.

[4]Compare http://theeword.co.uk/info/search_engine_market/, accessed on the 6.8.2015.

Search engines such as Google aggregate information on individual searches and make the data available to analysts and researchers. Google makes aggregate search data publicly available through GT. It provides information on the percentage of Google web searches for a given word or a combination of several words compared to the total number of Google searches taking place at the same time. Data series date back to the 1st of January 2004 and are available at daily, weekly, or monthly intervals.

The GT interface allows comparisons of search terms across different time periods and different locations. Multiple search terms can also be compared to each other and up to five different terms can be added and compared within the same query. Moreover, the use of operators makes it possible to look for combinations of different search terms (searches including each of the terms, at least one of the terms, or excluding one the terms). Additional flexibility is offered by filters for 30 different search categories and 250 subcategories (Section 3.3 provides more details).

For research purposes search data are overly aggregated, despite considerable flexibility in extracting search data from GT. Consider searches for the term Apple. The raw data provided by GT for this keyword conflate searches for the fruit with those for the firm's website. Similar problems arise in our empirical setting, the automotive industry; think for example of the Jaguar brand or the Volkswagen Beetle car model. This problem of how to interpret GT search indices extends beyond homonyms: keywords used to search for information related to a product or illness may change over time making a specific index less reliable, as was the case for the flu (Santillana et al., 2014).[5] In the case of product search we believe this problem to be generally less acute, as long as only the product name is used to specify the search index.[6]

Then there is also the difficulty of ascertaining the motives for searches observable in aggregate data. This is particularly important where the focus is on predicting an economic outcome such as product sales using the volume of search. For example, if a particular product is iconic and expensive, search volumes may be high, while sales measured in units, will be much lower than those of less iconic products. This problem is compounded when the available aggregate search index reflects a combination of search motives: searches motivated by fandom, searches motivated by interest in purchasing a product, searches motivated by news about a product or associated brand, searches due to post-sales problems with a product etc.

To see why this creates challenges to empirical research beyond potentially random measurement error, consider the example of product sales. A researcher might be interested in predicting product sales based on Google search volumes. The required assumption here is that consumers search on Google before purchasing a product. It is however possible that consumers search for a product *after* having purchased it and need additional information about it, for example on how to use it. If these two different search motives cannot be disentangled, there is potential for reverse causality.

The following section develops a model that formalizes these problems and helps link them to the empirical

---

[5] Brynjolfsson et al. (2014) suggest a crowd-sourcing solution for this problem where a random sample of internet users suggest keywords that they associate with a concept.

[6] Searches for combinations of product names and words such as 'dealer' or phrases such as 'buy online' reveal shifting patterns of sourcing for products over time. These were not our focus so we removed these additional terms from our searches. For more discussion see Appendix A.

models that deal with problems arising from the characteristic data properties of online search data, the length of the time series and cross-sectional correlation.

3   Methods

In this section we model three principal search motives and show how pre-sale search can be identified and used to predict product-level sales. We also discuss how product level-data necessary for this analysis can be obtained and show how our model can be estimated using weekly and monthly time-series data.

Throughout the section we focus on the relation between two variables: $y_{i,t}$ capturing sales of a product and $g_{i,t}$ capturing aggregate search intensity for these products on a search engine, for example GT.

3.1   Modeling Sales and Search

To begin with, consider a simple model of all searches for a product on a search engine. We disaggregate search for product (i) at a given time (t), $g_{i,t}$, into three components: search driven by need for information before a purchase $s_{i,t}^a$, after a purchase $s_{i,t}^p$ and search driven by interest $s_{i,t}^f$ in the product unrelated to a purchase:

$$g_{i,t} \equiv s_{i,t}^a + s_{i,t}^p + s_{i,t}^f \qquad . \tag{1}$$

Note that Equation (1) could easily also accommodate other search motives. Demand for a product is a function of past sales and demand resulting from information found during a pre-sale search:

$$y_{i,t} = \beta_0 + \sum_{l=1} \beta_{t-l} y_{i,t-l} + \beta_a s_{i,t}^a + u_{i,t} \qquad . \tag{2}$$

Past sales will strongly influence current sales for products that are bought regularly by a proportion of consumers. These buyers have no need to search for information about the products on offer in the market. Less frequent customers are more likely to need information ahead of a purchase and will therefore search the internet for such information.

Since the intensity of each type of product related search, $s_{i,t}^a$, $s_{i,t}^p$, and $s_{i,t}^f$, are all unobservable, we assume that they can be modeled as a function of observable product characteristics as follows:

$$s_{i,t}^a = \alpha_0 + \alpha_p P_{i,t} + \alpha_v V_{i,t} + \alpha_b B_{i,t} + v_{i,t} \tag{3}$$

$$s_{i,t}^p = \pi_0 + \pi_1 \sum_{r=1}^{R} \delta_{t-r} y_{i,t-r} + w_{i,t} \tag{4}$$

$$s_{i,t}^f = \phi_0 + \phi_b B_{i,t} + e_{i,t} \tag{5}$$

where P_t is the price of the product, V_t is a vector of objectively measurable product characteristics and B_t is

the value attributable to the product on account of the attached brands. $v_{i,t}$, $w_{i,t}$, and $e_{i,t}$ are i.i.d. errors.

In this model, the volume of search for pre-purchase information is driven by the physical characteristics, the price and the product's brand recognition among consumers $B_{i,t}$. Even consumers who have little knowledge of these characteristics will find out about them in the course of their search and will refine their search accordingly. This is reflected in aggregate search data as measured by GT. Post-purchase search results from the installed base of products, which is a function of past prices, characteristics and brand recognition, viz. Equations (2) and (3). Search not motivated by purchase is assumed to be driven only by brand recognition among consumers $B_{i,t}$.

Choi and Varian (2012) suggest using specific keywords to isolate $s_{i,t}^a$. Hu et al. (2014), who focus on pre-purchase searches, attempt to further exclude post-purchase and fandom searches by excluding searches that involve specific keywords indicating an interest in repairs, spare parts and the like. Reliance on category filters and attempts to exclude searches that do not precede a purchase can be problematic if Google alters the underlying algorithms or if consumer behavior changes so that new keywords need to be added to the exclusion list. Over time the set of keywords needed to exclude the wrong types of search is likely to change. This problem is at the root of the most recent criticisms of the attempts to predict outbreaks of epidemics using search engine data (Lazer et al., 2014).

We avoid this problem by building on the above model to analyze the link between pre-purchase search and sales. We use only the most general GT indices and discuss what can be identified using these data. Although we cannot separately identify from the available aggregate data search that arises before a purchase from the other types of search, we can still model the impact of search on sales by combining the above equations:

$$y_{i,t} = \beta_0 + \sum_{l=1} \beta_{t-l} y_{i,t-l} + \beta_a \left(g_{i,t} - \pi_0 - \pi_1 \sum_{r=1}^{R} \delta_{t-r} y_{i,t-r} - w_{i,t} - \phi_0 - \phi_b B_{i,t} - e_{i,t} \right) + u_{i,t}.$$

$$y_{i,t} = \underbrace{[\beta_0 - \beta_a(\pi_0 + \phi_0)]}_{\gamma_0} + \sum_{l=1}^{L} \underbrace{[\beta_{t-l} - \beta_a \pi_1 \delta_{t-l}]}_{\gamma_{t-l}} y_{i,t-l} + \beta_a g_{i,t} + \underbrace{u_{i,t} - \beta_a(w_{i,t} + \phi_b B_{i,t} + e_{i,t})}_{\nu_{i,t}}.$$

$$y_{i,t} = \gamma_0 + \sum_{l=1}^{L} \gamma_{t-l} y_{i,t-l} + \beta_a g_{i,t} + \nu_{i,t} \tag{6}$$

This model shows that the coefficients on lagged dependent variables may be negative, if the positive effects of persistence are weaker than those of post acquisition searches.

More importantly, by its definition (1) product search $g_{i,t}$ is correlated with consumers' brand recognition $B_{i,t}$. Absent a measure for the latter, the coefficient on the search index β_a is likely downward biased. To avoid this, we estimate our specification including a measure of brand recognition $B_{i,t}$:

$$y_{i,t} = \gamma_0 + \sum_{l=1}^{L} \gamma_{t-l} y_{i,t-l} + \beta_a g_{i,t} + \beta_b B_{i,t} + \epsilon_{i,t} \tag{7}$$

7

Addressing Endogeneity

We deal with the potential for bias of the coefficient on product search in two ways: first, we obtain another search index from GT which reflects searches for the manufacturer brand under which a given product is sold. For most car models this will be the brand name consumers are most likely to attach to the model and a brand level search index will yield a good proxy for $B_{i,t}$. Second, we exploit a quasi-natural experiment in the form of a short-term subsidy for purchase of new cars ($S_{i,T}$).[7] Although the subsidy will have a direct effect on sales, its main advantage is that it shifts pre-sale search without being correlated with fandom or post-purchase searches. Hence, the subsidy has two effects: First, it increases demand for cars that become significantly cheaper as a consequence of the subsidy. Second, it affects pre-sale search by affecting consumers purchasing decisions.

The direct effect leads us to rewrite the demand model (eqn. (2)) to include the scrappage search effect ($s_{i,t}^S$):

$$y_{i,t} = \beta_0 + \sum_{l=1} \beta_{t-l} y_{i,t-l} + \beta_a s_{i,t}^a + \beta_S s_{i,t}^S + u_{i,t} \quad . \tag{8}$$

The indirect effect means that we can rewrite pre-purchase search as:

$$s_{i,T}^a = \alpha_0 + \alpha_p \left(P_{i,T} - S_{i,T} \right) + \alpha_v V_{i,T} + \alpha_b B_{i,T} + v_{i,t} ,$$
$$\Leftrightarrow s_{i,T}^a = s_{i,T-1}^a - \alpha_p S_{i,T} \quad , \tag{9}$$

where T is the date at which the subsidy is announced. As the subsidy is temporary, $s_{i,T-1}^a$ will reflect the normal level of ex-ante search and we can then write the product search index after the subsidy has been announced as follows:

$$g_{i,T} = s_{i,T-1}^a + s_{i,T}^p + s_{i,T}^f - \underbrace{\alpha_p S_{i,T}}_{g_{i,T}^*} . \tag{10}$$

Note that $\alpha_p < 0$, if the law of demand applies to the products under consideration, so that $g_{i,T}$ is positively correlated with the level of the scrappage subsidy.

These considerations mean that we adjust our estimation model (7) as follows:

$$y_{i,t} = \gamma_0 + \sum_{l=1}^{L} \gamma_{t-l} y_{i,t-l} + \beta_a g_{i,t} + \beta_b B_{i,t} + \beta_S s_{i,t}^S + \epsilon_{i,t} \tag{11}$$

If we do not control for the subsidy event, the positive correlation between $s_{i,t}^S$ and $g_{i,t}^*$ implies that β_a will be upward biased. At the same time, the scrappage effect shifts $g_{i,t}$ by shifting pre-sale search $s_{i,t}^a$.

---

[7]Section 4 provides background detail on the scrappage subsidy.

3.2 Estimation

Given the product-level model outlined above we turn now to a discussion of how that model can be consistently estimated with large N and T panel data. GT data are available at daily, weekly or monthly frequency since 2004 and can be obtained for panels of search terms, e.g. names of car models or brand names. The high-frequency with which search data are available means that high levels of serial correlation and non-stationarity are likely to arise (Mellon, 2013). This applies to both the GT search data and to sales data. Hence, working with panels of GT data requires recent econometric methods that are currently not widely used (Pesaran, 2006; Banerjee et al., 2010; Pesaran and Tosetti, 2011; Hsiao, 2014; Chudik and Pesaran, 2015).

The estimation of individual time-series, which is a common approach with high frequency data in finance and macro-economics, discards valuable information available from the panel dimension of the data. Given the ease with which panels with relatively large N can be constructed from GT data, the estimation of individual time-series is also impracticable.

The solution we outline below is to use Mean Group estimators (Pesaran, 2006), to control for common correlated effects using cross-section averages of the dependent and explanatory variables, to adjust for seasonality and to estimate error correction models that identify short- and long-run effects of search on sales and are robust to any remaining non-stationarity.

Seasonality and Serial Correlation

Equation (11) contains L lags of the dependent variable to capture persistence and post-purchase search. Wooldridge (2012) notes that residual serial correlation in this kind of model is an indication that not enough lags of the dependent variable have been included in the specification.

Another source of serial correlation is seasonal variation in the demand for a product, which can be observed for almost any product. In our empirical application for example we see strong signs of such seasonal variation. Figures 1 and 2 show registration data (grey) and a search index (coloured) for the manufacturer name relative to 2004. The series for Toyota/BMW are shifted (down/up) by one unit in order to make the figures easier to read. The graphs suggest correlation between search and registration data over time; peaks of the search data generally predate peaks in the registration data.

Seasonal effects can be removed with methods like the X-13-ARIMA SEATS used by the U.S. Census Bureau or the Bank of England. Unfortunately, these methods affect the reliability of unit root tests applied to the transformed series (del Barrio Castro and Osborn, 2014).

