Sequencing jobs in an engineer-to-order engineering environment

Douglas H. Grabenstetter and John M. Usher*

Department of Industrial and Systems Engineering, Mississippi State University, MS 39762, USA

(Received 3 December 2014; accepted 25 March 2015)

Engineer-to-order (ETO) firms produce complex – one of a kind – products and desire shorter lead times as a key component to cost competitiveness. In ETO firms, the engineering process is the largest controllable consumer of lead time. Given that lead time is a function of completion rate and scheduling policy, one critical process is to accurately sequence jobs in front of the engineering function. However, unlike other manufacturing models, such as make-to-stock or make-to-order models, the design for an ETO product is not realized until after the engineering process has been completed. Hence, the only information available does not include data normally required by most sequencing algorithms. Therefore, the problem becomes the determination of an accurate schedule within a complex transactional process for jobs which have not even been designed yet. This paper investigates this topic in the context of the engineering process within the ETO model. Based on research conducted in conjunction with multiple firms, common factors are identified which drive complexity, and a new framework and algorithm are presented for using these factors to sequence jobs. Using discrete event simulation, the performance of this new algorithm is found to be a significant improvement over current industry and published methods.

Keywords: engineer-to-order; sequencing; design

1. Introduction

The engineer-to-order (ETO) model is used by a significant number of manufacturers across multiple sectors. Indeed, ETO firms comprise approximately one-fourth of all North American manufacturing and are growing at a rate of 20% (Cutler, 2005). One reason for this level of growth is the increased demand for customized products which, in turn, has caused markets to become more and more fragmented. Use of the ETO model provides firms with increased agility and flexibility thereby, allowing them to be ideally positioned to respond to rapid market shifts through production of one of kind products. Despite these advantages, the very nature of the ETO environment drives complexity. For instance, ETO products are manufactured and assembled in low quantities to individual customer specifications. In addition, the goods manufactured using ETO models tend to be complex in nature, resulting in deep and complicated multilevel product structures and routings which often contain job-specific and engineered-to-order components (Hicks & Braiden, 2000; Hicks, Song, & Earl, 2007). In the ETO environment, the engineering process is the largest controllable consumer of lead time taking up to one half of the total lead time (Little, Rollins, Peck, & Porter, 2000; Pandit & Zhu, 2009). Since ETO firms produce products which are often used in large projects, it

*Corresponding author. Email: usher@ise.msstate.edu

© 2015 The Author(s). Published by Taylor & Francis. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
is not unusual for their customers to impose large cost penalties for lateness. Further, given that lead time is a function of completion rate and scheduling policy, one critical process is to accurately sequence jobs in front of the engineering function and therefore is an important topic for these companies.

Yet, despite the aforementioned need, there has not been much sequencing research focused on the ETO engineering environment. Therefore, this paper investigates sequencing approaches used in the literature in order to generate good (not necessarily optimal) job sequences. A new complexity sequencing approach is introduced that has been tested using real-world data from a typical ETO industrial firm and found to meet the needs of ETO firms in that it is quick, easy to use, accessible, and accurate. Indeed, the approach has been found to significantly outperform traditional sequencing methods used in the ETO environment.

2. Background
When discussing ETO firms, it is important to consider a couple of important constraints. The first of these is that the engineering process is a bottleneck resource per Gosling and Naim (2009), Cutler (2005, 2009), and Pandit and Zhu (2009). Indeed, it is a core process and is never outsourced (Hicks, McGovern, & Earl, 2000). Further, competition from other firms has caused some firms to reduce their costs by 50% (Hicks, McGovern, & Earl, 2001). Since resources for the engineering processes are costly (Qi & Tu, 1998), these staffs in turn have been scaled down (Hicks et al., 2001). As a result, engineering staffs are now often resource constrained. Secondly, most ETO firms use enterprise resource planning (ERP) systems, per Stevenson, Hendry, and Kingsman (2005), which are substantial investments per Worthen (2002). Hence, trying to convince these firms to throw out their ERP systems likely will be a hard sell! Unfortunately, despite the need for research on the engineering process, the vast majority of research performed for ETO firms by such authors as Hicks and Pongcharoen (2006, 2009), Hicks and Braiden (2000), Hicks et al. (2000, 2001, 2007), Pongcharoen, Hicks, and Braiden (2004), Pongcharoen, Hicks, Braiden, and Stewardson (2002), and Song, Hicks, and Earl (2001, 2002, 2006) recommends complex global shop floor scheduling approaches which seek to replace ERP systems. Coates, Ritchey, Duffy, Hills, and Whittfield (2000) provide the only discussion found which focuses on the ETO engineering process. They present an agent-based design decision evaluation tool for large make-to-order products. This tool chooses the design concept which takes the least amount of projected time with minimum resource use. The purpose of the agents is to perform this task while dealing with a diverse set of design tools using data from heterogeneous distributed computing systems. Seven agents are identified with each of the agents fulfilling a specific role performing several activities. A resource manager agent optimizes the utilization of the resources. If the utilization of a particular resource falls below a given target, or if an alternative resource’s efficiency level is better than the current resource, the resource manager triggers a scheduling agent to create a task matrix which contains task dependencies and task durations. The scheduling agent determines in which order certain design candidates are evaluated. The scheduling problem is defined by the authors as a job shop problem (as opposed to a single-machine setting) which has been proven to be NP-Hard (Brucker, DrexI, Moehring, Neumann, & Pesch, 1999). Therefore, the authors make use of a multicriteria genetic algorithm (GA) which “optimizes” multiple objective functions simultaneously. The GA produces a pareto-optimal set of schedules using the task matrix and chooses the schedule which
maximizes the utilization of the resources, minimizes the number of resources, and
minimizes the time to complete design concept evaluations. The main problem with the
GA approach lies with its complexity as driven by the scheduling problem definition.
Since the scheduling problem is defined as a job shop problem, computation time will
likely be an issue due to problem logic complexity.

