A Review of Empirical Operations Management over the Last Two Decades

Forthcoming in
Manufacturing and Service Operations Management (M&SOM)

Christian Terwiesch
The Wharton School
University of Pennsylvania
573 JMHH, Philadelphia, PA 19104
terwiesch@wharton.upenn.edu

Marcelo Olivares
Universidad de Chile
Beauchef 851
Santiago, Chile
molivares@u.uchile.cl

Bradley R. Staats
University of North Carolina at Chapel Hill
Campus Box 3490, McColl Building
Chapel Hill, NC 27599-3490
bstaats@unc.edu

Vishal Gaur
Johnson School, Cornell University
377 Sage Hall
Ithaca, NY 14850
vg77@cornell.edu

August 31, 2018

Acknowledgments

We gratefully acknowledge funding through the Fishman Davidson Center at Wharton. We thank Maksim Yakovlev for his research assistance. We are especially thankful for the help of our colleagues who have participated over the last decade in the Workshop for Empirical Operations Management. Chris Tang and two anonymous referees have provided invaluable comments on an earlier version of this article.
Abstract
We develop a database of all empirical research related to Operations Management in the journals Management Science, Manufacturing and Service Operations Management (M&SOM), and Production and Operations Management (POM) from the beginning of 1999 to the end of 2016. This database includes 267 empirical papers. We analyze the set of empirical papers to look for longitudinal trends and other bibliometric patterns. In particular, we show that (a) empirical research as a whole is gaining in popularity as measured by the publication rates in these three journals, (b) empirical papers in M&SOM are more likely to get citations than non-empirical papers, and (c) researchers are now more commonly using instrumental variables and are more likely to consider endogeneity challenges in their research design. Using our database, we propose three dimensions on which empirical Operations Management papers can be compared, including their main objective, their data sources, and their identification strategy.
1. Introduction
The dominant research methodology of the first scholar of Operations Management, Frederick Winslow Taylor, as well as Carl Barth, Henry Gantt, and Frank and Lillian Gilbreth, was the empirical analysis of carefully collected data (Smiddy and Naum 1954). A century later, the early scholars’ empirical approach based on time sheets and stop watches was replaced by a normative research paradigm based primarily on analytical models. Among all the Operations Management articles published in Management Science in the year of 1998, only one included empirical data. M& SOM did not publish a single empirical paper from 2002 to 2005, despite the journal featuring two empirical papers in its inaugural issue in 1999. POM published an average of one empirical paper per year in this time frame. Over the course of the 20th century, the empirical base for Operations Management lagged behind other disciplines, as Marshall Fisher (2007) observed in M& SOM.

But, as we write this article in 2018, this pattern of scarcity of empirical research has changed. From 2006 to 2010, Management Science published, on average, five empirical Operations Management papers per year. In M& SOM, that number has grown to three per year and in POM the average was 1.6 per year. From 2011 to 2015, the average number of empirical papers in Management Science further increased to seven papers per year. Similarly, M& SOM also increased its rate of publishing empirical papers during this time to five per year while POM increased its rate to ten per year.

Though six empirical operations management papers published per year in M& SOM is still only slightly more than 10% of all published M& SOM papers (from 2011 to 2015, the journal published an average of 50 papers per year), this data suggests the emergence of a new stream of work. Changes towards a more empirical approach to research are by no means limited to Operations Management. Economics, the discipline that historically has had a major influence
on Operations Management, is going through a similar transition. In fact, the trend in Economics has already advanced much further. Daniel Hamermesh (2013) showed that while in 1980, more than 50% of the papers in top Economics journals were entirely theoretical, this number had fallen to only 28% in 2013. Moreover, Hamermesh reported another trend: In 1993, only 8.8% of the articles were based on data collected by the author (as opposed to being a theoretical paper, a paper based on a publicly available data set, or a paper based on experimental data). This number grew to 34% by 2011.