We show below that removing seasonal effects by differencing each series using annual lags is effective. Results for data without seasonal adjustment are also reported. The transformed data are denoted $\tilde{v}_{i,t} \in \{\tilde{y}_{i,t}, \tilde{g}_{i,t}, \tilde{B}_{i,t}, \tilde{s}_{S,i,t}\}$ where $\tilde{v}_{i,t} = v_{i,t} - v_{i,t-12}$ in the case of monthly data and $\tilde{v}_{i,t} = v_{i,t} - v_{i,t-52}$ in the case of weekly data.

Figure 1:

Registrations and Searches for BMW,VW and Toyota
Germany, 2004-2014

Source of data: Kraftfahrtbundesamt; Google Trends; Own calculations
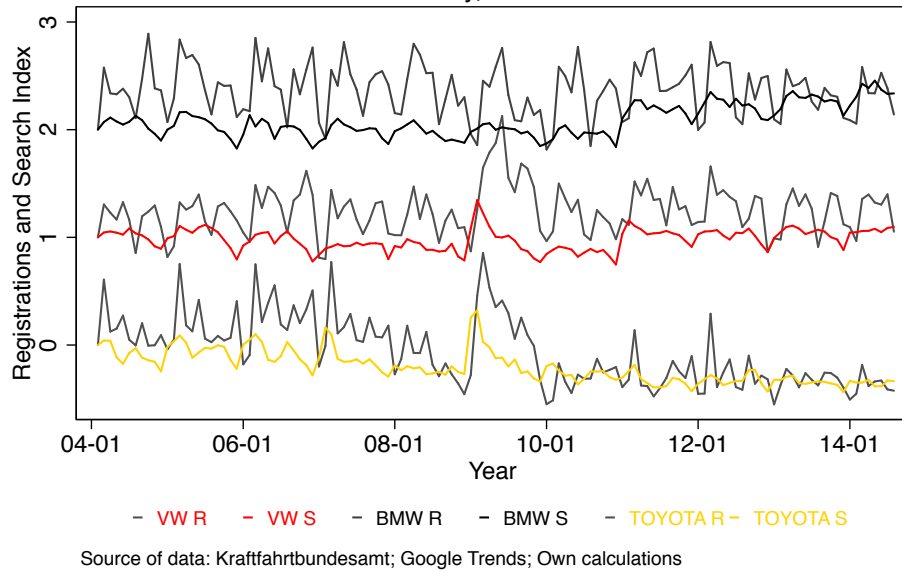
Figure 2:

Registrations and Searches for BMW,VW and Toyota
Great Britain, 2004-2011

Source of data: DoT - MOT project; Google Trends; Own calculations
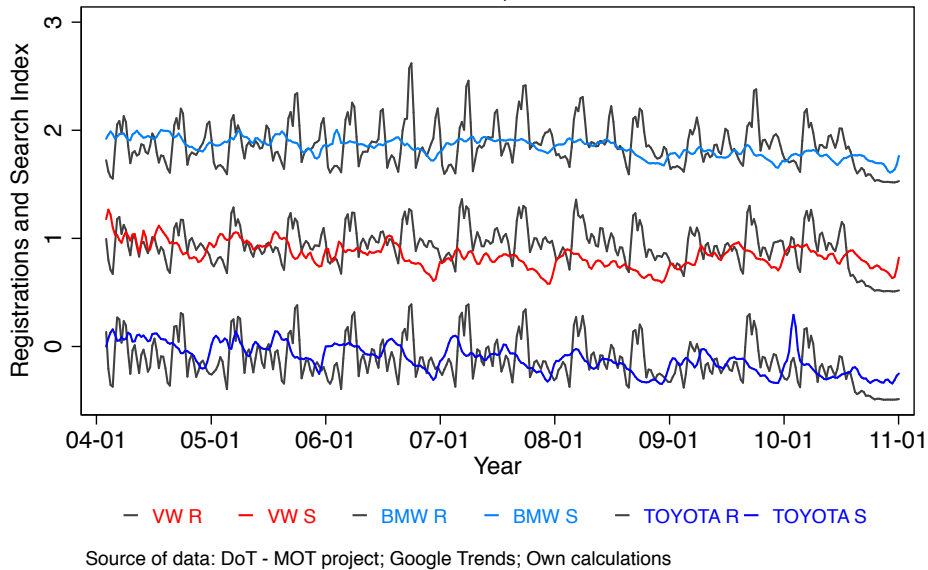
Co-integration

In section 4.3 we demonstrate that sales and the search index data are both affected by high levels of serial correlation. This makes it highly unlikely that the data series are stationary. Non-stationarity of the data series will lead to spurious correlation between variables and inference will be unreliable (Granger and Newbold, 1974).

To address the problem of non-stationarity we transform the auto-regressive distributed lag (ARDL) model set out in equation (11) by transforming the dependent variable into a first difference in case of monthly data

$(k = 1)$ or a fourth difference in case of weekly data $(k = 4)$ of the deseasonalized data:

$$\tilde{y}_{i,t} - \tilde{y}_{i,t-k} = \gamma_0 - (1 - \gamma_{t-k})\,\tilde{y}_{i,t-k} + \sum_{l=1,l\neq k}^{L} \gamma_{t-l}\tilde{y}_{i,t-l}$$
$$+ \sum_{m=0}^{M} \beta_{g,t-m}\tilde{g}_{i,t-m} + \sum_{n=1}^{N} \beta_{B,t-n}\tilde{B}_{i,t-n} + \sum_{o=0}^{O} \beta_S \tilde{s}_{i,t-o}^{S} + \epsilon_{i,t}\ . \quad (12)$$

This further differencing helps to stabilize the mean of the variables in the model. The model also includes lags of the product, brand and scrappage search indices to allow for lags in the impact of these search indices on sales.

A transformation of the dependent variable and the product search variables in this model results in an error correction specification, allowing us to estimate short- and long-run effects of product level search on sales:

$$\triangle\tilde{y}_{i,t} = \gamma_0 - (1 - \gamma_{t-k})\,\tilde{y}_{i,t-k} + \sum_{l=1,l\neq k}^{L} \gamma_{t-l}\tilde{y}_{i,t-l} + \beta_{g,t-k+1}\triangle\tilde{g}_{i,t-k+1} + \lambda_{g,t-k}\tilde{g}_{i,t-k}$$
$$+ \sum_{m=1,m\neq k}^{M} \beta_{g,t-m}\tilde{g}_{i,t-m} + \sum_{n=1}^{N} \beta_{B,t-n}\tilde{B}_{i,t-n} + \sum_{o=0}^{O} \beta_S \tilde{s}_{i,t-o}^{S} + \epsilon_{i,t}\ . \quad (13)$$

This specification corresponds to Equation 16 in Ericsson and MacKinnon (2002), who note that cointegration of $\tilde{y}_{i,t}$ and $\tilde{g}_{i,t}$ is implied by $-(1 - \gamma_{t-k}) < 0$. Gengenbach et al. (2015) provide critical values for the t-statistic on this coefficient for panel data, which has a non-standard distribution. In this representation, the speed of adjustment to long run equilibrium is captured by $(1 - \gamma_{t-k})$ and the long-run effect of search on sales is captured by $\frac{\beta_{g,t-k+1}}{(1-\gamma_{t-k})}$.

Note that we also test whether the brand-level and scrappage search indices are cointegrated with sales. We find no evidence to support this.

Unobserved Common Factors

The model of search developed here does not require that the researcher observe anything about the characteristics of the products being studied as long as the researcher's interest does not go beyond the relationship of search and sales. On the other hand, poor measures of consumers' brand recognition and of any shocks to product demand such as the scrappage subsidy will bias the coefficients of the model. Given the large number of products that may potentially be incorporated into an analysis, it is likely that subsets of these will be affected by unobserved common shocks leading to cross section dependence and biased coefficients.

To remove cross section dependence from unobserved common factors we apply the Common Correlated Effects Mean Group (CCE MG) estimator suggested by Pesaran (2006) and extended to the dynamic panel setting by Chudik and Pesaran (2015). The Mean Group estimator averages coefficients from the N time series making up a panel, allowing for heterogeneity in the intercept and slope parameters within the panel.

Pesaran (2006) shows that augmenting this estimator with cross-section averages of the dependent and the explanatory variables, removes correlations between unobserved common factors and the main variables of interest. Chudik and Pesaran (2015) then demonstrate that the CCE MG estimator can be extended to dynamic panels with weakly exogenous regressors. They observe that a suitable number of lags of the cross-section averages should be incorporated in the empirical model in this case. The number of lags required is the integer part of the cubic root of the length of the time-series, in our case this is 5 for the German data and 7 for the UK data. We apply this estimator to our data to account for the presence of unobserved common factors.

An important advantage of these estimators from the perspective of the applied researcher is that they are simple to use and allow the researcher to be agnostic about the structural interpretation of any unobservable common factors. [8]

The final representation of the empirical model we estimate is:

$$
\triangle \tilde{y}_{i,t} = \gamma_0 - (1 - \gamma_{t-k})\,\tilde{y}_{i,t-k} + \psi_{\triangle y,t}\triangle \bar{\tilde{y}}_t + \sum_{l=1,l\neq k}^{L} (\gamma_{t-l}\tilde{y}_{i,t-l} + \psi_{y,t-1-l}\bar{\tilde{y}}_{t-l-1}) + \lambda_{g,t-k}\tilde{g}_{i,t-k}
$$

$$
+ \beta_{g,t-k+1}\triangle\tilde{g}_{i,t-k+1} + \psi_{\triangle g,t-k}\triangle\bar{\tilde{g}}_{t-k+1} + \psi_{g,t-k}\bar{\tilde{g}}_{t-k} + \sum_{m=1,m\neq k}^{M}(\beta_{g,t-m}\tilde{g}_{i,t-m} + \psi_{g,t-m}\bar{\tilde{g}}_{t-m})
$$

$$
+ \sum_{n=1}^{N}\left(\beta_{B,t-n}\tilde{B}_{i,t-n} + \psi_{B,t-n}\bar{\tilde{B}}_{t-n}\right) + \sum_{o=0}^{O}\left(\beta_{S,t-o}\tilde{s}^S_{i,t-o} + \psi_{S,t-o}\bar{\tilde{s}}^S_{i,t-o}\right) + \epsilon_{i,t}\,. \quad (14)
$$

Here the ψ coefficients indicate a cross-section average term. This estimator may include a large number of cross-section average terms as well as lags of the dependent variable and of the three search indices. We have sought to restrict the number of lags used as much as possible in order not to loose too many observations, while still controlling for cross-section dependence. The tests used to identify the adopted lag structures are reported and discussed in Section 5.2 below. The testing process involved iterative removal of lags that did not contribute to good performance of the model, starting from the most extensive model. We tested a wide range of specifications and have included a number of tables with particular alternative specifications in the appendix to demonstrate robustness of the models presented in the main text.

Chudik et al. (2015) show that a cross-sectionally augmented distributed lag model can be used to directly estimate the long-run coefficient we obtain from the ARDL model. One advantage of their proposed estimator is its robustness to misspecified dynamics and serial correlation. We supplement our results below with those from their estimator and test whether the long-run parameters resulting from both approaches are significantly different in the panel. They suggest that the following model be estimated:

$$
\tilde{y}_{i,t} = \omega_0 + \theta_{g,t}\tilde{g}_{i,t} + \sum_{m=1}^{M}(\delta_{g,t-m}\tilde{g}_{i,t-m} + \phi_{g,t-m}\bar{\tilde{g}}_{t-m})
$$

---

[8]We implement the estimator using the package `xtmg` in Stata, but do not use the CCE option. Rather we augment the MG estimator with cross-section averages which we programmed and tested to control the precise lag structure of the estimator.

$$+ \sum_{n=1}^{N} \left(\delta_{B,t-n} \tilde{B}_{i,t-n} + \phi_{B,t-n} \bar{\tilde{B}}_{t-n} \right) + \sum_{o=0}^{O} \left(\delta_{S,t-o} \tilde{s}_{i,t-o}^{S} + \phi_{S,t-o} \bar{\tilde{s}}_{i,t-o}^{S} \right) + \epsilon_{i,t} . \quad (15)$$

Here $\theta_{g,t}$ is the coefficient capturing the long-run effect of the product level search index on sales and $M = 4$ for the German data, while $M = 6$ for the UK data as proposed by Chudik et al. (2015).

3.3 Google Trends

The search indices we use are obtained from GT. Here we briefly discuss a number of methodological issues that arise from this source of data.

First, the GT search indices are computed using a sampling method leading to variations of a few percent from day to day. Moreover, because of privacy considerations, only search queries with meaningful (Google does not specify the threshold) volume are shown to the public.[9] The data are normalised, i.e. the raw search counts are divided by an index of the total volume of search for a given area and time of search. Additionally, the data are scaled so that every data point for a given query is presented relative to the maximum value that the query will return multiplied by 100.