In contrast to what has been proposed by others, this paper offers a more robust
approach which uses the output of ERP and focuses on scheduling just the engineering
process not each process of the entire system. The new methodology addresses the
inherent scheduling complexity of the ETO environment through the use of dispatching
rules (DR). Compared to the aforementioned methods, DR are low in computational
complexity, having only to compute which job to process next as opposed to having to
determine the entire sequence of all the jobs in the queue. While use of DR will not
guarantee optimal solutions, good solutions are often achieved for either static or
dynamic, single or multiple machine environments. As well, DR are very easy to imple-
ment. Therefore, it is no surprise to find widespread use of DR by ETO firms.

There are three basic types of sequencing approaches found in the literature which
use due dates as an input: allowance-based, slack-based, and ratio-based. In his study,
Baker (1984) investigates approaches from each of these three categories and finds that
the performance of slack-based approaches is equivalent to the allowance-based
approaches. Referring to Table 1, Baker and Trietsch (2009) find that shortest processing
time (SPT) performs quite robustly for the proportion tardy (PT) and mean tardy (MT)
objective functions. Additionally, Baker reports that modified due date (MDD) performs
well for the single-machine dynamic maximum lateness (T) problem and that early due

| Metric | Best rule(s) | Other rules compared | Authors |
|--------|--------------|----------------------|---------|
| PT     | S/OPN        | EDD/MST              | Conway (1965a), Conway, Maxwell, and Miller (1967) |
|        | SPT          | MST, EDD, S/OPN      | Elvers (1973) |
|        | SPT          | EDD, ODD, MST, SCR,  | Muhlemann, Lockett, and Farn (1982) |
|        | SCR, A/OPN   | S/OPN                | Miyazaki (1981) |
|        | SPT          | MST, OST, EDD, SCR,  | Kanet and Hayya (1982) |
|        | SPT, S/OPN,  | EDD                  | Elvers and Taube (1983) |
|        | MST          | S/OPN, SPT           | Gere (1966) |
|        | SPT          | S/OPN, EDD           | Conway (1965a, 1965b), Conway and Maxwell (1962), Carroll (1965) |
|        | S/OPN        | SPT                  | Weeks (1979) |
|        | SPT, ODD     | MST, OST, SCR, ODD,  | Kanet and Hayya (1982) |
|        | A/OPN        | SCR, S/OPN           | Miyazaki (1981) |
|        | SPT, EDD     | MST, SCR             | Baker and Bertrand (1981) |
|        | MST, SPT,    | OST, S/OPN, SCR, ODD| Muhlemann et al. (1982) |
|        | EDD          |                      |         |
| CMT    | S/OPN        | SPT, EDD             | Conway (1965a, 1965b), Conway and Maxwell (1962) |
|        | SCR          | SPT, EDD, ODD, MST,  | Muhlemann, Lockett, and Farn (1982) |
|        |              | S/OPN, OST           |         |
|        | OCR          | SPT, EDD, ODD, MST,  | Kanet and Hayya (1982) |
|        |              | S/OPN, ODD           |         |
date (EDD) maximizes the minimum service level in the stochastic single machine environment. Further, EDD has been found by this author to be the method of choice among ETO firms.

3. Sequencing research in the ETO engineering process

As previously mentioned, the available research in the literature is not directly applicable to the ETO environment. In addition, very little information exists about the ETO engineering process and how engineering due dates are calculated. Therefore, subject-matter experts (SMEs) from five ETO firms of various industrials (listed in Table 2) were consulted remotely via telephone, or, where feasible, face-to-face. With each of the ETO firms, research was conducted regarding the front-end processes and flows they employed including due-date setting methods.

The engineering process is distinct from traditional shop floor processes as its inputs and outputs are purely intellectual and not physical in nature. Hence, in order to gain an idea of how sequencing is performed in ETO firms, SMEs at each site were interviewed about their individual process flows including the inputs and outputs, as well as the details of the tasks at each step. Not surprisingly, the challenges of the ETO environment have forced each of the firms to adopt similar process steps. The data gathered are summarized in the form of a generic IDEF0 diagram (see Figure 1).

In step A1, an initial job due date is set. The lead times are set intentionally wide enough to cover the amount of uncertainty present. Once the order is obtained, the firm must provide a committed job due date in step A2. The job information is forwarded on to the engineering step (A3). Research found that several constraints exist in regard to sequencing for the ETO design environment. First, the ETO environment is complex, and each job is unique and often complicated. In addition, there is a limited amount of information available. Referring to Figure 2, the first step in the engineering process (A3.1) is a review and sequencing process where the job information is scanned for long lead time non-inventoried components. In addition, the job data are assessed to ensure that vital data required by the engineering discipline in play (e.g. wiring diagrams, ladder diagrams, and other details) do not contain clean-order defects. These clean-order defects encompass missing information not provided by the sales force, incorrect information in that it is wrong, infeasible or non-sensical, and conflicting information between two or more pieces of data. For instance, a ladder diagram may call for logic which does not translate properly into the expected components.