In this article, we document the emergence of this new line of Empirical Operations Management research. Our main aim and contribution is to provide an overview of empirical research in Operations Management and to propose three dimensions along which empirical papers can be compared.

As empiricists, we want this documentation to be based on a careful bibliometric analysis of the empirical research published in the journals Management Science, M&SOM, and POM over the last two decades. This analysis yields a number of interesting observations:

- We show that empirical OM is on the rise in all three journals considered in our study period, with the number of empirical papers published across the three journals moving up from less than five per year in the first years to well over 20 in the most recent years.
- This trend comes along with an increasing focus on health care and new methodologies. In particular, we present a text-mining analysis that searches for keywords throughout abstracts and bodies of the papers, revealing an increase in the use of instrumental variables and a larger effort to address endogeneity.
- We also show that empirical OM papers are frequently cited. Controlling for publication year, empirical OM papers get more citations than non-empirical papers. This is
remarkable because there are fewer empirical papers published, so one might expect those to be at a disadvantage.

Our article is intended to be useful to Operations Management researchers interested in conducting empirical research, especially those new to the field, such as doctoral students or junior faculty. It also is important to note what this article does not do – it does not survey the substantive contributions of empirical research to different topics in Operations Management, such as manufacturing, supply chain management, services, or health care. In other words, we focus on the empirical methods, not the findings.

2. Developing the Database

The foundation of our work is a database of all empirical Operations Management papers in Management Science, M&SOM, and POM that we constructed. In assembling this database, we included the following articles:

1. Articles published in M&SOM and POM, articles identified by the Operations Management department at Management Science, and all articles published by operations management scholars (as defined by their primary affiliations) in any department of Management Science.

2. Articles that used empirical data from databases or corporate records as their primary data collection methodology.


Looking at the Financial Times ranking methodology as well as the journal list published by UT Dallas, Management Science, M&SOM, and POM, and the Journal of Operations Management are the leading scholarly journals in Operations Management. Of those, the Journal of Operations
Management has published the most empirical research, much of which has already been the subject of a review article (Gupta, Verma and Victorino 2006). We also did not consider Operations Research as it rarely publishes empirical research. Thus, we restricted ourselves to the journals Management Science, M&SOM, and POM, which publish a mixture of empirical and non-empirical work and have not been reviewed recently from an empirical perspective.

Management Science has multiple departments relevant to Operations Management. In addition to including all articles identified by the Operations Management department, we also included Operations Management work published by other departments. To avoid drawing boundaries around the topic of Operations Management (For example, is a paper measuring price dispersion an Operations Management paper? Is product development part of Operations Management?), we included all empirical work from scholars who are active in the field of Operations Management regardless of which department published the work.

The second criterion limits our scope to articles that draw upon data derived from some kind of record, not a perceptual survey measure. To illustrate, a paper that evaluates employee engagement based on the number of improvement suggestions made last quarter would be included in the database, but a paper that uses a perceptual scale such as “how actively are employees engaged in submitting improvement suggestions (1: not at all, to 5: very engaged)” would not be included. In addition, it excluded exclusively experimental work in the lab because that has been reviewed elsewhere (Gino and Pisano 2008; Bendoly et al. 2010).

We began the search in 1999, which is the year M&SOM began publishing (third criterion). Having defined the criteria for inclusion, the next step was to identify those papers that met the three criteria. Together, with two research assistants, we went through each issue of the included journals in order to identify relevant articles. To ensure the quality of our work, we also completed
the following activities:

- Each author of this article picked five schools with a track record of empirical research and checked the school’s faculty websites against the database.

- We looked at our own empirical papers (close to 40 across the four authors) and compared them to how the research assistants coded them.

- We reached out to five individuals with related papers to check their work in the database.

- We contacted attendees of the Workshop on Empirical Research in Operations Management, a group of over 150 scholars with interest and experience in the topic, and asked them to ensure that all their work was captured in the list. In a second iteration, as part of the review process, we went back to this group, showed them our database, and asked them to check the accuracy of how our database captured their work. We received three requests to include a paper, in two of which the authors were “on the fence.” We included all three papers.