Second, GT provides data after 2004 or the date at which the search intensity for a particular term is sufficiently large. Depending on the search volumes, series are provided either at the monthly, weekly or daily level. Due to the normalisation, it is not possible to accurately aggregate data to a common frequency level, which limits the comparability of the different search terms (see also Appendix C).

Search queries are limited to a maximum of five terms in any one search, although each term may be quite complex: it can involve up to five separate search terms that are linked by either an *or* (disjunction) or an *and* (conjunction) operator.

Furthermore, GT rounds numbers to integer values. This restricts the information obtained from a search if the researcher downloads several series together and relative search volumes for one or more of these are so low that the series only ever consists of numbers below 10. By construction the detail regarding the variability of such a series is much lower than that for the series that vary between 10 and 100 in the same search.

In practice, many features of GT (e.g. in terms of output format, restrictions of queries etc.) tend to evolve rapidly, which poses problems for consistency over time. We have generally downloaded each product or brand level index together with one other index common to all queries. This allowed us to ensure that the data we obtain is normalized in a consistent way. For our purposes, the problem of volume thresholds is less relevant and we have simply excluded products for which we could not obtain enough data on search. In our context, these are cars for which sales volumes are also low.

---

[9]Stephens-Davidowitz (2013) proposes a method that provides data even on such low volume searches by jointly searching for the low volume term and a higher volume term.

4 Data

In this section we describe data on automobile sales from the UK and Germany. Sales are measured using registration data at monthly (D) or weekly frequency (UK). The availability of weekly data in the case of the UK demonstrates how the combination of high frequency search engine data with extremely detailed official data can provide information that is much more detailed than data on car sales that have previously been used. One question we also address here is whether the triple challenge of serial correlation, non-stationarity and common factors is exacerbated by this.

4.1 Variables

The datasets described here contain four variables: sales at the product level, measured on the basis of official registrations; search indices at the product level to measure product level search, search indices at the manufacturer level to capture consumers' valuation of brands, and search indices capturing search induced by the scrappage subsidy.

Sales The dependent variable is based on product registration data collected by national agencies in the UK and Germany. Data for the UK were obtained from the Department of Transport (DoT) on the basis of national data on MOT tests. Data for Germany were downloaded from the website of the KBA. Both sets of data are available from 2004 onwards. As the first MOT test takes place 3 years after the first sale of a car, the UK data end in 2011. The data reported by KBA end in 2014 (for details see Appendix D).

Search Indices We measure search intensity at the product level and at the the brand level. The indices we obtain are measured relative to a popular product or brand: we search simultaneously for the product/brand in question and for a reference product/brand. The reference product/brand anchors all search indices included in our data (for more information see Appendix B).

The product level search index encompasses pre-purchase, post-purchase, and non-purchase searches by product fans as set out in the model in Section 3. The brand level search index measures consumers' brand recognition.

Scrappage In 2009 the governments of Germany and the UK each introduced a temporary scrappage subsidy. These subsidies were introduced to encourage purchases of new cars and the scrappage of older ones. The subsidies were part of both governments' reactions to the significant fall in economic activity that followed the bankruptcy of Lehman Brothers in 2007 and the ensuing turmoil in the financial markets around the world. Each government had slightly different objectives as the German car market is dominated by manufacturers owned by German corporations, whereas this is not the case in the UK.

The subsidies were introduced in reaction to significant lobbying efforts by representatives of the auto industry in both countries, which took place against the backdrop of similar efforts by representatives of the

US auto industry to prevent the failures of GM and Chrysler in the context of the US presidential election campaign of 2008.

In Germany first reports in the national press that the government was considering a subsidy surfaced in October of 2008. The foreign minister and deputy Chancellor voiced his support for a scrappage subsidy as the way to support the auto industry in late December of 2008 and the government announced details of the subsidy in early January of 2009. The subsidy was christened 'Abwrackprämie' by the national press, a term not widely used otherwise.

The German government initially pledged € 1.5 billion to the subsidy program, promising to subsidize the purchase of a new car with € 2500, if a car older than 9 years was scrapped. As the uptake of the subsidy was significant, by March car sales were 40% higher than a year before and funds were likely to be exhausted, the government announced a top up of the program with a further € 3.5 billion.

The UK government followed suit with a similar scrappage scheme, however this was less generous as car manufacturers were required to match the subsidy offered by the government. Also, the amount pledged by the government was much lower, reflecting the scepticism of parts of the government for this form of subsidy.

The announcement of the subsidy had an immediate effect on subsidy related searches. The search index based on the names of these subsidies is extremely specific, as such searches are not observable either before or after the period in which subsidies were paid out. The index identifies the causal effect running from search to sales of a product better than either of the other indices we employ as it reveals the intentions of those searching more clearly than the product or manufacturer level searches. Figure 3 illustrates searches for the combination of the name of each manufacturer and the word 'Abwrackprämie' in Germany and likewise the combination of the manufacturer name and the word 'scrappage' in the UK.

Figure 3:
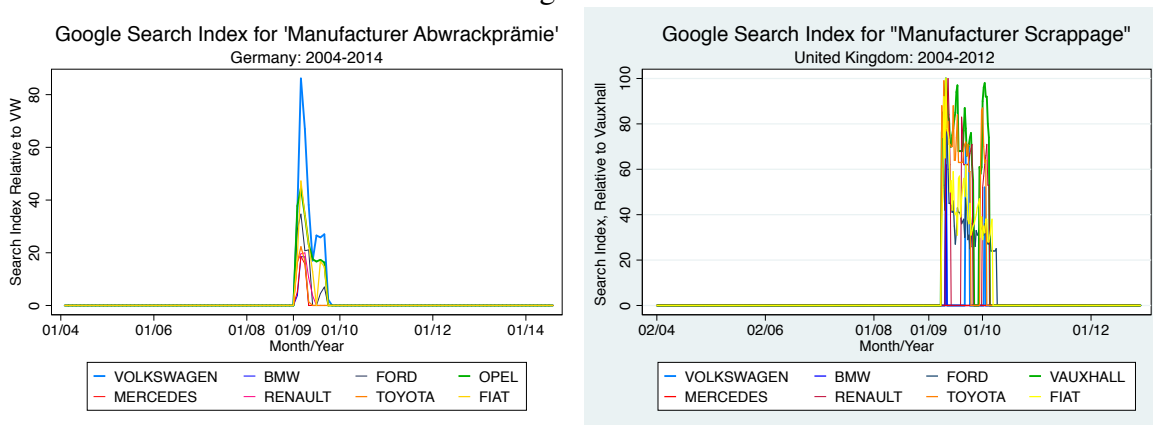
These two subsidy programs offer an interesting chance to compare how the announcement of temporary subsidies in the two countries affected search and subsequently car sales.

4.2 Descriptive Statistics

The data for the German automobile market are measured at monthly frequency for 127 months. KBA's registration data contain 486 product names. There are 50 brands of which Peugeot has the highest number of distinct model names (24) and Bentley and Lotus each have the lowest (1). We exclude all those months in which we do not observe any search information for a product from the sample. We did not get search data from GT for 182 products and have excluded a further 99 due to short length of the time series on sales. This leaves 205 product names in the data.[10]

Table 1 provides summary statistics for product registrations and the three search indices. We report statistics for the raw data and for the data which we have adjusted for seasonality by differencing against the same month one year prior. The descriptive data reveals a slight negative trend in the monthly sales of cars across all products and positive trends for the product and brand search indices. These are statistically significant at the 1% level; we report the t-statistics for these tests in the table.

The data for the UK automobile market are measured at weekly frequency for 463 weeks and 363 product names. There are 49 brands of which Mercedes has the highest number of distinct model names (31) and Bentley, Dodge and Lotus each have the lowest (1).[11] We exclude all those products for which we have fewer than 12 months of sales and three months of consecutive sales in our data. This leaves 226 product names in the data[12].

Table 1: *Summary Statistics - Germany, 205 Products, 127 Months*

Variable	Mean	Std. Dev.	Min.	Max.	N	t
Registrations	1429.067	2956.753	1	44280	16458	
Registrations (s.a.)	-77.932	1179.31	-22294	20059	16458	-8.48
Product SI	7.169	10.339	0	100	16458	
Product SI (s.a.)	0.575	5.879	-100	100	16458	12.55
Brand SI	36.161	24.711	1	95.75	16458	
Brand SI (s.a.)	0.325	5.159	-39.036	38.345	16458	8.07
Scrappage SI	0.688	4.763	0	86.129	16458	
Scrappage SI (s.a.)	-0.022	6.882	-86.129	86.129	16458	-0.41

Table 2 provides summary statistics for product registrations and the three search indices. We report statistics for the raw data and for the data which we have adjusted for seasonality by differencing against the same week one year prior. The descriptive data reveals a slight negative trend in the monthly sales of cars

---

[10]Descriptive statistics for all products are provided in Appendix E. A comparison with Table 1 in this section shows that the negative trend in seasonally adjusted registrations is more pronounced in the sample, and that the brand search index has a higher mean and lower standard deviation in the sample than in the population. Comparing product search indices with those for the population is not possible as we did not obtain product level search data for all products.

[11]We do not use the same set of product names for the UK and Germany as the registration data are reported with different names.

[12]Summary statistics for all products in the data are provided in Appendix E

across all products as well as a negative trend for the brand search index. There is a positive trend for the product search index. These trends are statistically significant at the 1% level; we report the t-statistics for these tests in the table.

Table 2: *Summary statistics - United Kingdom, 226 Products, 389 Weeks*

Variable	Mean	Std. Dev.	Min.	Max.	N	t
Registrations	121.125	345.824	0	9291	87914	
Registrations, (s.a.)	-24.171	203.352	-6834	6652	87914	-35.24
Product SI	13.132	15.257	0	100	87914	
Product SI, (s.a.)	0.952	6.783	-92	85	87914	41.59
Brand SI	38.037	20.901	0	100	87914	
Brand SI, (s.a.)	-1.023	4.906	-58	60	87914	-61.84
Scrappage SI	2.411	12.631	0	100	87914	
Scrappage SI, (s.a.)	0	18.177	-100	100	87914	0.00

4.3 Serial Correlation, Stationarity and Common Factors

Regression analysis based on non-stationary time series may reveal only spurious correlations. This section shows that data on car sales and product search are non-stationary in monthly data. Tables 3 and 4 show results from Pesaran's CIPS test (Pesaran, 2007) of non-stationarity for different degrees of serial correlation. The test is robust to the presence of common factors and serial correlation.

Table 3: *Unit Root Tests - Germany, Monthly Data*

	Sales		Product SI	
	not adj.	(s.a.)	not adj.	(s.a.)
Observations	12,827	11,615	12,827	11,615
	Pesaran 2007 CIPS Test: H_0 - All panels are non-stationary			
p-value, lags 2	0.000	0.000	0.000	0.000
Z-score, lags 2	-15.87	-17.09	-9.484	-10.15
p-value, lags 4	0.000	0.000	0.003	0.003
Z-score, lags 4	-9.502	-7.398	-2.712	-2.775
p-value, lags 6	0.004	0.000	0.421	0.446
Z-score, lags 6	-2.662	-3.658	-0.198	-0.136
p-value, lags 8	0.973	0.090	0.949	0.292
Z-score, lags 8	1.926	-1.340	1.632	-0.548

The table is based on 101 strongly balanced panels.

Both tables are restricted to those panels for which we have full coverage in our data, i.e. products that were sold over the entire sample period.

Table 4: *Unit Root Tests - United Kingdom, Weekly Data*

	Sales		Product SI	
	not adj.	(s.a.)	not adj.	(s.a.)
Observations	$87,914$	$87,914$	$87,914$	$87,914$
	Pesaran 2007 CIPS Test: H_0 - All panels are non-stationary			
p-value, lags 2	0	0	0	0
Z-score, lags 2	-67.05	-66.93	-51.11	-54.76
p-value, lags 4	0	0	0	0
Z-score, lags 4	-43.03	-46.77	-41.40	-45.98
p-value, lags 6	0	0	0	0
Z-score, lags 6	-33.16	-35.99	-31.31	-35.38
p-value, lags 8	0	0	0	0
Z-score, lags 8	-21.45	-22.77	-24.31	-27.46

The table is based on 226 strongly balanced panels.

The tests show that some panels are stationary if we restrict serial correlation to 6 lags. For 8 lags we cannot reject the null hypothesis of non-stationarity for registrations or the search index in the case of Germany. Sales and search in the UK data are stationary up to lag 8 in the weekly data we analyze. Appendix F shows that after aggregation to monthly data there is evidence of non-stationarity for registrations and the search index, as in the German data. This suggests that aggregation of time series can induce non-stationarity (Rossana and Seater, 1995).

Next we test the data for evidence of cross-section dependence. Table 5 provides results from Pesaran's test for cross section dependence in the sales and the product search index data. The test clearly demonstrates that this problem affects both types of data, regardless of seasonal adjustment. The table also reveals differences in the functioning of the market for cars in the two countries: while the seasonally adjusted absolute serial correlation of sales is quite similar, the level of product search is more correlated in Germany than in the UK.