Clean-order defects are a major source of variability in the ETO environment and cause rework loops in the engineering process (steps A3.2 and A3.4). These issues in

| Firm | Market  | Product                     | Annual sales | Employees |
|------|---------|-----------------------------|--------------|-----------|
| A    | Industrial | Motor control center       | $60MM        | 256       |
| B    | Construction | Switchboards               | $25MM        | 20        |
| C    | Industrial | Process instrumentation    | $38MM        | 80        |
| D    | Construction | Switchgear                | $81MM        | 317       |
| E    | Industrial | Busway                     | $63MM        | 232       |
ETO firms are problematic. Indeed, a four-year study conducted by Firm A found that 37% of orders contain some sort of clean-order error. Normally, the cycle time to solve clean-order defects may take up to 2 weeks. However, if the customer is found to be the source of the defect, the ETO sales force is contacted and negotiations then occur, which can drive up the resolution cycle time. According to Firm A, such a defect occurrence can take 60 days or more to resolve. Additionally, if the customer is at fault or if
a change is initiated by the customer, the due date could be moved out. If the clean-order negotiations become intractable, it is also possible that the order will be put on hold or even canceled.

Once the clean-order review has been completed, the job is then sequenced in step A3.1. As previously mentioned, research with the five partner firms has revealed that the EDD method of sequencing is used as a general practice. The main reason for using EDD is its simplicity. Each of the firms stated that they prefer simple, easy to use methods that can be run using a spreadsheet. This follows the findings of Stoop and Wiers (1996) that simple and robust approaches are the most broadly used in practice. Otherwise, the firms will likely not use the methodology (O’Neil, 1991). The scheduling process is typically conducted on a weekly basis for several reasons. First, Pandit and Zhu (2009) and Little et al. (2000) point out that the ETO engineering process can consume up to one-half of the production lead time. Hence, the engineering task is considered to be a long duration process. Second, the engineering step is a complex, intellectual process and interruptions to the process can lead to an inflated error rate. Thus, when an engineer begins work on a job, if at all possible, it is desired that they be allowed to complete the engineering task without interruption (non-preemption). Otherwise, constant rescheduling is known to cause schedule nervousness.

Upon the job being sequenced and assigned to an engineer, the next step in the process (A3.3) is the actual engineering task where the job inputs are transformed into a tangible product design. Given the high complexity of ETO products, the engineers themselves typically possess a high degree of experience and are difficult to replace. In addition, due to the high product complexity present, each of the partner ETO firms makes use of a proprietary computer-aided job engineering configuration software system (CAJECSS). When engineering the job, the engineer chooses predefined job characteristics found in the system. The CAJECSS then uses logic tables to generate the appropriate engineering output (e.g. wiring diagrams and assembly drawings). In addition, it also generates a top-level bill of material which is used in the subsequent production planning step to order materials. Only when the engineering process is finished, can the remainder of lower level discrete components be procured. Once the components are on-hand, the manufacturing process itself can begin.

These findings lead to an overriding question: How can scheduling capabilities which co-exist with existing ERP systems be developed for the highly complex ETO engineering environment while, at the same time, remaining realistic, robust, acceptable, and easily implemented? One might conjecture that the answer simply lies in applying what has been researched already in the job shop literature. However, much of this research presupposes that the process time of each job is known or can be easily estimated. However, because each job is unique, in the ETO environment, this simply is not the case. Restated, the literature survey and ETO firm research have shown that what is needed is a sequencing policy which:

- robust and relatively simple to implement and use (spreadsheet driven),
- accurate (an improvement over EDD), and
- complementary to an ERP system (Recall that many firms have invested heavily in their ERP systems and are in no hurry to replace them).
4. Regression-driven complexity-based flow time prediction

Based upon the previous discussions, the engineering process can be modeled as a dynamic job shop with the following assumptions:

- arrivals are dynamic,
- machines process one job at a time,
- setup times are independent of job sequence,
- machines are continuously available, and
- neither idle time nor preemption is allowed.

SPT is well known to be a robust sequencing algorithm. However, given the nature of the ETO engineering environment, nothing is known ahead of time in regards to job processing time. To overcome this problem, Grabenstetter and Usher (2013) identified seven engineering factors that can be used in this context to predict job processing times using a regression-based approach. These factors are:

- $x_1$ – number of functional requirements
- $x_2$ – number of basic components
- $x_3$ – number of design interdependencies
- $x_4$ – number of technologies
- $x_5$ – number of regulations and standards
- $x_6$ – number of subsystems
- $x_7$ – presence of a reference job

Although multiple regression has been used by others such as Salman and Dogru (2004), no such procedure has ever been employed which uses ETO-specific complexity factors (CFs). Once the regression procedure generates a predicted processing time for each job $j$, they can then be used as an input to SPT. The framework for this approach (referred to as RegSPT) is comprised of three parts that are complementary in nature. The first component employs multiple linear regression using a basic CF equation composed of the seven CFs.

$$
CF_j = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7
$$

where $CF_j > 0$, $x_1$, $x_2$, $x_3$, $x_4$, $x_5$, $x_6$, $x_7 \geq 0$ and initially $x_1$ through $x_6$ are defined as discrete integer variables, and initially $x_7$ is defined as a binary ordinal categorical variable ($1, 0$).