Despite our inspection process, it is possible that articles have been missed. However, we believe that our rigorous and robust process has identified almost all articles. To maximize transparency, we make this database publicly available. It can be accessed via the web-site of the Workshop on Empirical Research in Operations Management at Wharton. We believe that this database constitutes a valuable resource for the field. It also allowed us to engage other authors into the content creation and inspection process.

3. Analysis of Empirical Publications

Having collected the data on empirical operations publications, we then analyzed it, looking at the overall trend, trends across methodologies and industries, and a comparison of citation rates between empirical and analytical papers.
3.1 Overall Trend

We began by looking at the trend in publications over time. As seen in Figure 1, over the last two decades, the number of empirical operations publications has increased substantially. Although variability exists from year to year, the trend is very clear and consistent across the three journals. As seen in Figure 1, the number of empirical papers published across the three journals moved up from an average of less than five in the early years to well over 20 in the most recent years.

![Figure 1: Empirical operations publications by year](image)

3.2 Trends across industry focus

We next investigated how research in empirical Operations Management has evolved in terms of industry focus. Figure 2 compares the three categories with the greatest number of papers in our database: manufacturing, health care, and retail. Although the total number of papers over
the study period was similar across each of these industries, we saw different patterns in the evolution of each over time. Not surprisingly, given the roots of the Operations Management field, manufacturing has shown consistent publications over time (with some year-over-year variability). The empirical analysis of retail operations and health care operations, in contrast, had very little activity some 18 years ago, but has grown substantially since. Although these three industries have the greatest number of papers, they still make up less than half of the population, highlighting the diversity of industries studied within empirical operations.

![Figure 2: Papers by industry in empirical operations management over time](image)

3.3 Methodology
Third, we wanted to compare the methodologies used by the papers, both across methodologies as well as across time. Towards this goal, we conducted a text analysis of the articles. As is typically done with text analysis, we first constructed a dictionary that groups alternative words into a set of topics. Capturing the most frequently used empirical methods, we defined those topics as summarized by Table 1.

<table>
<thead>
<tr>
<th>Methodology group (topic)</th>
<th>Associated key words (specific methods)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrumental Variables (INSTRVAR)</td>
<td>Instrument, instrumental variables, two stage least squares</td>
</tr>
<tr>
<td>Diff-in-Diff / Matching (DIFMATCH)</td>
<td>Control group + treatment, difference in difference, matching methods, propensity score</td>
</tr>
<tr>
<td>Field Experiments (FIELDEXP)</td>
<td>Randomized, field experiment, intervention</td>
</tr>
<tr>
<td>Panel Data Analysis (PANEL)</td>
<td>Fixed effects, longitudinal, random effects, panel data</td>
</tr>
<tr>
<td>Regression Discontinuity Design (RDD)</td>
<td>Discontinuity design, Regression Discontinuity Design, RDD, regression discontinuity</td>
</tr>
<tr>
<td>Structural Estimation (STRUCT)</td>
<td>Structural estimation, structural model, primitives, generalized method of moments, GMM</td>
</tr>
<tr>
<td>Endogeneity (ENDOG)</td>
<td>Endogeneity, Endogenous Identification</td>
</tr>
</tbody>
</table>

**Table 1**: Empirical methodologies and associated key words

Although the last topic is not a method per se, it provides information to distinguish between a descriptive study and one that is aiming at establishing causality. Difference-in-difference and matching methods were included in the same group because they tend to use similar terms as they both match treatment and control groups. Digital files of each article in the database were downloaded and parsed to find the words in the dictionary. The parsing included the title, abstract, and main body of the paper, but it excluded the reference section. The number of occurrences of each word was counted for each article, and each paper was evaluated on whether
it matched any of the topics defined above. If a paper used words from the dictionary across multiple topics, it was assigned to each of those multiple topics.