Table 5: *Cross Section Dependence Test*

	Germany				United Kingdom			
Variable	CD-test	p-value	ρ	$|\rho|$	CD-test	p-value	ρ	$|\rho|$
Sales	196.27	0.000	0.198	0.292	1050.08	0.000	0.334	0.386
Sales (s.a.)	51.55	0.000	0.052	0.164	307.28	0.000	0.098	0.175
Product S.I.	363.44	0.000	0.363	0.421	576.85	0.000	0.183	0.329
Product S.I. (s.a.)	51.26	0.000	0.052	0.154	99.60	0.000	0.032	0.121

The table is based on 138 balanced panels for Germany and 226 balanced panels for the United Kingdom.

5 Results

Given the triple challenge of serial correlation, non-stationarity and unobserved common factors, we adopt the following approach to estimating reliable empirical models: seasonal effects are removed by differencing against observations made one year before, serial correlation is modeled by including lags of the dependent variable and effects of common factors are captured by including lagged cross-section averages of all variables. We rely on Mean Group estimators which include Error Correction Models (ECM) to identify both short- and long-run effects of search on sales. Alternative estimators are also presented and tested.

This section provides results from estimation of these MG ECM's as well as FE and OLS models. Results are presented in two steps: first, we compare results from OLS models and MG estimators based on the preferred model, derived using the MG estimator. These tables show how each element of our modeling strategy contributes to removing sources of bias. Next, we test the model specification described in Section 3. We show how inclusion of search indices related to the scrappage subsidy and to fandom alter the effect of product level search on sales.

5.1 Baseline Regressions

This section sets out regression results obtained with OLS, fixed effects (FE) and mean group (MG) estimators. Tables 6 and 7 contain five models: two columns with OLS results and three columns with results from FE or MG estimators. The OLS and MG estimators are presented without and with data transformed to remove seasonality.

The MG estimator results in column 5 of Tables 6 and 7 represent our preferred specification as serial correlation and cross-section dependence are minimized.[13] We present results from estimating models based on the specification of Equation (12), which does not include an error correction mechanism. This specification lends itself to comparisons with results from OLS and FE estimators. Tables 6 and 7 illustrate how misleading OLS and FE results for such long panels with cross section dependence can be: the product search index effects in the monthly data from Germany are biased towards zero, while they are too high by a factor of at least 4 in the weekly data from the UK.

The results presented in the 5 columns of Tables 6 and 7 are based on the same set of regressors, including cross-section averages of registrations, product, brand and scrappage level searches. As is indicated, the fixed effects (Date, Manufacturer) control for time-invariant differences across the models presented in these tables. The tests reported below the tables show that inclusion of cross-section averages does not remove common correlated effects from the OLS or FE estimators.[14] There is also evidence of serial correlation in the residuals of the OLS and FE estimators, particularly in the weekly UK data. Both problems contribute to the significant

---

[13]The same set of regressors is used in the preferred MG models with error correction mechanisms that are presented further below in Tables 8 and 10.

[14]In the specifications for the German data we use up to 13 lags (months). Specifications for the UK data are based on up to 28 lags (weeks), i.e. slightly over 6 months. Figure 2 shows the 6 months cycle that exists for cars in the UK as a result of the changes of letters on the number plates for new cars every 6 months.

biases of the main coefficients of interest, relative to the MG models in column 5 of these tables.

Table 6: *Germany, Monthly Data, N=* 16458, 205 *Products*

| | OLS | | Panel Estimators | | |
| | not adj. | (s.a.) | MG not adj. | FE (s.a.) | MG (s.a.) |
	(1)	(2)	(3)	(4)	(5)
Product SI $_t$	-1.143	2.380	25.04^{***}	2.489^*	23.99^{***}
	(1.287)	(1.739)	(6.184)	(1.227)	(5.237)
Product SI $_{t-1}$	3.935^*	5.973^*	14.84^{**}	6.271^{***}	15.73^{**}
	(1.878)	(2.500)	(5.258)	(1.225)	(5.792)
Brand SI $_{t-2}$	11.96^{***}	4.043^*	9.694^{**}	4.440^*	6.985^*
	(2.942)	(1.913)	(3.136)	(2.169)	(3.441)
Brand SI $_{t-3}$	-12.62^{***}	0.486	-6.491^{**}	0.0793	-2.347
	(3.490)	(2.765)	(2.375)	(2.006)	(2.222)
Scrappage SI $_t$	15.27^{***}	10.79^{***}	72.60	10.94^{***}	15.40^*
	(2.624)	(2.573)	(56.93)	(1.306)	(7.818)
Registrations$_{t-1}$	0.643^{***}	0.645^{***}	0.234^{***}	0.638^{***}	0.260^{***}
	(0.0295)	(0.0288)	(0.0156)	(0.00737)	(0.0183)
Registrations$_{t-9}$	-0.0511^{***}	-0.0377^*	0.0574^{***}	-0.0392^{***}	0.0347^{**}
	(0.0117)	(0.0158)	(0.0109)	(0.00717)	(0.0118)
Registrations$_{t-12}$	0.365^{***}	-0.429^{***}	0.0359^{**}	-0.430^{***}	-0.339^{***}
	(0.00945)	(0.00870)	(0.0118)	(0.00786)	(0.0133)
Registrations$_{t-13}$	-0.217^{***}	0.317^{***}	-0.0348^{**}	0.310^{***}	0.112^{***}
	(0.0113)	(0.0104)	(0.0116)	(0.00701)	(0.0122)
Manufacturer, Date FE	DATE	DATE	MAN.	BOTH	MAN.
Constant	-250.9	-32.09	-489.0^{***}	-12.84	-45.48^*
	(374.6)	(33.44)	(135.3)	(8.972)	(19.71)
CD Test p	0	0	0	0	0.218
CD Test	141.39	48.43	113.45	52.06	1.23
CH Test AR(1)	0.333	0.0389	0.0519	0.734	0.348
CH Test AR(2)	0.178	0.0263	0.835	0.243	0.711
CH Test AR(11)	0.787	0.926	0.930	0.0108	0.873
CH Test AR(12)	0.894	0.665	0.103	0.0121	0.160

Notes: Std. errors are clustered on the manufacturer. ***, **, * denote significance at the 0.1%, 1%, 5% level. Time period is 2004-2014. Sample contains 205 products with at least 24 months of positive sales. OLS models contain month, year and make fixed effects. All models also contain lags 1,3,4,6,7,8,11 of Registrations and Registration cross section averages at lags 1,3,6,11,12,13.

Diagnostics: CH (x) Test: Cumby and Huizinga (1992) test for H_0 of no residual serial correlation at lag x (p-values)[15]. CD Test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals, (p-values) and statistic. Test restricted to 101 highly balanced product panels. CIPS Test: Pesaran (2007) test for H_0 of non-stationary residuals with trend and maximum lags, (p-values) and statistic.

---

[15]This test is implemented as `actest` in Stata. In our setting it is equivalent to the Arellano and Bond (1991) test for autocorrelation.

Table 7: *UK, Weekly Data, N= 86558, 226 Products*

| | OLS | | Panel Estimators | | |
| | not s.a. | s.a. | MG not s.a. | FE (s.a.) | MG (s.a.) |
	(1)	(2)	(3)	(4)	(5)
Product SI $_t$	1.886***	0.595***	0.446**	0.614***	0.317*
	(0.131)	(0.0973)	(0.147)	(0.0977)	(0.131)
Product SI $_{t-4}$	1.340***	0.532***	0.411***	0.548***	0.304*
	(0.145)	(0.103)	(0.123)	(0.103)	(0.122)
Brand SI $_{t-25}$	−0.417	−0.707***	−0.00592	−0.736***	−0.262
	(0.235)	(0.176)	(0.230)	(0.180)	(0.178)
Scrappage SI $_{t-16}$	0.288***	0.274***	0.0528	0.275***	0.117*
	(0.0721)	(0.0528)	(0.0683)	(0.0531)	(0.0548)
Registrations$_{t-1}$	0.213***	0.134***	0.184***	0.131***	0.192***
	(0.00334)	(0.00336)	(0.00788)	(0.00337)	(0.00786)
Registrations$_{t-4}$	0.128***	0.141***	0.102***	0.139***	0.123***
	(0.00281)	(0.00338)	(0.00696)	(0.00338)	(0.00731)
Registrations$_{t-9}$	0.0479***	0.117***	0.0649***	0.116***	0.0630***
	(0.00287)	(0.00340)	(0.00635)	(0.00341)	(0.00617)
Registrations$_{t-26}$	0.512***	0.135***	0.198***	0.134***	0.0958***
	(0.00276)	(0.00335)	(0.0106)	(0.00335)	(0.00758)
Manufacturer, Date FE	DATE	DATE	MAN.	BOTH	MAN.
Constant	−30.82	−2.982	−0.148	−45.46	−2.364
	(93.38)	(18.87)	(13.71)	(530.5)	(3.642)
CD Test	557.86	783.40	6.82	492.47	2.81
CD Test p	0.00	0.00	0.00	0.00	0.005
CH Test AR(1)	0.016	0.559	0.120	0.652	0.169
CH Test AR(6)	0.000	0.003	0.951	0.001	0.877
CH Test AR(16)	0.925	0.167	0.022	0.297	0.016
CH Test AR(18)	0.000	0.000	0.691	0.000	0.360

Notes: ***, **, * denote significance at the 0.1%, 1%, 5% level. Time period 2004-2011. The sample contains 226 products with at least 3 months of consecutive sales and 52 weeks of positive sales. All models contain lags 1-4,7-9, 11,13-15,20-24,26,27 of registrations, lags 1-4,6-9,12-15,19,20,23,26,27 of registration cross-section averages (CA), the CA of the change in registrations. Furthermore models contain lags 8,13,17,23 and 27 of the product SI and lags 4,5,9,11,12,17-19,21-17 of product SI CA, lags 20,23,27 of brand SI as well as lags 2,7,8,9,10 of the brand SI CA, lags 3,8,12,20 of the scrappage SI and lags 1,3-6,10,16-19,21,23-27 of the scrappage SI CA.

Diagnostics: CH (x) Test: Cumby and Huizinga (1992) test for H_0 of no residual serial correlation at lag x (p-values). CD Test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals, (p-values) and statistic. Test restricted to 144 highly balanced product panels.

The result of Pesaran's 2004 CD test reported for column 5 of Table 6 (Germany) shows that any common correlated effects are captured by the cross section averages included in that specification. The same test reported for Table 7 (UK) shows that some evidence of common correlated effects remains for that model,

although this is much weaker than for the OLS and FE models we report for the UK.

Results differ strongly across the five empirical models for Germany and the UK. The pattern of deviations from the preferred model varies across the two countries. For instance the OLS models suggest that search does not affect sales in the German data, while the effect is overestimated by a factor of six in the case of the UK data when compared to the preferred model. In both countries the preferred MG model indicates that there is a positive and significant effect of contemporaneous search on sales.

5.2   Cointegration and Exogeneity

In this section we analyze our preferred models in detail, introducing the error correction specification and focusing on the biases discussed in Section 3.1. Tables 8 and 10 provide results from estimation of the ARDL model with error correction mechanism as set out in Equation (14). Tables 9 and 11 provide tests of these specifications based on the estimation of the long-run parameters from the specification set out in Equation (15) as suggested by Chudik et al. (2015).

Tables 8 and 10 show that the biases predicted by the model presented in Section 3 are present in the data for both countries. Moving from left to right in each table we add the search indices and cross-section averages for brands (column 2) and for brand level scrappage searches (column 3) to the baseline model that is based solely on predicting sales with product level search indices (column 1). As expected the coefficients on the short- and long-run effects of the product level search index are biased downwards, absent a measure of brand level search and are biased upwards, absent a measure of scrappage related search. The omission of scrappage related search indices results in much larger biases in both sets of data. The models in columns 1 and 2 are affected by significant cross-section dependence in the error terms as Pesaran's 2004 CD test demonstrates. The models in columns 3 and 4 exhibit no (Germany) or much less/no cross-section dependence (UK). Note that the specifications reported in columns 3 of Tables 8 and 10 correspond to those reported in columns 5 of Tables 6 and 7, the only difference being the use of the error correction model in the former tables.

In all columns of Tables 8 and 10 the t-statistics on the lagged dependent variable are below -40 for Germany and -110 for the UK. These statistics are below the 1% critical values reported by Gengenbach et al. (2015) for panel models with three explanatory variables, indicating that search and sales are cointegrated both in the UK and the German data.

In the German data (Table 8) the coefficients on brand level searches are only significant at the 10% level in column 3. However, the search index for the scrappage subsidy in column 3 is significant at the 5% level. The inclusion of brand and scrappage search cross-section averages in column 3 and of scrappage search cross-section averages in column 4 of Table 8 result in specifications for which there is no evidence of cross-section dependence in the errors. The specifications reported in columns 3 and 4 of the Table 8 are based on 7 cross-section averages (CAs) of the dependent variable, but only four each of the brand and scrappage search indices.