The second element of the framework is an estimation algorithm (EA) which is used to determine the values of the coefficients of the CF equation. This step is necessary for each different ETO firm given that the exact continuous distribution function of the dependent variable data and the exact probability distribution function of the dependent variable data will vary somewhat with each firm. As well, the EA will need to be rerun periodically because it is likely that business conditions will change over time and process improvements will occur. These changes may affect the efficacy of the input data used to build the actual regression equation currently in use. This means every $p$ periods, the EA will need to be run using up-to-date historical data for the last $p$ periods for both the response variable and the independent variables $x_1$, $x_2$, $x_3$, $x_4$, $x_5$, $x_6$, and $x_7$. Because not all ETO firms are exactly alike, the length of the update cycle, $p$, and the sample size, $z$, will need to be determined by each firm. The EA is defined as:
Step 1: Gather dependent variable data for the last $p$ periods.
Step 2: Gather independent variable data for the last $p$ periods.
Step 3: Generate updated slope coefficients ($\beta$): using Equation (1) as a starting point, generate a statistically valid regression model using the seven aforementioned factors.

The final component of the framework is the RegSPT algorithm (RSA) which uses the resulting CF equation as a part of the sequencing process to generate the complexity-based processing times for each job order. Operationally, when a job is received, the production planning step assembles data that are found in the specific job’s packet. The independent variable data ($x_i$’s) are then entered at the engineering step into the regression equation for the job at hand to determine the job unique complexity-driven processing time. Once processing times have been predicted for each job to be sequenced, the SPT sequencing algorithm is then utilized. RSA is defined as:

Let $n =$ the number of jobs.
Step 1: Initialization: Upon arrival of job $j$, Set $j = 1$.
Step 2: For Equation (1), determine $x_1, x_2, x_3, x_4, x_5, x_6,$ and $x_7$ using the job information provided by the sales department.
Step 3: Evaluate Equation (1) to generate the predicted process time.
Step 4: Let $j = j + 1$.
Step 5: If $j = n$, Go to step 6. Else, go to Step 1.
Step 6: Sequence job $j$ . . . $n$ in ascending order of predicted processing times. If ties result, schedule by earliest due date.
Step 7: Stop.

5. Example implementation of the RegSPT method

To illustrate the application of the framework for generating job sequences, actual data from Firm A was chosen because it closely matches the criterion listed in the literature for ETO firms (Hicks et al., 2007). When performing regression analysis, it is important to follow a strategy of solid statistical practice. Therefore, the strategy listed in Kutner, Nachtsheim, and Neter (2004) was followed.

Step 1 – The first step of the EA process is to gather data from the last $p$ periods. At this point, it should be stated that due to differences in software and operational practices, the execution of this task will vary with each firm. Some firms may make use of MS Excel® tables to store these data, while others may store them in their ERP system. At Firm A, response variable data were gathered which represent actual historical engineering labor flow time data in days. These data were resident in their ERP system and consisted of two years’ worth (1276 distinct jobs) of data which were exported to a spreadsheet program. Again, for our tests, in order to rule out the impact of sampling bias, the time period from which the data were gathered came from the most recent period of economic stability and represents steady state/normal business conditions. The response variable statistics were mined directly from the data-set. Outliers were identified, and, if determined by Firm A to be inappropriate, were removed.

Step 2 – The independent variable data to be gathered must represent actual historical engineering data used to engineer each of the corresponding jobs identified in Step 1. Once the raw data are obtained, the data need to be separated into the aforementioned seven factors. Again, due to differences in the format in which this data is stored and operational practices, the difficulty of this task will vary with each firm. In this example,
similar to some other ETO companies, Firm A makes use of a sophisticated proprietary CAJECSS. With this system, the engineer chooses predefined job characteristics found in the system. Then using logic tables, the CAJECSS generates appropriate wiring diagrams and assembly drawings. In addition, it also generates a top-level Bill of Material which is used in the subsequent production planning step to order materials. For two of the seven factors (number of subsystems and presence of a reference job), the data could easily be extracted directly. However, data categorization for the other five factors was not as simple. These data were discovered to be embedded within multiple levels of tables found within the CAJECSS. Further, given the multilevel nature of the tables, and that factor $x_3$ is built upon interdependence of design characteristics, the data for each factor could not be extracted using simple sorting techniques. Therefore, in order to accomplish separation of these data into their appropriate factors, a binary mapping table using MS Excel$^{\text{®}}$ was developed which took advantage of the logic statements found within the configuration system. Upon reviewing the raw data from the CAJECSS, the amount of raw data was found to be a staggering amount. As a result, the mapping table was then utilized to develop a computerized algorithm to separate the data into the five aforementioned factors. The algorithm was written in the Visual Basic$^{\text{®}}$ for Applications programming language. The run time was very quick, requiring less than 30 s. These tools are reusable and once developed, can be used when future updates of the system are required.