For each year, we then counted the fraction of papers in the database that mentioned a particular topic. Figure 3 shows the fraction of papers in our database that were matched to a particular topic. Very few papers in our sample use regression discontinuity (RDD) (less than 10%) and therefore were excluded from the figure.

Figure 3: Fraction of papers in the database that mention words associated with different methodologies, for the years 2005-2016

Figure 3 shows that panel data and difference-in-difference/matching have been the most frequent methodologies mentioned in the database and had a relatively stable occurrence frequency over time. In contrast, instrumental variables have moved from below 20% to about
40%, indicating a strong increase in use. This is consistent with the more frequent occurrence of dictionary words related to endogeneity, which increased from 20% to more than 60%. Although it is not clear that all of these papers address issues related to endogeneity and causal inference, it is informative to see that causal analysis is receiving more attention. This is in line with previous findings (Ho et al. 2017) that about 75% of the empirical papers in Operations Management that use observational data were on causal inference.

3.4 Citation Analysis

After analyzing the frequency of publications, we now turn to the question of whether or not empirical papers accumulate more citations than non-empirical papers. To answer this question, we took a detailed look at citation patterns in *M&SOM*. We focus on one journal, *M&SOM*, for two reasons. First, *Management Science* citations are, in our opinion, more heavily influenced by the decentralized department structure and the associated non-homogenous research communities while *M&SOM* is a more centrally managed journal with a research community consisting of Operations Management scholars. Second, this article is written for publication in *M&SOM* and so it is a natural choice to focus on this journal.

Specifically, we ran a Poisson regression model with the dependent variable being the number of Google Scholar Citations, the independent variables being a dummy variable for the year of publication (to control for a fixed time effect), and another dummy variable indicating whether or not the paper was empirical. We ran the analysis for the time period of 2006 to 2016 as there were very few empirical *M&SOM* publications prior to this (see above). The results of this regression are illustrated in Figure 4. Empirical papers have about 20% more citations than non-empirical papers. To confirm this analysis, we also used a negative binomial regression model that created similar results, though it had slightly higher standard errors.
Figure 4: Citation analysis based on a Poisson regression (M&SOM only)

4. Empirical Research in Operations Management: Three Dimensions

Based on the articles in our database, we propose three dimensions on which empirical Operations Management papers can be compared: their main objective, their data sources, and their identification strategy. We believe that these three dimensions are helpful to illustrate the diversity of research that has been done so far. Moreover, these dimensions also should be helpful for researchers as they create new empirical studies. One paper might use multiple data
sources and pursue multiple aims, hence these three dimensions are not mutually exclusive in the classification of an individual paper. Table 2 summarizes the three dimensions.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Characteristics and Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main objective of paper</td>
<td>• Descriptive</td>
</tr>
<tr>
<td></td>
<td>• Test hypothesis based on OM theory</td>
</tr>
<tr>
<td></td>
<td>○ Test predictions of a normative model</td>
</tr>
<tr>
<td></td>
<td>○ Disentangle the mechanisms behind an effect</td>
</tr>
<tr>
<td></td>
<td>• Provide inputs and validate assumptions of a decision model</td>
</tr>
<tr>
<td>Data sources</td>
<td>• Structured vs. non-structured</td>
</tr>
<tr>
<td></td>
<td>• Proprietary vs. public</td>
</tr>
<tr>
<td>Identification strategy &amp; estimation method</td>
<td>Observational/experimental</td>
</tr>
<tr>
<td></td>
<td>Estimation methods</td>
</tr>
<tr>
<td></td>
<td>• Panel data</td>
</tr>
<tr>
<td></td>
<td>• Diff-in-diff/event studies</td>
</tr>
<tr>
<td></td>
<td>• Matching</td>
</tr>
<tr>
<td></td>
<td>• Instrumental variables/natural experiments</td>
</tr>
<tr>
<td></td>
<td>• Regression discontinuity</td>
</tr>
<tr>
<td></td>
<td>• Field experiments</td>
</tr>
<tr>
<td></td>
<td>Reduced form/structural estimation</td>
</tr>
</tbody>
</table>