Table 8: *CCE MG Results for Germany, Monthly Data*

	(1)	(2)	(3)	(4)
Long Run Coefficient				
Product SI $_{t-1}$	43.08∗∗∗	46.51∗∗∗	34.54∗∗∗	35.16∗∗∗
	(6.448)	(7.328)	(7.614)	(6.300)
Short Run Coefficient				
Adjustment rate	−0.656∗∗∗	−0.703∗∗∗	−0.772∗∗∗	−0.721∗∗∗
(Registrations$_{t-1}$)	(0.0157)	(0.0164)	(0.0188)	(0.0169)
△ Product SI $_t$	20.61∗∗∗	25.08∗∗∗	22.84∗∗∗	21.40∗∗∗
	(4.357)	(4.902)	(5.309)	(4.272)
Brand SI $_{t-2}$		4.154	6.691†	
		(2.757)	(3.800)	
Scrappage SI $_t$			16.72∗	14.94†
			(7.849)	(7.639)
Registrations$_{t-9}$	0.0556∗∗∗	0.0517∗∗∗	0.0413∗∗∗	0.0404∗∗∗
	(0.00905)	(0.0102)	(0.0122)	(0.0105)
Registrations$_{t-12}$	−0.368∗∗∗	−0.333∗∗∗	−0.326∗∗∗	−0.348∗∗∗
	(0.0103)	(0.0120)	(0.0148)	(0.0127)
Registrations$_{t-13}$	0.131∗∗∗	0.105∗∗∗	0.0929∗∗∗	0.120∗∗∗
	(0.0108)	(0.0118)	(0.0132)	(0.0121)
Constant	−27.30	−21.81	−31.98	−22.16
	(17.06)	(19.65)	(25.21)	(21.96)
Products	200	200	200	200
Observations	16,284	16,284	16,284	16,284
CD Test	5.86	4.98	0.77	1.55
CD Test p	0.00	0.00	0.44	0.12
CH Test AR(1)	0.196	0.207	0.0301	0.0673
CH Test AR(2)	0.269	0.103	0.359	0.565
CH Test AR(11)	0.353	0.497	0.632	0.995
CH Test AR(12)	0.00302	0.00315	0.0911	0.0488

Notes: ***, **, *, † denote significance at the 0.1%, 1%, 5%, 10% level. Time period is 2004-2014. The sample contains 200 products with at least 24 months of positive sales. All models contain lags 1-6,8,9,12,13 of registrations, lags 0,1,3,6,11,12,13 of registration cross-section averages (CA) and the CA of the change in registrations. Models 2 and 3 contain brand SI CAs at lags 2,4,5 and 6. Models 3 and 4 contain lags 0,3,6-9 and 11 of the scrappage SI and lags 0,9,11,12 of the scrappage SI CAs.

Diagnostics: CH (x) Test: Cumby and Huizinga (1992) test for H_0 of no residual serial correlation at lag x (p-values). CD Test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals, (p-values) and statistic. Test restricted to 149 sufficiently balanced product panels.

An additional test of the specifications reported in Table 8 is provided by Table 9 in which we report tests of the mean long-run estimator ($\hat{\theta}_{g,t}$) in all four models presented in Table 8. $\hat{\theta}_{g,t}$ is estimated using the cross-sectionally augmented distributed lag model proposed by Chudik et al. (2015). This is compared to the

long-run estimator derived from the estimators reported in Table 8, based on all lags of the dependent variable included in these models. The smallest absolute average difference between the two methods of estimating the long-run parameter is obtained for model 4, which also has the lowest variance. We cannot reject that the coefficients differ on average for the specifications from Columns 2,3 or 4 in Table 8, but we can reject this for the base specification from Column 1.

Taking the evidence presented in Tables 8 and 9 together we view Model 3 as the most reliable representation of the German data.

Table 9: Comparing Long Run Effects from CS-DL and ARDL Models for Germany

Model	Products	$\hat{\theta}_{g,t}$	$\frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	$\hat{\theta}_{g,t} - \frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	Std. Err.	p	t
1	200	105.798	23.477	82.321	26.424	0.002	3.115
2	200	111.142	102.297	8.845	18.328	0.630	0.483
3	200	83.147	92.431	-9.284	36.473	0.799	-0.255
4	200	80.730	73.093	7.637	14.927	0.609	0.512

In the UK data (Table 10) the coefficients on brand level searches are not significant, while the search index for the scrappage subsidy in Columns 3 and 4 is significant at the 5% level. The inclusion of brand and scrappage search cross-section averages in Column 2 and of scrappage search cross-section averages in column 3 of Table 8 result in significant reductions of the level of cross-section dependence. In Column 4 we drop 29 products from the data for which the average product level search index is falling over the sample period. Many of these products are affected by significant one off reductions in search interest. Dropping these products does not significantly change any of the coefficients, but does reduce cross-section dependence very significantly, so that it no longer matters. Overall, we take this as an indication that cross section dependence doesn't create significant biases in either column 3 or column 4 of Table 10. The specifications reported in Table 10 are based on 17 cross-section averages (CAs) of the dependent variable, five brand level CAs and 16 scrappage search CAs.

We test the specifications reported in Table 10 by comparing the implied long-run effect of product level search with that obtained from an alternative model as suggested by Chudik et al. (2015) in Table 11. The smallest absolute average difference between the two methods of estimating the long-run parameter is obtained for model 4, but model 3 has the lowest variance of the estimates. The estimated long-run coefficients for product search are very close similar for both of these specifications, providing further evidence of their consistency.

In all models higher levels of product level search predict higher sales of the average product. In the short run a one standard deviation increase in the monthly product level search index (s.a.) predicts a 16.5% increase of car sales relative to average monthly sales in Germany. In the UK the corresponding effect is an increase of weekly sales by 4% of average weekly sales. In the UK data a permanently higher level of search has an effect almost twice as high as the short run effect on sales, whereas in Germany the long-run increase in sales is only 1.5 times higher than the short-run effect. Direct comparisons between both countries on the basis of

these results should be treated with caution due to the differing frequencies of the time series we rely on.[16]

Table 10: *CCE MG ECM Results for the UK, Weekly Data*

	(1)	(2)	(3)	(4)
Error Correction Term				
Product SI $_{t-4}$	0.673∗∗∗	0.681∗∗∗	0.621∗∗∗	0.632∗∗∗
	(0.179)	(0.185)	(0.176)	(0.191)
SR				
Adjustment rate	−0.862∗∗∗	−0.867∗∗∗	−0.877∗∗∗	−0.875∗∗∗
(Registrations$_{t-4}$)	(0.00702)	(0.00701)	(0.00731)	(0.00786)
$\tilde{\triangle}$ Product SI $_{t-4}$	0.342∗∗	0.346∗∗	0.317∗	0.292∗
	(0.125)	(0.132)	(0.131)	(0.138)
Brand SI $_{t-25}$		−0.311	−0.262	−0.272
		(0.188)	(0.178)	(0.168)
Scrappage SI $_{t-16}$			0.117∗	0.108∗
			(0.0548)	(0.0545)
Registrations$_{t-1}$	0.227∗∗∗	0.215∗∗∗	0.192∗∗∗	0.190∗∗∗
	(0.00791)	(0.00788)	(0.00786)	(0.00852)
Registrations$_{t-9}$	0.0672∗∗∗	0.0655∗∗∗	0.0630∗∗∗	0.0644∗∗∗
	(0.00606)	(0.00609)	(0.00617)	(0.00671)
Registrations$_{t-26}$	0.0925∗∗∗	0.0939∗∗∗	0.0958∗∗∗	0.0906∗∗∗
	(0.00729)	(0.00744)	(0.00758)	(0.00813)
Constant	−0.594	−2.155	−2.364	−1.813
	(1.795)	(2.021)	(3.642)	(3.708)
Products	226	226	226	197
Observations	86, 558	86, 558	86, 558	75, 451
CD Test	8.13	7.12	4.95	1.92
CD Test p	0.00	0.00	0.00	0.055
CH Test AR(1)	0.570	0.380	0.169	0.263
CH Test AR(6)	0.043	0.644	0.877	0.739
CH Test AR(16)	0.178	0.056	0.016	0.025
CH Test AR(18)	0.007	0.066	0.360	0.361

Notes: ***, **, * denote significance at the 0.1%, 1%, 5% level. Time period 2004-2011. The sample contains 226 products with at least 3 months of consecutive sales and 52 weeks of positive sales. All models contain lags 1-4,7,8,9,11,13-15,20-24,26,27 of registrations, lags 1-4,6-9,12-15,19,20,23,26,27 of registration cross-section averages (CA) and the CA of the change in registrations. Models 2,3 and 4 contain brand SI CAs at lags 2,7,8,9 and 10. Models 3 and 4 contain lags 3,8,12,16 and 20 of the scrappage SI and lags 1,3-6,10,16-19,21,23-27 of the scrappage SI CAs.

Diagnostics: CH (x) Test: Cumby and Huizinga (1992) test for H_0 of no residual serial correlation at lag x (p-values). CD Test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals, (p-values) and statistic. Test restricted to 144 highly balanced product panels.

---

[16]Rossana and Seater (1995) show that aggregation of a time-series to a lower frequency often removes evidence of low frequency cycles and can alter the properties of the time series significantly.

Table 11: Comparing Long Run Effects from CS-DL and ARDL Models for UK

Model	Products	$\hat{\theta}_{g,t}$	$\frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	$\hat{\theta}_{g,t} - \frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	Std. Err.	p	t
1	226	2.77	4.24	-1.46	1.17	0.21	-1.26
2	226	3.41	-90.85	94.26	94.74	0.32	0.99
3	226	2.77	2.28	0.49	0.70	0.48	0.70
4	197	2.58	2.29	0.29	0.80	0.72	0.36

Appendix G provides tables for both countries with specifications that contain additional cross section average terms such as the average change in the product, brand and scrappage search indices. In the German data these models can perform marginally better against the cross section dependence test, than our preferred model, but at the cost of losing observations. In the UK data two models perform marginally better against the cross-section dependence test than the model in column 3 of Table 10, but these models are affected by significant serial correlation. In both sets of data the long-run parameters obtained from these richer models differ more in absolute value from the long-run parameters we estimate with the alternative estimation method proposed by Chudik et al. (2015) than our preferred specifications. This illustrates the robustness of the results we present here.

Appendix H provides tables with specifications in which we use data that are not seasonally adjusted. The results show that significant evidence of cross-section dependence and serial correlation in the errors remains. This indicates that adjusting the data for seasonality is an important step in our analysis of the data.

Overall we find that the empirical models behave in a similar fashion for both countries, in spite of different levels of aggregation. In the following section we present product level results that demonstrate how heterogeneous the effect of search on sales is across products in both markets.

5.3 Product Level Results

The mean group estimator differs from a fixed effects estimator mainly because its coefficients are the average of the product level coefficients that are estimated independently in a first step. Both estimators remove fixed effects that are constant in time from the data. One advantage of the mean group estimator is that we can study not just the average effect of a variable in the sample, but also the effect for specific units of observation, here products. We exploit this possibility to compare the coefficients for the product-level search indices in the short-run and the long-run for the top ten most frequently sold car models in the two countries.

Tables 12 and 13 show the estimated coefficients corresponding to columns 1 and 3 of tables 8 and 10. We have argued that the coefficients from column 1 of these tables are biased. Tables 12 and 13 reveal that the biases we identified play out very differently across different products in both markets. Two observations stand out from the tables below: there is very significant variation in the short-run and long-run parameters in both markets. The largest values we observe are 5 or 6 times larger than the average coefficients reported in Table 8 and between 36 and 30 times larger than the average coefficients reported in Table 10. Secondly, there

are negative values. Some of the negative coefficients are not likely to be significantly different from zero, but others clearly are.

We focus on the case of the difference between the A-Class and the B-Class cars produced by Mercedes Benz and sold in significant volume in Germany as Table 12 reveals. Figure 4 shows both the product level search index and the level of sales for both car models in the German data relative to the mean for each series (we have shifted the two series for the A-Class by adding 2 to each observation to make the figure easier to read).