Once the data are gathered, the next step is to validate the data and gain an understanding of the data quality. Direct diagnostic plots of the response variable are normally not very important in regression analysis as the values of the observations on the response variable are functions of the levels of the predictor variables. However, in order to ensure that the study is of the highest quality, a cursory look at the data should be performed. The data for each dependent and independent variable were plotted using box plots and reviewed with the SMEs in order to identify and remove gross data errors and outliers.

**Step 3** – Next, the regression model is generated using Equation (1) and a standard residual analysis performed to ensure that the model meets the statistical criterion for multiple regressions. Since the distribution of residuals was initially found to be non-normal, a logarithmic transformation was performed on the response, which is referred to in our example as $\log \text{PT}$. Referring to Tables 3 and 4, the model is found to be appropriate. The model itself has a $p$-value of zero, which is well below the normal .050 alpha threshold. Further, all of the factors except for $x_2$ and $x_5$ similarly are well

| Predictor           | Coef  | SE Coef | $T$   | $p$   |
|---------------------|-------|---------|-------|-------|
| Constant            | .63492| .03167  | 20.05 | .000  |
| $x_1$ – # Func. req. | .00148| .00025  | 6.01  | .000  |
| $x_2$ – # Basic comp. | .04222| .01216  | 3.47  | .001  |
| $x_3$ – # Design inter. | .14271| .03200  | 4.46  | .000  |
| $x_4$ – # Technologies | -.04273| .01225 | -3.49 | .001  |
| $x_5$ – # Reg. and Std. | -.001185| .00067 | -1.77 | .076  |
| $x_6$ – # Subsystems | .26399| .03257  | 8.11  | .000  |
| $x_7$ – Ref. job? | -.58987| .03055 | -19.31| .000  |

$S = .472984$ 
$R^2 = 39.8\%$ 
$R^2_{adj} = 39.4\%$

Table 3. Model with factor significance.
below the normal .050 alpha threshold and although $x_2$ and $x_5$ are marginally above the threshold, investigations with the ETO firms have concluded that these factors are valid. Hence, the model can be used for predictive purposes.

Step 4 – Given that the CF equation with coefficients is available, the RSA can be run. In order to generate the processing times, a second historical data-set consisting of (74) jobs was gathered from Firm A and represents the input data that would be used by the RSA. Recall that the regression formulation is based on $p$ time periods before the current time period. Therefore, the data from the 1276 distinct jobs used in Step 1 are from a time period preceding this set of 74 jobs. This condition simulates actual operating conditions which would be experienced by practitioners. Equation (1) is embedded in an MS Excel® spreadsheet and is evaluated for each of the 74 jobs to compute the predicted processing times. Examples of the input data and predicted process time values are given in Table 5.

Table 4. ANOVA table.

| Source          | DF | SS   | MS   | $F$  | $p$  |
|-----------------|----|------|------|------|------|
| Regression      | 7  | 187.366 | 26.767 | 119.65 | .000 |
| Residual error  | 1268 | 283.669 | .224  |      |      |
| Total           | 1275 | 471.034 |      |      |      |

Table 5. Example of input data for RegSPT and output.

| Order No. | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | LOG predicted process time | Predicted process time |
|-----------|-------|-------|-------|-------|-------|-------|-------|-----------------------------|------------------------|
| 3001774379| 128   | 71    | 0     | 68    | 0     | 1     | 1     | .59                         | 3.90                   |
| 3001912763| 142   | 113   | 1     | 112   | 0     | 1     | 0     | 1.24                        | 17.31                  |
| 3002194681| 30    | 6     | 0     | 6     | 0     | 1     | 0     | .94                         | 8.72                   |
| 3002246903| 2     | 0     | 1     | 0     | 0     | 0     | 0     | .78                         | 6.04                   |

Table 6. Example of the RegSPT schedule.

| Order No. | Arrival time hour | Predicted process time (days) |
|-----------|-------------------|-------------------------------|
| 3001774379| 0                 | 3.90                          |
| 3002267611| 0                 | 8.27                          |
| 3002194681| 0                 | 8.72                          |
| 3002341332| 0                 | 12.36                         |
| 3001912763| 0                 | 17.31                         |
| 3002380038| 7                 | 2.35                          |
| 3002400367| 7                 | 4.32                          |
| 3002390220| 7                 | 8.03                          |
| 3002454575| 14                | 8.16                          |
| 3002452423| 14                | 8.50                          |
| 3002447003| 14                | 8.58                          |
| 3002450199| 14                | 11.18                         |
| 3002491099| 21                | 4.35                          |
| 3002472412| 21                | 6.44                          |
| 3002488807| 21                | 7.98                          |
Step 5 – Referring to Table 6, the next step is to group the jobs based on their arrival time and then sequence the jobs in each group in ascending order of their processing time (SPT). The run time to sort the 74 jobs is extremely fast and was completed in less than 5 s.