**Table 2:** Three dimensions on which empirical research can be compared

### 4.1 Main Objective

Empirical research in Operations Management is carried out in the pursuit of different aims. Though it may well be that any given paper pursues multiple aims, we find it helpful to distinguish among the following:

1. **Descriptive study:** A descriptive paper uses empirical data and describes an interesting phenomenon relevant to Operations Management. An important contribution of this work is to motivate new lines of research (both empirical and analytical) and to explain the
reported phenomenon. The work of Nicole DeHoratius and Ananth Raman (2008) provides an example of a descriptive study. Prior to this study, research assumed that the inventory records within companies were accurate. By revealing that this was not the case, the authors inspired a significant body of new research that showed how decision models needed to change in the face of uncertain inventory data. The same paper also developed and tested a theory of record inaccuracy, which illustrates our previous observation that a paper may pursue multiple aims.

2. **Testing OM theory**: Analytical research in Operations Management focuses on developing normative models to make predictions about an outcome of interest and to provide managerial insights based on these predictions. Empirical studies can be used to test if the predictions of an analytical model are supported by empirical evidence. For example, the work by Cachon, Randall and Schmidt (2007) investigated the bullwhip effect, a phenomenon that has been predicted by numerous analytical supply chain management articles. The empirical analysis of the bullwhip effect was further refined by Bray and Mendelson (2012).

3. **Provide inputs and validate assumptions of a decision model**: Operations Management has a long tradition of combining analytical models and optimization methods from operations research and Economics to help managers improve the efficiency of their operations. Over the recent years, with the rise of supply chain management and service management, models incorporate more sophisticated decision processes of customers, decentralized business units, employees, and competition. Applying these models in practice requires an empirical validation of some of the assumptions required by these models as well as an estimation of some of its parameters.
and primitives. For example, the estimation results derived by Chong, Ho, and Tang (2001) enable a more accurate estimation of lost sales due to changes in assortment and enhance the measurement of product similarities across different brands, which leads to better assortment planning decisions.

4.2 Data Sources

Next, consider the type of data sources that provide the foundation for the empirical studies. As discussed earlier, we focus on company created and collected data, rather than data constructed by researchers (e.g., surveys). We can categorize the empirical papers in our database along two dimensions, structured vs. non-structured data and public vs. proprietary sources. This is shown in Table 3.

Structured vs. non-structured data

At the point of econometric analysis, the researcher always will work with structured data. Sometimes, this structured data is directly available from records. At other times, it has to be generated from non-structured data. Examples in which the structured data is available include sales transactions, orders, shipments, inventory, payroll, employee performance, and many other types. Public structured data has been useful to conduct cross-industry studies, industry-specific research, and studies that link operational with financial performance. Financial data from public firms (Gaur, Fisher and Raman 2005) and retail point-of-sales transaction data (Perdikaki, Kesavan and Swaminathan 2012) are other examples of structured data used in OM empirical studies. Textual data from news articles (Girotra, Terwiesch and Ulrich 2007), an analysis of customer reviews (Cachon and Olivares 2010), and video analysis (Lu et al. 2013) are examples in which the data is initially not ready to be included into an econometric analysis. In such cases, the researcher needs to translate this data into structured data.
Proprietary vs public source