Table 12: *Product Level Search Index Coefficients - Germany*

			Coefficients			
			Column 1		Column 3	
Product	Manufacturer	Av. Sales	*s.r.*	*l.r.*	*s.r.*	*l.r.*
GOLF	VW	20855.92	12.54	81.09	77.04	75.02
C CLASS	MERCEDES	18911.55	95.64	67.37	37.43	126.33
A CLASS	MERCEDES	14215.11	−201.65	−64.45	−236.38	−105.09
B CLASS	MERCEDES	11301.34	230.61	509.71	154.16	173.89
PASSAT	VW	7623.24	−10.79	−15.52	−24.52	−8.02
ASTRA	OPEL (GM)	7151.69	43.44	111.17	49.71	90.77
3	BMW	6888.24	47.36	−.70	−4.31	−37.90
POLO	VW	6819.28	91.09	121.40	108.60	31.75
CORSA	OPEL (GM)	5546.82	−39.91	80.58	39.64	7.97
TOURAN	VW	5282.95	172.35	143.02	156.68	203.19

Table 13: *Product Level Search Index Coefficients - UK*

			Coefficients			
			Column 1		Column 3	
Product	Manufacturer	Av. Sales	*s.r.*	*l.r.*	*s.r.*	*l.r.*
FIESTA	FORD	1618.79	9.27	18.22	12.45	19.69
FOCUS	FORD	1604.20	4.92	6.49	4.13	5.47
ASTRA	VAUXHALL (GM)	1356.52	−1.90	−.11	−1.67	−.87
CORSA	VAUXHALL (GM)	1269.79	1.42	5.61	3.41	9.76
GOLF	VW	1008.23	−2.84	.89	−2.01	2.29
TRANSIT	FORD	901.42	−.40	2.65	.29	2.98
POLO	VW	580.31	−.12	.59	−.15	.73
207	PEUGEOT	557.81	8.83	18.80	7.76	16.41
CLIO	RENAULT	518.75	1.32	1.62	2.21	2.07
MONDEO	FORD	518.16	.41	−.45	1.38	1.92

The figure reveals that sales of both the A- and B-Class have been fairly stable for most of the sample period with the minimum sales for the B-class coinciding with a switch from the first to the second generation

at the end of 2011, while the same happened in case of the A-Class with a switch to the third generation in 2012. The third generation A-Class is a very different-looking vehicle from its two predecessors, in an attempt by Mercedes to appeal to a younger cohort of customers. As the data illustrate, this led to a much larger level of search, but not to a corresponding increase in sales.

Figure 4:

Registrations and Searches: A Class and B Class
Germany: 2004-2014

Source of data: Kraftfahrtbundesamt; Own calculations
A Class indices shifted by 2
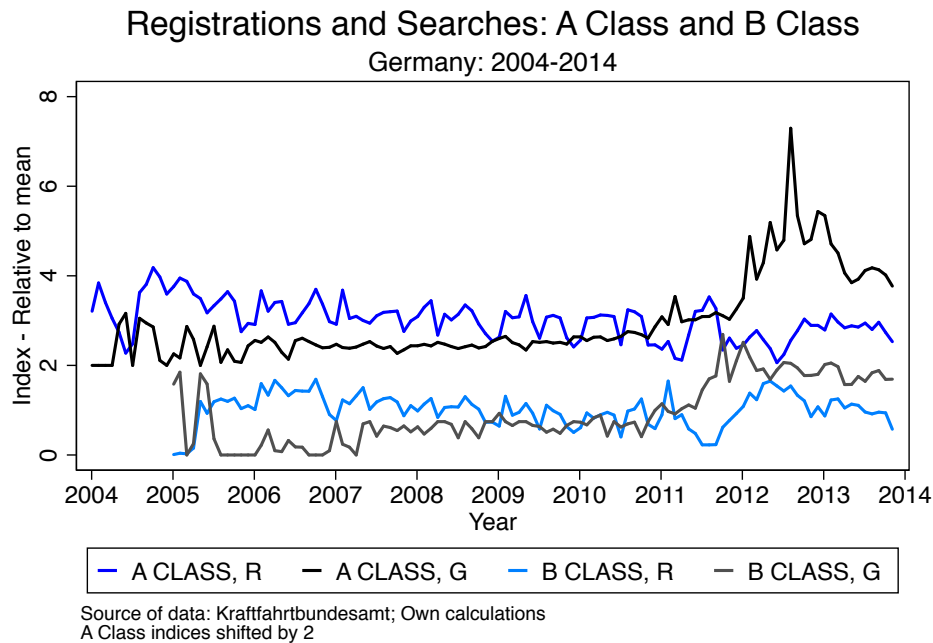
The negative coefficients for short- and long-run search reveal that the A-Class has fared less well than Mercedes might have been hoped for in light of the large amount of interest the new car had generated. The promotion strategy for the B-Class, in contrast, was much more successful in converting searches into sales.

6 Conclusion

Search engine data offer an invaluable opportunity to analyze information resulting from online search activity of hundreds of millions of individuals around the world. Therefore, it is not surprising that there has been rapidly growing interest in the use of such data across the entire range of the social sciences and beyond.

In this growing literature little attention has been paid to modeling the distinct motivations for search. It is often overlooked that online searches that look identical in the available data result from different motivations that lead to different post-search behavior. This creates challenges if the objective is to link search data to economic outcomes. Researchers have also often ignored the econometric challenges that arise from the use of large T, large N panel data. This paper shows that biases can result in both cases, with the larger biases arising from failure to address common correlated effects and serial correlation in our datasets.

Focusing on pre-purchase searches we model biases that arise if fandom and post-purchase searches are not controlled for, predict their effects and confirm that the biases arise in our data. Fortunately, the rich data that can be obtained from GT offer solutions to this problem: additional search terms can be found that help to

condition out forms of search that are not of immediate interest. Nevertheless, much work remains to be done here to help determine how aggregate search can be decomposed into its main components.

The increased length of the available time series promises to offer more opportunities to exploit quasi-random shocks to identify interesting effects across a wide range of economic activity. Nevertheless, using online search data means that the afore-mentioned data properties have to be addressed: serial correlation and non-stationarity. The combination of long high-frequency time series in large panels also introduces the potential that common correlated effects will bias standard panel estimators. As we show in this paper, mean group estimators can be used to remove biases arising from unobserved common effects. As an added benefit these estimators reveal not just the population average, but also the product level coefficients derived from the time series data. This means that individual level heterogeneity can be studied in much greater detail than in other panel data applications.

A last point to emerge from our analysis is that the frequency of the data available has implications for the challenges that the data might hide, e.g. non-stationarity, and also affects the questions that a researcher can ask. The frequency of the data is often determined by the frequency of the data that are matched to the aggregate search data. We use datasets that have weekly and monthly frequencies and find that short-run shocks disappear more quickly from the weekly data than from the monthly data. In our view this is not a reflection of fundamental differences in the markets being analyzed, but a reflection that data averaged over a longer period of time reverts to equilibrium more slowly. Depending on the object of study high frequency search data may, therefore, create an impetus to increase the frequency with which data on 'real world' outcomes is made available.

We expect that online search data will find many more applications in the future. The richness of this type of data is only beginning to be explored. We expect that methods for panel data analysis will continue to evolve as the potential for applications of aggregate search data combined with data on outcomes is more widely appreciated.

References

ARELLANO, M. AND S. BOND (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The Review of Economic Studies*, 58, 277–297.

BANERJEE, A., M. EBERHARDT, AND J. J. READE (2010): "Panel Estimation for Worriers," Tech. Rep. 514.

BRYNJOLFSSON, E., T. GEVA, AND S. REICHMAN (2014): "Crowd-Squared: A New Method for Improving Predictions by Crowd-sourcing Google Trends Keyword Selection," Tech. rep., Working Paper, Center for Digital Business, Massachusetts Institute of Technology.

CHOI, H. AND H. VARIAN (2012): "Predicting the Present with Google Trends," *Economic Record*, 88, 2–9.

CHUDIK, A., K. MOHADDES, M. H. PESARAN, AND M. RAISSI (2015): "Long-run effects in large heterogenous panel data models with cross-sectionally correlated errors," .

CHUDIK, A. AND M. H. PESARAN (2015): "Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models with Weakly Exogenous Regressors," *Journal of Econometrics*, 188, 393–420.

CUMBY, R. E. AND J. HUIZINGA (1992): "Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions," *Econometrica: Journal of the Econometric Society*, 185–195.

DE LOS SANTOS, B., A. HORTAÇSU, AND M. R. WILDENBEEST (2012): "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior," *The American Economic Review*, 102, 2955–2980.

DEL BARRIO CASTRO, T. AND D. R. OSBORN (2014): "The Distribution of Unit Root Test Statistics after Seasonal Adjustment," Tech. rep.

DIETZEL, M. A., N. BRAUN, AND W. SCHÄFERS (2014): "Sentiment-based Commercial Real Estate Forecasting with Google Search Volume Data," *Journal of Property Investment & Finance*, 32, 2–2.

DUTTON, W., G. BLANK, AND D. GROSELJ (2013): "OxIS 2013 Report: Cultures of the Internet," .

EBERHARDT, M., C. HELMERS, AND H. STRAUSS (2013): "Do Spillovers Matter when Estimating Private Returns to R&D?" *Review of Economics and Statistics*, 95, 436–448.

ERICSSON, N. R. AND J. G. MACKINNON (2002): "Distributions of error correction tests for cointegration," *The Econometrics Journal*, 5, 285–318.

GENGENBACH, C., J.-P. URBAIN, AND J. WESTERLUND (2015): "Error Correction Testing in Panels with Common Stochastic Trends," *Journal of Applied Econometrics*.

GINSBERG, J., M. H. MOHEBBI, R. S. PATEL, L. BRAMMER, M. S. SMOLINSKI, AND L. BRILLIANT (2009): "Detecting Influenza Epidemics Using Search Engine Query Data," *Nature*, 457, 1012–1014.

GRANGER, C. W. AND P. NEWBOLD (1974): "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2, 111–120.

HELMERS, C., P. KRISHNAN, AND P. MANASA (2015): "Attention and Saliency on the Internet: Evidence from an Online Recommendation System," *CEPR Discussion Paper No. 10939*.

HSIAO, C. (2014): *Analysis of Panel Data*, vol. 54, Cambridge University Press, 3 ed.

HU, Y., R. Y. DU, AND S. DAMANGIR (2014): "Decomposing the Impact of Advertising: Augmenting Sales with Online Search Data," *Journal of Marketing Research*, 51, 300–319.

LAZER, D., R. KENNEDY, G. KING, AND A. VESPIGNANI (2014): "The Parable of Google Flu: Traps in Big Data Analysis," *Science*, 343, 1203–1205.

LEBO, H. (2014): "Surveying the Digital Future: 2014 Digital Future Project," .

MATHEWS, A. AND C. TUCKER (2014): "Government surveillance and Internet search behavior," *Available at SSRN 2412564*.

MELLON, J. (2013): "Where and When Can We Use Google Trends to Measure Issue Salience?" *PS: Political Science & Politics*, 46, 280–290.

——— (2014): "Internet Search Data and Issue Salience: The Properties of Google Trends as a Measure of Issue Salience," *Journal of Elections, Public Opinion & Parties*, 24, 45–72.

MORAGA-GONZÁLEZ, J. L., Z. SÁNDOR, AND M. R. WILDENBEEST (2015): "Consumer Search and Prices in the Automobile Market," CEPR DP 10487.

NUTI, S. V., B. WAYDA, I. RANASINGHE, S. WANG, R. P. DREYER, S. I. CHEN, AND K. MURUGIAH (2014): "The Use of Google Trends in Health Care Research: A Systematic Review," *PLoS ONE*, 9, e109583.

ORMEROD, P., R. NYMAN, AND R. A. BENTLEY (2014): "Nowcasting Economic and Social Data: When and Why Search Engine Data Fails, an Illustration Using Google Flu Trends," *arXiv preprint arXiv:1408.0699*.

PAVLICEK, J. AND L. KRISTOUFEK (2014): "Can Google Searches Help Nowcast and Forecast Unemployment Rates in the Visegrad Group Countries?" *arXiv preprint arXiv:1408.6639*.

PESARAN, M. H. (2004): "General Diagnostic Tests for Cross Section Dependence in Panels," CESifo Working Paper Series 1229, CESifo Group Munich.

——— (2006): "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74, 967–1012.

——— (2007): "A Simple Panel Unit Root Test in the Presence of Cross-section Dependence," *Journal of Applied Econometrics*, 22, 265–312.

PESARAN, M. H. AND E. TOSETTI (2011): "Large Panels with Common Factors and Spatial Correlation," *Journal of Econometrics*, 161, 182–202.

POLGREEN, P. M., Y. CHEN, D. M. PENNOCK, F. D. NELSON, AND R. A. WEINSTEIN (2008): "Using Internet Searches for Influenza Surveillance," *Clinical Infectious Diseases*, 47, 1443–1448.

PREIS, T., H. S. MOAT, AND H. E. STANLEY (2013): "Quantifying Trading Behavior in Financial Markets Using Google Trends," *Scientific Reports*, 3.

SOVINSKY GOEREE, M. (2008): "Limited Information and Advertising in the US Personal Computer Industry," *Econometrica*, 76, 1017–1074.

STATISTISCHES BUNDESAMT (2015): "Private Haushalte in der Informationsgesellschaft - Nutzung von Informations- und Kommunikationstechnologien," .

PURCELL, K., J. BRENNER, AND L. RAINIE (2012): "Search Engine use 2012," .

ROSSANA, R. J. AND J. J. SEATER (1995): "Temporal Aggregation and Economic Time Series," *Journal of Business & Economic Statistics*, 13, 441–451.

SANTILLANA, M., D. W. ZHANG, B. M. ALTHOUSE, AND J. W. AYERS (2014): "What Can Digital Disease Detection Learn from (an External Revision to) Google Flu Trends?" *American Journal of Preventive Medicine*, 47, 341–347.

VARIAN, H. R. (2014): "Big data: New Tricks for Econometrics," *The Journal of Economic Perspectives*, 28, 3–27.