6. Methods

Given the previous section’s description of RegSPT as an alternative to current method of sequencing for the ETO engineering environment, it remains to be proven whether the new methodology is better than current approaches typically used by ETO firms and popular literature approaches. Therefore, an analysis using simulation employing data from one of the five aforementioned partner firms, Firm A, was performed to study the performance and behavior of sequencing methods in a real-world ETO engineering department. A discrete simulation model of the engineering process was built, verified, and validated using the approach of Harrell, Ghosh, and Bowden (2004). The output of the model yields queue lengths, utilization, job cycle times, and many other data useful for analysis of the long-term behavior of the engineering environment. In particular, the cycle times were used to determine job completion times. In order to provide a 95% percent confidence interval with a three-day half-length using the standard deviation of lateness (SDL) value based on an initial run of 30 replications, it was determined that 53 replications of the experiment would be required. Due dates for input into the model were calculated using the approach of Grabenstetter and Usher (2014). The process steps in the model follow those shown in Figure 2 and the processing times for those steps utilize industry values. The arrival patterns form the basis for testing the performance of RegSPT against popular sequencing methods.

Three popular sequencing approaches from the literature were chosen for the analysis: SPT, MDD, and EDD. Baker and Trietsch (2009) find that SPT performs quite robustly for a variety of objective functions (see Table 1) and, in particular, in the presence of imperfect information which is an advantage in the ETO environment. In addition, it is easy for practitioners to implement and understand. However, as can be observed, this rule requires foreknowledge of a job processing time. Hence, in order to simulate the operation of SPT in practice at an ETO firm, it is necessary to develop a method that ETO firms would most likely use to predict processing times using the limited data (material costs) that is available at the beginning of step A3. To accomplish this task, it is important to know that the firms will likely only implement a method which is simple and robust (O’Neil, 1991; Stoop & Wiers, 1996). Therefore, it was decided to use a simple material-based approach similar to what is currently utilized in industry. As such, jobs are grouped into three classes based upon perceived difficulty as predicted by the job material content (Grabenstetter & Usher, 2013). To accomplish this task, the 1276 jobs in the data-set were grouped by the three classes and an average process time by class determined. These average process times are then used to help drive SPT. In the literature, the release time \( r_j \) is another variable that normally appears as a part of sequencing formulations. However, in this case study, it is assumed universally that the release time equals zero. Therefore, this term drops out of the formulations.

The MDD method chooses the assigned job due date \( d_j \) or the soonest time that the job could be completed \( t + p_j \), whichever time is later.

\[
d_j' = \max\{d_j, t + p_j\}
\]
The same approach for processing time prediction used previously for SPT is employed here to get the values of $p_j$.

EDD is the method of choice among ETO firms and has been reported to perform robustly. It is simple, easy to understand and implement, and does not require foreknowledge of job processing times. In order to sequence the jobs via the EDD method, within each arrival period, the jobs are simply arranged in ascending order according to their due dates. If ties are encountered, the jobs are ordered arbitrarily.

In order to judge the performance of the sequencing methods, it is necessary to define the performance metrics used. This study employs the criteria used by other published ETO studies such as Hicks and Pongcharoen (2006, 2009), Hicks and Braiden (2000), Hicks et al. (2007), and Song et al. (2002, 2006). These criteria are:

- Mean lateness (ML): lateness is calculated as the job completion time minus the due date. As such it measures the deviation, early or tardy from the due date.
- Mean absolute lateness (MAL): defined as the absolute value of the job lateness.
- SDL: used to add a measure of variability.

7. Effectiveness of the RegSPT method

The test results show the RegSPT method to be provide robust performance for the studied metrics. Referring to the summary statistics in Table 7, RegSPT produces the best performance in all three categories. In particular, the value for MAL and the standard deviation of ML are both statistically lower when compared to the closest competitor: EDD. Since the data for neither of these metrics is normally distributed, Levene’s test is used for variance testing and the Mann–Whitney Test is used for median testing. Referring to Table 8, given that the $p$-value is less than the .05 alpha value, the ratio of the variances are considered to be unequal. Hence, the variance for RegSPT is significantly lower compared to EDD and Table 9 shows MAL for RegSPT to be likewise significantly different (lower) from that of EDD. Since lower values of absolute lateness are important for lean manufacturing (goal of zero), this is an important finding.

8. Implementation issues

RegSPT has been shown to provide more balanced performance compared to the existing sequencing methods and is easier to implement than the MDD and SPT approaches as there is no need to estimate a processing time for each job. In terms of tools required to implement the approach, all a company needs is a computer and standard software applications such as MS Excel® and possibly Minitab® (although Excel® statistical add-ons will work fine). RegSPT can be thought of as having two components: the CF regression equation and the sequencing algorithm. Because business conditions change over time, there is a need to regularly refresh the coefficients of the regression equation.

| Metric | EDD | MDD | SPT | REG SPT |
|--------|-----|-----|-----|--------|
| MAL    | 10.55 | 10.68 | 29.83 | 7.84 |
| SDL    | 11.33 | 12.12 | 24.49 | 7.11 |
| AL     | −7.06 | −7.08 | 27.01 | −7.09 |
This process is conducted offline to the everyday operation of RegSPT and given that most ETO firms use ERP systems, the historical data needed to compute the coefficients should be simple to access and export to a spreadsheet. The amount of time to accomplish this task depends on the firm’s systems/data cleanliness and availability, but it should be on the order of several hours.