Public data may be available for free or via a subscription service. The U.S. Census Bureau (http://www.census.gov) has many free and publicly available datasets that have been used in empirical studies in operations (e.g., Rajagopalan and Malhotra 2001). Other data can be obtained via subscription service for academic purposes, such as the Ward’s auto database (e.g., Moreno and Terwiesch 2015; Cachon and Olivares 2010) and the IRI retail dataset (e.g., Bronnenberg, Kruger and Mela 2008; Jagabathula and Vulcano 2017). Some data that were originally collected for proprietary use may become publicly available (e.g., hotel bookings published by Bodea, Ferguson and Garrow 2008). Some publicly available data may require significant processing to be suitable for empirical analysis, such as data scraped from the web (e.g., Olivares and Cachon 2009). The increasing volume of data generated by social networks and online market platforms provides an interesting opportunity to collect publicly available data for future OM research. Many of the articles in our database, especially those using patient-level data in the health care setting (Chan, Farias and Escobar 2017) or calls in a call center (Shen and Huang 2008, Garnett, Mandelbaum and Reiman 2002), are based on a collaboration between researchers and one organization that shares objective transaction data. We label such data sources as proprietary.

<table>
<thead>
<tr>
<th>Proprietary</th>
<th>Public</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structured</strong></td>
<td><strong>Public</strong></td>
</tr>
<tr>
<td>Point-of-sales data (Lee et al. 2015)</td>
<td>U.S. Census (Cachon et al. 2007)</td>
</tr>
<tr>
<td>Inventory records (DeHoratius and Raman 2008)</td>
<td>Store locations (Olivares and Cachon 2009)</td>
</tr>
<tr>
<td>Patient medical records (Chan, Farias and Escobar (2017)</td>
<td>Automobile production data (Wards) (Moreno and Terwiesch 2015)</td>
</tr>
<tr>
<td>Staffing (Kesavan, Staats and Gilland 2014)</td>
<td>Financial reports (Gaur et al. 2005)</td>
</tr>
<tr>
<td>Calls in contact center (Aksin et al. 2013, Garnett, Mandelbaum and Reiman 2002)</td>
<td>Stock prices (Chen, Frank and Wu 2007)</td>
</tr>
<tr>
<td>Customer panel data from market research companies (Jagabathula and Vulcano 2017)</td>
<td></td>
</tr>
<tr>
<td>Non-Structured</td>
<td>Radio frequency identification data (Staats et al. 2017)</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Video analysis (Lu et. al. 2013)</td>
</tr>
</tbody>
</table>

Table 3: Types of data sources

4.3 Identification strategy and estimation methods

The third dimension on which we find it helpful to compare empirical papers relates to the methodology used for estimation. The types of estimation methods used depend on the goal of the study. In addition to the method chosen, the researcher may take a passive or an active role in the data generation process. A passive role is called an observational design, and in this case the researcher has no intervention in the process that generates the data. Alternatively, in an experimental design, the researcher manipulates the data-generating process of some of the factors under study with the purpose of producing exogenous variation needed to identify a causal effect. Note that both designs can be used with any of the data sources described earlier.

A descriptive study typically uses an observational design. This often introduces several challenges when causal inference is sought. The association between two variables of interest may not be entirely driven by a causal effect as there may be other factors not observed in the data – denominated the “omitted variables” – that simultaneously affect the two variables of interest in a systematic direction. Many empirical studies in Operations Management seek to estimate the impact of certain managerial decisions on operational performance. However, it often occurs that the managerial decisions are influenced by other factors that are difficult
(sometimes impossible) to obtain for the purpose of a research study, creating an omitted variable bias.

To avoid such bias, a fundamental aspect in the research design of studies of causal inference is the identification strategy: a clear definition of the sources of variation in the data that can be used to estimate the causal effect of interest. We refer readers to Ho et al. (2017) for a full description of causal methods in empirical research.

Based on our analysis of the papers in our database, authors used the following estimation methods:

1. **Panel data:** In cross-sectional studies, it may not be possible to incorporate all the relevant factors that describe the units of analysis. In this context, panel data provides an identification strategy that relies on the time-series variation within each unit while controlling for time-specific factors that affect all units similarly (e.g., Jagabathula and Vulcano 2017).