VOSEN, S. AND T. SCHMIDT (2011): "Forecasting Private Consumption: Survey-based Indicators vs. Google Trends," *Journal of Forecasting*, 30, 565–578.

WOOLDRIDGE, J. (2012): *Introductory Econometrics: A Modern Approach*, Cengage Learning.

A GT keyword selection

One of the main challenges in the literature relying on GT data is to find the right keywords on which to base the analysis. The approach generally followed consists in finding a reliable benchmark and then determining the combination of keywords which best approximates the benchmark.

An important issue in key word selection concerns language confusion, i.e. the fact that brands or products may correspond to common names. This issue is often country-specific as brand or product names are generally not translated from one language to another and may therefore correspond to a common name in one country but not in others.

There are several possible ways to address language confusion systematically. One possible strategy to overcome the language confusion issue is to use combinations of words, linking related products. Another possible strategy is to use the categories of searches in GT, restricted to the appropriate category of searches. This means building a rough concordance between categories of products and Google categories. One limitation of the latter approach is a lack of information on how Google constructs its search categories.

Finally, some features regarding strings in GT need to be considered to construct keywords, in particular the treatment of special characters:

- *Accents, special characters, numbers and punctuation signs:* these are taken into account by Google Trends ("Société" gives different results from "Societe", "apple." gives different results than "apple") – as a result they need to be properly encoded when retrieving the series;

- *Case:* Google Trends is case-insensitive ("apple" gives the same results as "APPLE");

- *Logical operators:* some signs cannot be searched for as such due to their use as logical operators in Google Trends (for example ","; "+"). The hyphen ("-") is considered as a regular character by Google, except if preceded by a space, then it corresponds to the logical operator "excluding";

- *Apostrophe ('):* this sign is currently automatically deleted by Google Trends. This is an issue for certain brand names containing an apostrophe (e.g. L'OREAL), as it is currently not possible to catch the search volume for the name containing the apostrophe; it is only possible to get results for OREAL or LOREAL;

B GT data extraction

Data from Google Trends can be downloaded in three ways: through the dedicated user interface, an application programming interface (API), or using a site crawler. The dedicated user interface requires manually entering each query of interest, which makes it unsuitable for systematic analysis of a large number of terms. As for crawlers, Google limits the ability of such tools to obtain data from its Google Search service. An API therefore appears the most appropriate approach for data collection. No API is made available by Google

on Google trends. In the absence of an official Google API, a number of unauthorized packages/scripts exist that exploit this facility to automatically download data from Google Trends. We used a similar approach to that of the **googletrend** package which requires a browser that downloads the data to a local directory. Our own script is available on GitHub (https://github.com/jirzii/R4GT). The script uses the R statistics package's **browseURL** command.

We downloaded our GT data between September and November of 2014 for Germany and the United Kingdom. We downloaded weekly and where necessary monthly series between 2004 and 2014 for 53 (DE) and 373 (UK) manufacturer and model names. We chose the longest possible series that we were able to match with our market level data and we used manufacturer and model names as the search terms.

We downloaded data on each search query individually to obtain maximal detail on the variance of the series. We also downloaded each series together with searches for a very popular and a middle of the range manufacturer and model. These additional searches provide data about relative search intensities for each class of information.

Two main complications arose in creating functioning scripts:

1. Identifying the correct syntax to specify complex queries that involve logical *and* and logical *or* operations;

2. Dealing with data that was only available at a monthly frequency (see Section C below).

The syntax to create correct queries for downloads from GT can most easily detected if complex manual downloads are initially undertaken with the Google Chrome browser. The syntax for the query appears in the browser's window once the query has been entered on the webpage for GT via the GT interface.

C GT weekly to monthly data conversion

Depending on the search volumes, GT series are provided either at the monthly or weekly level. In order to compare the different search terms, it is necessary to aggregate data to a common frequency level, which is not straightforward due to the normalisation applied to the series.

Google Trends series correspond to the raw number of searches done for one term normalised by the total number of search queries in the geographical region of interest, rescaled to an index taking a maximum value of 100 over the period, which may be summarised with the following formula:

$$G_t = \frac{(x_t/X_t)}{\max(x_t/X_t)} \times 100 \tag{16}$$

where G_t is the Google Trends figure in time t, x_t is the number of searches for the keyword x in time t and X_t is the total number of Google searches in time t. Depending on the series, t corresponds to a week or a month. Considering a month m with four weeks (w_1, w_2, w_3, w_4) the monthly indicator may be written as follows:

$$G_m = \frac{(x_m/X_m)}{\max{(x_m/X_m)}} \times 100 = \frac{((x_{w_1} + x_{w_2} + x_{w_3} + x_{w_4})/(X_{w_1} + X_{w_2} + X_{w_3} + X_{w_4}))}{\max{(x_m/X_m)}} \times 100 \quad (17)$$

where $\max{(x_m/X_m)}$ corresponds to the monthly maximum of the ratio over the entire period. The average of the weekly indicators, on the other hand, is:

$$\overline{G}_w = \frac{(x_{w_1}/X_{w_1} + x_{w_2}/X_{w_2} + x_{w_3}/X_{w_3} + x_{w_4}/X_{w_4})}{4 \times \max{(x_w/X_w)}} \times 100 \quad (18)$$

where $\max{(x_w/X_w)}$ corresponds to the weekly maximum of the ratio over the whole period. This is equivalent to the monthly indicator G_m if the two following conditions are verified:

- either the overall number of Google searches X_w or the proportion of searches corresponding to the keyword x_w/X_w is stable over all weeks in each month;

- the weekly and monthly maximum proportions of searches corresponding to the keyword over the whole period are equivalent: $\max{(x_m/X_m)} \sim \max{(x_w/X_w)}$.

The second condition is more likely not to be satisfied, especially for series with a high variation. In that case the series of converted weekly data $\overline{G}_w$ are proportional to the series of monthly data G_m, only with lower variation:

$\overline{G}_w = G_m \times \frac{(x_m/X_m)}{\max(x_w/X_w)}$.

In practice, in order to convert the weekly to monthly series, we consider the weekly dataset, expand each observation by seven, increment the date by one for each line and then average out the data at the monthly level.

D   Data from Kraftfahrtbundesamt

In Germany car license plates are issued by municipalities. Information about registered vehicles is collected on a central register (ZFZR) maintained by the KBA. Based on these data the KBA provides a number of statistical tables about the number of new registrations each month (FZ10) by manufacturer and model.

KBA also provide annual data on the number of deregistered cars (FZ7) by manufacturer, resale (FZ18) by manufacturer and total stock of vehicles (FZ17) by manufacturer.

We downloaded these data from the internet in September 2014 and cleaned and collated the data. Registration data were added to a panel with monthly frequency at the manufacturer and model level covering the period between January 2004 and January 2014. Data on total stock of registered vehicles, resale and deregistration was added to a panel with annual frequency at the manufacturer level.

KBA's website for vehicle statistics is available here (we accessed the website in September 2014):

http://tinyurl.com/k8lxms7

E   Descriptives

Table A-1: *Summary Statistics - Germany, 486 Products, 127 Months*

Variable	Mean	Std. Dev.	Min.	Max.	N
Registrations	1164.168	2668.832	1	44280	33951
Registrations (s.a.)	-56.361	1063.042	-22294	20059	28026
Product SI	6.385	12.059	0	100	33951
Product SI (s.a.)	0.406	7.961	-100	100	28026
Brand SI	32.063	24.333	0	95.75	33951
Brand SI (s.a.)	0.122	4.688	-70.387	38.345	28026
Scrappage SI	0.469	3.888	0	86.129	33951
Scrappage SI (s.a.)	0.014	5.87	-86.129	86.129	28026

Correlation Matrix - Germany ($N = 20169$)

	Registrations$_t$	Product SI$_{t-1}$	Brand SI$_{t-2}$	Scrappage SI$_t$
Registrations$_t$	1			
Product SI$_{t-1}$	0.0772***	1		
Brand SI$_{t-2}$	0.154***	0.0819***	1	
Scrappage SI$_t$	0.155***	0.0631***	0.487***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-2: *Summary statistics - United Kingdom, 363 Products, 463 Weeks*

Variable	Mean	Std. Dev.	Min.	Max.	N
Registrations	57.479	239.547	0	9291	266688
Registrations, (s.a.)	-10.484	214.892	-8774	7836	236160
Product SI	8.305	14.195	0	100	174551
Product SI, (s.a.)	0.560	6.019	-100	85	154570
Brand SI	38.002	22.067	0	100	266688
Brand SI, (s.a.)	-1.168	4.867	-53	80	236160
Scrappage SI	1.592	10.255	0	100	266688
Scrappage SI, (s.a.)	0	15.597	-100	100	236160

<div align="center">

Correlation Matrix - United Kingdom ($N = 92886$)

</div>

	Registrations$_t$	Product SI$_{t-4}$	Brand SI$_{t-4}$	Scrappage SI$_{t-3}$
Registrations$_t$	1			
Product SI$_{t-4}$	0.0768***	1		
Brand SI$_{t-4}$	0.0335***	0.0809***	1	
Scrappage SI$_{t-3}$	0.0225***	0.0544***	0.0457***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

F   Monthly UK Data

<div align="center">

Table A-3: *Unit Root Tests - United Kingdom, Monthly Data*

</div>

	Sales		Product SI	
	not adj.	(s.a.)	not adj.	(s.a.)
Observations	19516	19516	19516	19516
	Pesaran 2007 CIPS Test: H0 - All panels are non-stationary			
p-value, lags 2	$5.46e-25$	$2.64e-55$	$7.53e-38$	$1.87e-59$
Z-score, lags 2	-10.26	-15.62	-12.81	-16.22
p-value, lags 4	0.000000196	$4.53e-16$	0.00709	$1.14e-12$
Z-score, lags 4	-5.072	-8.039	-2.453	-7.017
p-value, lags 6	$3.95e-09$	$6.98e-20$	0.175	0.000379
Z-score, lags 6	-5.771	-9.053	-0.934	-3.367
p-value, lags 8	0.335	$2.64e-11$	1.000	0.00457
Z-score, lags 8	-0.426	-6.563	3.926	-2.607

The table is based on 226 strongly balanced panels and 389 weeks.

G Robustness Checks

Table A-4: *CCE MG Results for Germany, Monthly Data*

	(1)	(2)	(3)	(4)	(5)
Long Run Coefficient					
Product SI $_{t-1}$	34.54$***$	38.02$***$	36.83$***$	32.20$***$	32.36$***$
	(7.614)	(8.147)	(9.859)	(7.571)	(7.347)
Short Run Coefficient					
Adjustment rate	$-0.772***$	$-0.771***$	$-0.754***$	$-0.773***$	$-0.771***$
(Registrations$_{t-1}$)	(0.0188)	(0.0193)	(0.0188)	(0.0193)	(0.0190)
$\triangle$ Product SI $_t$	22.84$***$	24.88$***$	24.57$***$	21.76$***$	20.62$***$
	(5.309)	(5.363)	(6.443)	(5.546)	(5.808)
Brand SI $_{t-2}$	6.691	6.386	6.350	7.854$*$	6.326
	(3.800)	(3.912)	(4.650)	(3.608)	(3.808)
Scrappage SI $_t$	16.72$*$	16.65$*$	-24.17	15.95$*$	18.29$*$
	(7.849)	(8.391)	(31.26)	(7.932)	(8.625)
Registrations$_{t-9}$	0.0413$***$	0.0474$***$	0.0379$**$	0.0393$**$	0.0282$*$
	(0.0122)	(0.0126)	(0.0139)	(0.0121)	(0.0118)
Registrations$_{t-12}$	$-0.326***$	$-0.320***$	$-0.296***$	$-0.313***$	$-0.325***$
	(0.0148)	(0.0143)	(0.0156)	(0.0145)	(0.0142)
Registrations$_{t-13}$	0.0929$***$	0.0886$***$	0.0954$***$	0.0940$***$	0.0948$***$
	(0.0132)	(0.0146)	(0.0148)	(0.0136)	(0.0138)
Further variables		$\triangle$ Prd. SI CA	$\triangle$ Prd./Br./Scr. SI	Lag 8	Lags 8, 13
			CA, $\triangle$ Br./Scr. SI	Scr. SI CA	Prd. SI CA
Constant	-31.98	-39.18	-31.73	-5.554	-2.193
	(25.21)	(26.80)	(31.81)	(20.94)	(23.15)
Products	200	198	189	198	195
Observations	16284	16204	15822	16204	16081
CD Test	0.77	0.77	1.84	0.10	0.42
CD Test p	0.44	0.44	0.065	0.919	0.672
CH Test AR(1)	0.0301	0.0317	0.0457	0.0267	0.0455
CH Test AR(2)	0.359	0.406	0.116	0.329	0.272
CH Test AR(11)	0.632	0.611	0.631	0.611	0.512
CH Test AR(12)	0.0911	0.0880	0.694	0.179	0.143

Notes: $***$, $**$, $*$, $\dagger$ denote significance at the 0.1%, 1%, 5%, 10% level. Time period is 2004-2014. The sample contains 200 products with at least 24 months of positive sales. All models contain lags 1-6,8,9,12,13 of registrations, lags 0,1,3,6,11,12,13 of registration cross-section averages (CA) and the CA of the change in registrations, brand SI CAs at lags 2,4,5 and 6 and lags 0,3,6-9 and 11 of the scrappage SI and lags 0,9,11,12 of the scrappage SI CAs.