Next, given that the regression equation is up-to-date, the jobs which are to be run through the engineering step must be sequenced on a regular predetermined basis (typically weekly). In order to generate the input data required for the sequencing algorithm,
the predicted process time for each job needs to be generated. In order to accomplish
this task, personnel simply need to enter the aforementioned seven pieces of complexity
information \( (x_i)'s \) into the regression equation which can be encoded into a spreadsheet
formula. At this point, it needs to be stated that the process of searching for these data
is not an extra step that results from the use of RegSPT, but is actually already currently
being performed as a normal part of the engineering review process. Because the job
package is reviewed as a normal part of the engineering review process, the seven input
data values can be quickly gathered and the RSA executed to generate the predicted
processing times. The jobs are then sequenced in ascending order of predicted
processing times. Once the input data are extracted, based upon discussions with the
five firms studied, it is estimated that the entire sequencing process should typically take
15–20 min vs. the roughly one hour that it takes today. Overall, this algorithmic
approach is quick, user-friendly, and runs using spreadsheet programs and tools that are
well known by ETO practitioners.

9. Summary
The ETO environment drives complexity and variability which, in turn, provides a
justification for administrative convenience. Administrative convenience leads personnel
to choose methods which are familiar and easy. Unfortunately, per O’Neil (1991),
administrative convenience also often pushes in a direction counter to best performance
and, in this case, the use of ad hoc approaches that yield poor results. To combat
administrative convenience, O’Neil states that simple, easy to understand robust
heuristic approaches are needed and more likely to be implemented. In fact, Stoop and
Wiers (1996) state that such simple and robust approaches are the most broadly used in
practice. Therefore, research is needed which helps practitioners in terms of choosing
the correct simple and robust scheduling rules for the ETO engineering environment.
Yet, despite this need, there has been little sequencing research focused on the ETO
engineering environment. Therefore, this research has focused on a more robust
approach for enhancing scheduling in the ETO environment concentrating on what has
been called the main bottleneck in the ETO environment, namely the engineering
process.

The engineering process is distinct from traditional shop floor processes as its inputs
and outputs are purely intellectual and not physical in nature. These and other issues
have led to an overriding question: how can scheduling capabilities which co-exist with
existing ERP systems be developed for the highly complex ETO engineering environ-
ment while, at the same time, remaining realistic, robust, acceptable, and easily imple-
mented? Restated, the literature survey and the results of the study of the partner
firms have shown that what is needed is a scheduling policy which focuses on the engineering
process, is robust and relatively simple to implement and use, and is complementary to
an ERP system.

As has been demonstrated, the RegSPT sequencing approach, as applied to the ETO
engineering environment, meets all of the above goals. First, RegSPT is robust and sim-
ple. It requires only seven pieces of information in order to accurately predict the job
processing time. These data are available in a typical job quotation package which is
received from sales. Because the job package is reviewed as a normal part of the engi-
neering review process, the seven data points can be quickly gathered and the RSA exe-
cuted. In addition, the algorithm runs using a spreadsheet program which are extremely
user-friendly and well known by ETO practitioners. Once the input data are extracted, it
is estimated that the entire sequencing process should typically take 15–20 min vs. roughly one hour today.

Second, RegSPT is accurate. The industry and literature approaches were rigorously tested against one another and RegSPT exhibited the best performance in all categories compared to its next closest competitor, EDD. In particular, it achieved statistically better results for MAL (7.84 days vs. 10.55 days). Since lower values of absolute lateness are important for lean manufacturing, this is an important finding. In addition, RegSPT achieved a statistically significantly lower SDL compared to the same competitor (7.11 days vs. 11.33 days). Given the high amount of variability found in the ETO engineering environment, these results are fairly impressive. Lastly, it is known that ERP systems have the capability to export and import data to and from the spreadsheet software. Hence, since RegSPT can easily be implemented using spreadsheet programs, it is compatible with typical ERP systems. Therefore, there is no need to jettison an existing costly ERP system.

In terms of next steps, while RegSPT has been shown to be an improved method of sequencing in the ETO engineering environment, admittedly there are other possible approaches which can take advantage of the REG approach and perhaps the use of a complexity-driven regression might yield in an incremental improvement in performance. Also, an important part of the complexity-driven regressive approach is the determination of \( p \), the number of time periods needed before data refresh. In this study, this is left up to the practitioner to determine, but further studies on an analytic method to determine the parameter \( p \) would be beneficial to the industrial community.

ORCID
John M. Usher http://orcid.org/0000-0002-4762-6989

References
Baker, K. R. (1984). Sequencing rules and due-date assignments in a job shop. Management Science, 30, 1093–1104.
Baker, K. R., & Bertrand, J. V. M. (1981). A comparison of due date selection rules. AIIE Transactions, 13, 123–131.
Baker, K. R., & Trietsch, D. (2009). Principles of sequencing and scheduling. New York, NY: Wiley.
Brucker, P., Drexel, A., Moehring, R., Neumann, K., & Pesch, E. (1999). Resource-constrained project scheduling: Notation, classification, models, and methods. European Journal of Operational Research, 112, 3–41.
Carroll, D. C. (1965). Heuristic sequencing of single and multiple component jobs (Unpublished Ph.D. dissertation). Sloan School of Management, MIT, Cambridge, MA.
Coates, G., Ritchey, I., Duffy, H. B., Hills, W., & Whitfield, R. I. (2000). Integrated engineering environments for large complex products. Concurrent Engineering: Research and Applications, 8, 171–182.
Conway, R. W. (1965a). Priority dispatching and work-in-process inventory in a job shop. Journal of Industrial Engineering, 16, 123–130.
Conway, R. W. (1965b). Priority dispatching and job lateness in a job shop. Journal of Industrial Engineering, 16, 228–237.
Conway, R. W., & Maxwell, W. L. (1962). Network dispatching by shortest operation discipline. Operations Research, 10, 51–73.
Conway, R. W., Maxwell, W. L., & Miller, L. W. (1967). Theory of Scheduling. Reading, MA: Addison-Wesley.
Cutler, T. R. (2005). Are you ETO? The increasing importance of custom (engineer-to-order) manufacturers. In MFG, pp. 1–3.
Cutler, T. R. (2009). Special orders, engineer to order challenges for the industrial engineer. *Industrial Engineer*, 41, 36–38.