2. **Difference-in-differences:** As with panel data, the effects are identified by estimating the change in a unit of analysis before and after exposure to a treatment of interest, relative to a valid control group that did not receive the treatment. When using a difference-in-differences approach, the key assumption is that the treatment group and the control group followed a “parallel trend” prior to the treatment (Bertrand, Duflo and Mullainathan 2004; Song et al. 2016; Staats et al. 2017).

3. **Event studies:** This identification approach is similar to difference-in-differences and is typically applied in the context of financial data. The control group is provided by an empirical model that predicts the expected value of a financial instrument (e.g., stock
returns), which is used as a benchmark to estimate the “abnormal” value attributed to the treatment group (Hendricks and Singhal 2001).

4. **Matching**: This approach seeks to replicate a balanced experimental design using observational data by finding close matches between pairs or groups of units and separating out the ones that received a specified treatment from those that did not, thus defining the control groups. A common approach to create matches between treatment and control groups is to match based on observable covariates so that the likelihood of receiving treatment is similar between the two groups. One common method to accomplish this is known as propensity score matching, although there are several extensions and related approaches in the literature (Levine and Toffel 2010).

5. **Instrumental variables**: Instruments are additional variables that produce an exogenous variation in the endogenous variable and do not directly affect the outcome (other than through their effect on the endogenous variable). Intuitively, this identification strategy uses only a part of the variation in the endogenous variable that is considered to be exogenous. The estimation can be carried out using two-stage least squares when the model is linear, using simultaneous equation models when dealing with multiple endogenous variables, or using a control function or maximum likelihood approach when some of the endogenous variables require a nonlinear model such as logit, probit or multinomial (Kesavan et al. 2014; Tan and Netessine 2014; Kim et al. 2015).

6. **Natural experiments**: This requires finding a random source of variation in the endogenous variable that would be equivalent to conducting a field experiment. A natural experiment can also be considered an instrumental variable because the random nature of the variation generates an exogenous shock in the endogenous variable that facilitates
identification. Examples include Parker, Ramdas and Savva (2016) and Staats, KC and Gino (2016).

7. **Discontinuities in the endogenous variable:** In many managerial settings, there are specific and known rules that describe how the value on an endogenous variable is determined. These rules can sometimes force two similar units to receive different levels of the endogenous variable, which can be used as an exogenous source of variation to measure the causal effect of the endogenous variable (Bennett et al. 2013).

8. **Field experiments:** Commonly used in the field of Medicine, the idea of an experiment is for the researcher to actively intervene in the study setting, measuring the effect of the intervention on some outcome variable. Field experiments have recently gained popularity in Economics, and recent OM applications include the work by Zhang, Allon and van Mieghem (2016) and Buell, Kim and Tsay (2016). Though lab experiments are not part of this review, they can be used to complement such field experiments. For example, Buell et al. (2016) conducted a set of lab studies to conceptually replicate and extend their finding when examining operational transparency in food preparation.

Examples of articles that use each of these identification strategies are described in Table 4, alongside seminal methodological references.

The identification strategies described above can be implemented using two different modeling approaches:

- **A reduced form approach,** where the dependencies among the different factors that are deemed relevant are specified through a simplified parametric econometric model that describes the association or causal effects among these variables. A reduced form estimation provides inferences on the parameters that describe this econometric model.
Such a model can be used to make predictions by estimating the variation in the dependent variable triggered by changes in some of the independent variables, which can be computed after the parameters of the model have been estimated. Reduced form models can be linear (a regression), or nonlinear (logit, probit, multinomial logit, Poisson regression, among others).

- **A structural approach**, which specifies directly the decision process that generates the data used in the estimation (e.g., Olivares, Terwiesch & Cassorla 2008). The specification of the decision model includes primitives that determine the endogenous outcomes of interest. The objective of a structural estimation is to “reverse engineer” the model to make inferences about the primitives that would generate the endogenous outcome observed in the data. Predictions from a structural model consist of a counterfactual (or what-if/simulation) analysis that predicts how changes in the primitives would affect the outcomes of interest. The estimation of structural models frequently involves ad-hoc nonlinear models that can be estimated via maximum likelihood or generalized method of moments.