Diagnostics: CH (x) Test: Cumby and Huizinga (1992) test for H_0 of no residual serial correlation at lag x (p-values). CD Test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals, (p-values) and statistic. Test restricted to 149 sufficiently balanced product panels.

Table A-5: *CCE MG ECM Results for the UK, Weekly Data*

	(1)	(2)	(3)	(4)	(5)
Error Correction Term					
Product SI $_{t-4}$	0.621∗∗∗	0.604∗∗∗	0.564∗∗	0.746∗∗∗	0.748∗∗∗
	(0.176)	(0.178)	(0.184)	(0.200)	(0.200)
SR					
Adjustment rate	−0.877∗∗∗	−0.877∗∗∗	−0.871∗∗∗	−0.865∗∗∗	−0.864∗∗∗
(Registrations$_{t-4}$)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
$\tilde{\triangle}$ Product SI $_{t-4}$	0.317∗∗	0.306∗	0.228	0.382∗∗	0.371∗
	(0.131)	(0.137)	(0.144)	(0.147)	(0.148)
Brand SI $_{t-25}$	−0.262	−0.249	−0.259	−0.291	−0.315
	(0.178)	(0.176)	(0.174)	(0.192)	(0.194)
Scrappage SI $_{t-16}$	0.117∗	0.125∗	0.117∗	0.142∗	0.141∗
	(0.0548)	(0.0561)	(0.0589)	(0.0640)	(0.0638)
Registrations$_{t-1}$	0.192∗∗∗	0.189∗∗∗	0.182∗∗∗		
	(0.008)	(0.008)	(0.008)		
Registrations$_{t-9}$	0.0630∗∗∗	0.0628∗∗∗	0.0622∗∗∗	0.0776∗∗∗	0.0769∗∗∗
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Registrations$_{t-26}$	0.0925∗∗∗	0.0958∗∗∗	0.0966∗∗∗	0.0979∗∗∗	0.0976∗∗∗
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Further variables		$\triangle$ Prd. SI CA	$\triangle$ Prd./Br./Scr. SI		Dropped Reg.
			CA, $\triangle$ Br./Scr. SI		$_{t-11}$
Constant	−2.364	−2.174	−2.742	−2.347	−2.269
	(3.642)	(3.573)	(3.600)	(4.089)	(4.046)
Products	226	226	226	226	226
Observations	86, 558	86, 558	86, 558	86, 558	86, 558
CD Test	4.95	5.05	5.10	4.33	4.27
CD Test p	0.00	0.00	0.00	0.00	0.00
CH Test AR(1)	0.169	0.138	0.102	0.000	0.000
CH Test AR(6)	0.877	0.775	0.757	0.590	0.721
CH Test AR(16)	0.016	0.008	0.003	0.000	0.001
CH Test AR(18)	0.360	0.497	0.482	0.554	0.549

Notes: ***, **, * denote significance at the 0.1%, 1%, 5% level. Time period 2004-2011. The sample contains 226 products with at least 3 months of consecutive sales and 52 weeks of positive sales. All models contain lags 1-4,7,8,9,11,13-15,20-24,26,27 of registrations, lags 1-4,6-9,12-15,19,20,23,26,27 of registration cross-section averages (CA) and the CA of the change in registrations. Models 2,3 and 4 contain brand SI CAs at lags 2,7,8,9 and 10. Models 3 and 4 contain lags 3,8,12,16 and 20 of the scrappage SI and lags 1,3-6,10,16-19,21,23-27 of the scrappage SI CAs.

Diagnostics: CH (x) Test: Cumby and Huizinga (1992) test for H_0 of no residual serial correlation at lag x (p-values). CD Test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals, (p-values) and statistic. Test restricted to 144 highly balanced product panels.

Table A-6: Comparing Long Run Effects from CS-DL and ARDL Models - Germany

Model	Products	$\hat{\theta}_{g,t}$	$\frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	$\hat{\theta}_{g,t} - \frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	Std. Err.	p	t
1	200	83.15	92.43	-9.28	36.47	0.80	-0.255
2	198	82.14	13.64	68.50	75.82	0.37	0.90
3	189	81.03	86.48	-5.45	41.38	0.90	-0.13
4	198	81.67	42.76	38.91	25.09	0.12	1.55
5	195	89.41	62.82	26.59	25.82	0.30	1.03

Table A-7: Comparing Long Run Effects from CS-DL and ARDL Models - UK

Model	Products	$\hat{\theta}_{g,t}$	$\frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	$\hat{\theta}_{g,t} - \frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	Std. Err.	p	t
1	226	2.77	2.28	0.49	0.70	0.48	0.70
2	226	2.75	2.15	0.60	0.70	0.39	0.86
3	226	2.78	2.01	0.76	0.70	0.27	1.10
4	226	2.77	-2.52	5.29	2.57	0.04	2.06
5	226	2.77	0.32	2.45	1.14	0.03	2.14

H Results Without Seasonal Adjustment

In this section we present results of estimating MG models with data that is not seasonally adjusted in Tables A-10 and A-11. We also present comparisons of the long-run coefficients for the effect of online search on sales for each model using the methodology proposed by Chudik et al. (2015). The first column in Tables A-10 and A-11 provides results from estimating the preferred model presented in Section 5.2. The remaining columns present models that are adapted to minimize Pesaran (2004)'s CD test statistic. These models were arrived at by testing additions and exclusions of individual cross-section average terms to the model in column 1, one at a time.

Table A-8: Comparing Long Run Effects from CS-DL and ARDL Models - Germany

Model	Products	$\hat{\theta}_{g,t}$	$\frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	$\hat{\theta}_{g,t} - \frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	Std. Err.	p	t
1	189	69.94	39.34	30.60	28.03	0.28	1.09
2	189	83.90	18.13	65.77	68.22	0.34	0.96
3	189	83.90	38.51	45.39	80.92	0.58	0.56
4	181	81.72	50.58	31.14	39.44	0.43	0.79

Table A-9: Comparing Long Run Effects from CS-DL and ARDL Models - UK

Model	Products	$\hat{\theta}_{g,t}$	$\frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	$\hat{\theta}_{g,t} - \frac{\hat{\beta}_{g,t}}{1-\hat{\gamma}_{t,k}}$	Std. Err.	p	t
1	226	0.17	6.94	-6.77	3.52	0.06	-1.92
2	226	0.35	-1.56	1.91	2.01	0.34	0.95
3	226	0.28	-0.74	1.02	1.50	0.50	0.68
4	226	0.22	0.11	0.12	1.15	0.92	0.10

Table A-10: *CCE MG Results for Germany, Monthly Data*

	(1)	(2)	(3)	(4)
Long Run Coefficient				
Product SI $_{t-1}$	30.17∗∗∗	25.12∗∗∗	23.94∗∗∗	25.90∗∗∗
	(9.031)	(7.255)	(6.943)	(6.518)
Short Run Coefficient				
Adjustment rate	−0.780∗∗∗	−0.874∗∗∗	−0.863∗∗∗	−0.837∗∗∗
(Registrations$_{t-1}$)	(0.0160)	(0.0186)	(0.0188)	(0.0188)
$\triangle$ Product SI $_t$	20.34∗∗	23.56∗∗∗	26.12∗∗∗	23.77∗∗
	(7.728)	(7.017)	(7.139)	(8.966)
Brand SI $_{t-2}$	8.743∗	8.877∗∗	9.385∗∗	8.548∗∗
	(3.463)	(3.223)	(3.383)	(3.130)
Scrappage SI $_t$	74.46	17.35	18.14	−13298.9
	(41.60)	(15.35)	(15.57)	(13181.9)
Registrations$_{t-9}$	0.0508∗∗∗	0.0515∗∗∗	0.0450∗∗∗	0.0363∗∗
	(0.0109)	(0.0128)	(0.0126)	(0.0125)
Registrations$_{t-12}$	0.0337∗∗	0.0371∗∗	0.0323∗	0.0350∗
	(0.0125)	(0.0139)	(0.0154)	(0.0143)
Registrations$_{t-13}$	−0.0421∗∗∗	−0.0315∗	−0.0255	−0.0358∗
	(0.0127)	(0.0129)	(0.0134)	(0.0155)
Constant	−722.4∗	−697.7∗	−690.5∗	−434.8
	(292.4)	(279.9)	(268.6)	(260.9)
Products	189	189	189	181
Observations	15, 822	15, 822	15, 822	15, 457
CD Test	7.04	3.76	3.15	3.89
CD Test p	0.00	0.00	0.00	0.00
CH Test AR(1)	0.0529	0.183	0.234	0.184
CH Test AR(2)	0.885	0.114	0.118	0.169
CH Test AR(11)	0.643	0.0981	0.0807	0.0726
CH Test AR(12)	0.216	0.0423	0.0151	0.0266

Notes: ***, **, *, † denote significance at the 0.1%, 1%, 5%, 10% level. Time period is 2004-2014. The sample contains 200 products with at least 24 months of positive sales. All models contain lags 1-6,8,9,12,13 of registrations, lags 0,1,3,6,11,12,13 of registration cross-section averages (CA) and the CA of the change in registrations. Models 2 and 3 contain brand SI CAs at lags 2,4,5 and 6. Models 3 and 4 contain lags 0,3,6-9 and 11 of the scrappage SI and lags 0,9,11,12 of the scrappage SI CAs.

Diagnostics: CH (x) Test: Cumby and Huizinga (1992) test for H_0 of no residual serial correlation at lag x (p-values). CD Test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals, (p-values) and statistic. Test restricted to 149 sufficiently balanced product panels.

Table A-11: *CCE MG ECM Results for the UK, Weekly Data*

	(1)	(2)	(3)	(4)
Error Correction Term				
Product SI $_{t-4}$	0.857∗∗∗	0.987∗∗∗	0.999∗∗∗	0.958∗∗∗
	(0.212)	(0.222)	(0.224)	(0.225)
SR				
Adjustment rate	−0.898∗∗∗	−0.880∗∗∗	−0.880∗∗∗	−0.878∗∗∗
(Registrations$_{t-4}$)	(0.00696)	(0.00763)	(0.00763)	(0.00766)
$\tilde{\triangle}$ Product SI $_{t-4}$	0.446∗∗	0.518∗∗∗	0.584∗∗∗	0.583∗∗∗
	(0.147)	(0.151)	(0.154)	(0.161)
Brand SI $_{t-25}$	−0.00592	−0.0502	−0.0387	−0.0284
	(0.230)	(0.237)	(0.240)	(0.237)
Scrappage SI $_{t-16}$	0.0528	0.0718	0.0797	0.0548
	(0.0683)	(0.0761)	(0.0751)	(0.0691)
Registrations$_{t-1}$	0.184∗∗∗	0.159∗∗∗	0.157∗∗∗	0.154∗∗∗
	(0.00788)	(0.00854)	(0.00856)	(0.00858)
Registrations$_{t-9}$	0.0649∗∗∗	0.0826∗∗∗	0.0821∗∗∗	0.0830∗∗∗
	(0.00635)	(0.00683)	(0.00683)	(0.00688)
Registrations$_{t-26}$	0.198∗∗∗	0.176∗∗∗	0.176∗∗∗	0.177∗∗∗
	(0.0106)	(0.0104)	(0.0105)	(0.0105)
Constant	−0.148	6.448	7.249	2.892
	(13.71)	(17.06)	(17.26)	(17.37)
CD Test	6.82	4.01	3.75	3.81
CD Test p	0.00	0.00	0.00	0.00
CH Test AR(1)	0.120	0.0518	0.0419	0.0452
CH Test AR(5)	0.0129	0.103	0.0753	0.122
CH Test AR(10)	0.00284	0.0214	0.0286	0.0602
CH Test AR(16)	0.0221	0.105	0.119	0.0757

Notes: ***, **, * denote significance at the 0.1%, 1%, 5% level. Time period 2004-2011. The sample contains 226 products with at least 3 months of consecutive sales and 52 weeks of positive sales. All models contain lags 1-4,7,8,9,11,13-15,20-24,26,27 of registrations, lags 1-4,6-9,12-15,19,20,23,26,27 of registration cross-section averages (CA) and the CA of the change in registrations. Models 2,3 and 4 contain brand SI CAs at lags 2,7,8,9 and 10. Models 3 and 4 contain lags 3,8,12,16 and 20 of the scrappage SI and lags 1,3-6,10,16-19,21,23-27 of the scrappage SI CAs.

Diagnostics: CH (x) Test: Cumby and Huizinga (1992) test for H_0 of no residual serial correlation at lag x (p-values). CD Test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals, (p-values) and statistic. Test restricted to 144 highly balanced product panels.