Elvers, D. A. (1973). Job shop dispatching using various due date setting criteria. *Production and Inventory Management*, 14, 62–69.

Elvers, D. A., & Taube, L. (1983). Time completion for various dispatching rules in a job shop. *Omega*, 11, 81–89.

Gere, W. S. (1966). Heuristics in job shop scheduling. *Management Science*, 13, 167–190.

Gosling, J., & Naim, M. M. (2009). Engineer to order supply chain management: A literature review and research agenda. *International Journal of Production Economics*, 122, 741–754.

Grabenstetter, D. H., & Usher, J. M. (2013). Determining job complexity in an engineer to order environment for due date estimation using a proposed framework. *International Journal of Production Research*, 51, 5728–5740.

Grabenstetter, D. H., & Usher, J. M. (2014). Developing due dates in an engineer-to-order engineering environment. *International Journal of Production Research*, 52, 6349–6361.

Harrell, C., Ghosh, B. K., & Bowden, R. O. (2004). *Simulation using ProModel* (2nd ed.). New York, NY: McGraw-Hill.

Hicks, C., McGovern, T., & Earl, C. F. (2001). A typology of UK engineer-to-order companies. *International Journal of Logistics: Research and Applications*, 4, 43–56.

Hicks, C., Song, D. P., & Earl, C. F. (2007). Dynamic scheduling for complex engineer to order products. *International Journal of Production Research*, 45, 3477–3503.

Kanet, J. J., & Hayya, J. C. (1982). Priority dispatching with operation due dates in a job shop. *Journal of Operations Management*, 2, 167–175.

Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied linear regression models* (4th ed.). New York, NY: McGraw-Hill Irwin.

Little, D., Rollins, R., Peck, M., & Porter, J. K. (2000). Integrated planning and scheduling in the engineer to order sector. *International Journal of Computer Integrated Manufacturing*, 13, 545–554.

Miyazaki, S. (1981). Combined scheduling system for reducing job tardiness in a job shop. *International Journal of Production Research*, 19, 201–211.

Muhlemann, A. F., Lockett, A. G., & Farn, C. K. (1982). Job shop scheduling heuristics and frequency of scheduling. *International Journal of Production Research*, 20, 227–241.

O’Neil, P. (1991). *Performance of lot dispatching and scheduling algorithms through discrete event simulation*. IEEE/SEMI International Semiconductor Manufacturing Science Symposium (pp. 21–24). San Francisco, CA.

Pandit, A., & Zhu, Y. (2009). An ontology-based approach to support decision-making for the design of ETO (engineer to order) products. *Automation in Construction*, 16, 759–770.

Pongcharoen, P., Hicks, C., Braiden, P. M., & Stewardson, D. J. (2002). Determining optimum genetic algorithm parameters for scheduling the manufacturing and assembly of complex products. *International Journal of Production Economics*, 78, 311–322.

Pongcharoen, P., Hicks, C., & Braiden, P. M. (2004). The development of genetic algorithms for the finite capacity scheduling of complex products, with multiple levels of product structure. *European Journal of Operational Research*, 152, 215–225.

Qi, X., & Tu, F. S. (1998). Scheduling a single machine to minimize earliness penalties subject to the SLK due-date determination method. *European Journal of Operational Research*, 105, 502–508.

Salman, N., & Dogru, A. (2004). Design effort estimation using complexity metrics. *Journal of Integrated Design and Process Science*, 8, 83–88.
Song, D. P., Hicks, C., & Earl, C. F. (2001). Stage due date planning for multistage assembly systems. *International Journal of Production Research, 39*, 1943–1954.

Song, D. P., Hicks, C., & Earl, C. F. (2002). Product due date assignments for complex assemblies. *International Journal of Production Research, 76*, 4877–4895.

Song, D. P., Hicks, C., & Earl, C. F. (2006). An ordinal optimization based evolution strategy to schedule complex make to order products. *International Journal of Production Research, 44*, 4877–4895.

Stevenson, M., Hendry, L. C., & Kingsman, B. G. (2005). A review of production planning and control: The applicability of key concepts to the make-to-order industry. *International Journal of Production Research, 43*, 869–898.

Stoop, P. P. M., & Wiers, V. C. S. (1996). The complexity of scheduling in practice. *International Journal of Operations and Production Management, 16*, 37–53.

Weeks, J. K. (1979). A simulation study of predictable due dates. *Management Science, 25*, 363–373.

Worthen, B. (2002). Nestlé’s ERP Odyssey. *CIO, 15*, 1–9.