Based on the papers in the database, we observed that there is no preferred method to study a research question. Table 5 provides several examples where the same research question was studied by articles using reduced form models and others using structural estimation. Like in Table, the last column shows seminal methodological references. Structural and reduced form should be viewed as complementary approaches when conducting empirical research.

### 5. Conclusion

The publication rate of empirical research in Operations Management has increased in the three journals, moving up from less than five papers per year in the first years to well over 20 in the
most recent years. This trend comes along with an increasing focus on health care and new methodologies. In particular, we see an increase in the use of instrumental variables and a much larger concern about endogeneity. Finally, empirical OM papers are strongly cited. Controlling for publication year, they get more citations than non-empirical papers.

We are empirical researchers, not forecasters. So, we will not make a prediction about the trend we described beyond our study period (ending in December 2016). However, we believe that two forces will help empirical research enjoy a healthy future in Operations Management. First, based on our database, there exists a large number of scholars who have contributed three or more empirical papers to our database. Thus, we now have reached a critical mass that is important for a research area, as it allows for the recruitment of reviewers, editors, and letter writers for tenure processes. We have moved beyond the tipping point. Second, new technologies such as the internet of things (IOT) are not only transforming the way we work, but also how we research. The associated sensors and video cameras also create databases of billions of transactions, including worker and customer movements, GPS tracking of ships, planes, and trucks, container routing, patient health data, and much more. For example, Lu et. al. (2013), used video analysis to study how customers behave in a queue, leveraging second-by-second data of customer behavior. Staats et al. (2017) used radio frequency sensor data that continuously tracked when health care providers enter a room and whether they complied with hand hygiene processes. Batt et al (2017) used tracking sensors to study how nurses move through a hospital, emitting a location signal every 6 seconds. As the academic discipline that studies how people work (after all, the word operations has its roots in the Latin word opus, which translates as work), our field has unprecedented opportunities in this ocean of data (Terwiesch, forthcoming). By providing an easy-to-use interface to the list of empirical papers that so far have been
published in our field, we hope to provide guidance through the literature so that future authors can benefit from the work that already has been done.

6. References


Table 4: Identification Strategies
The following table provides a list of the identification strategies we found in our database. The table shows the Operations Management application (column 2), the key variables in the study (column 3), and a reference example from the Economics and Statistics literature.

<table>
<thead>
<tr>
<th>Identification Method</th>
<th>Examples of papers</th>
<th>Key variables of interest</th>
<th>References from Econ/Stats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>Reference</td>
<td>Effect Description</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>Reference</td>
<td>Description</td>
<td>Reference</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Field Experiments</td>
<td><strong>Zhang, Allon and Van Mieghem (2016).</strong> &quot;Does social interaction improve service quality: field evidence from massive open online education.&quot; <em>M&amp;SOM (forthcoming).</em></td>
<td>Effect of nudging students to increase (online) social engagement on their grades and completion rates in a MOOC.</td>
<td><strong>Bandiera, Barankay and Rasul (2005)</strong></td>
</tr>
</tbody>
</table>
**Table 5:**
The following table shows examples of different research questions that have been studied via structural estimation and reduced form methods to illustrate how these two estimation approaches complement each other for conducting empirical research. The first column describes the research question. The second column shows two examples, corresponding to a structural estimation and a reduced form estimation (described in column 3). Column 4 describes the key parameters to be estimated, and the last column shows an example of the methodology from the Economics and Statistics literatures.

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Examples from Database</th>
<th>Methodology</th>
<th>Key Parameters</th>
<th>Method References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Variability of production relative to variability in sales, for different industries.</td>
<td>Blinder (1986)</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Petrin and Train (2010)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>