mathematics

MDPI

Article

Work Faster, Work in Parallel, or Work Overtime? An Assessment of Short-Term Capacity Adjustments by Simulation

Nuno O. Fernandes [1,2,*], Matthias Thürer [3] and Federica Costa [4]

[1] Department of Industrial Engineering, Instituto Politécnico de Castelo Branco, Av. do Empresário, 6000-767 Castelo Branco, Portugal
[2] Campus de Gualtar, ALGORITMI Research Unit, University of Minho, 4710-057 Braga, Portugal
[3] Chair of Factory Planning and Intralogistics, Chemnitz University of Technology, Erfenschlager Str. 73, 09125 Chemnitz, Germany; matthiasthurer@workloadcontrol.com
[4] Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Via Lambruschini 4/b, 20156 Milano, Italy; federica.costa@polimi.it
* Correspondence: nogf@ipcb.pt

Abstract: Many firms use short-term capacity adjustments to deal with demand changes over time, and a broad literature assesses when and where to adjust capacity. This study highlights that this may be dependent on the type of capacity flexibility used to actually realize the adjustment. By comparing for the first-time capacity adjustments by speeding up processing rates, working in parallel, or using overtime, significant differences in the operational performance are identified, with the latter resulting in the best percentage tardy performance. This provides important insights for the interpretation of the existing literature, guides the future literature, and helps managers to make better decisions.

Keywords: short-term capacity adjustments; overtime; state-dependent processing rates; parallel work; simulation

MSC: 90B30

Citation: Fernandes, N.O.; Thürer, M.; Costa, F. Work Faster, Work in Parallel, or Work Overtime? An Assessment of Short-Term Capacity Adjustments by Simulation. *Mathematics* **2024**, *12*, 2515. https://doi.org/10.3390/math12162515

Academic Editor: Andrea Scozzari

Received: 5 July 2024
Revised: 6 August 2024
Accepted: 9 August 2024
Published: 15 August 2024

1. Introduction

Capacity control determines the resources required to sustain a given demand over a planning period [1]. This study focuses on short-term capacity control to deal with demand changes over time. Short-term capacity control decides when, where, and how to use flexible capacity. Short-term flexible capacity itself can, among others, refer to: (i) workers that work faster or slower, e.g., in the context of state-dependent processing times [2–5]; (ii) workers that are able to help each other working in parallel [6,7]; or (iii) overtime [8]. However, it remains unknown whether there is any operational difference across these different types of flexible capacity and what the size of this difference is. Therefore, this paper seeks to answer the following research question:

What is the operational effect of workers working faster, parallel work, and overtime?

Answering this question is important since it guides Industrial Engineers in practice on how to design capacity control mechanisms to confidently balance idiosyncratic capacity flexibility costs and potential performance gains. From a research perspective, this study contributes by clarifying whether the literature that models one type of capacity flexibility also holds for other types. This is specifically important in the context of Industry 4.0, which often relies on the accurate modeling of production planning and scheduling methods to make decisions.

This study uses discrete event simulation to substantiate the operational effect of capacity adjustments and answer the above research question. Simulation was chosen since it has been an important mathematical tool for Industrial Engineers for over fifty years, as pointed out by Nelson [9], and allows for the modeling of complex systems. The remainder of this study is structured as follows. Section 2 introduces state-dependent processing times, parallel work, and overtime before the simulation model used to evaluate the different types of flexible capacity is introduced in Section 3. The results are presented and discussed in Section 4. Section 5 concludes this paper.

2. Background

According to Bertrand and Wortman [10], there are three main production control functions: due date setting, order release (input) control, and capacity (output) control. This study focuses on output control, which is realized through capacity management and which can range from long-term capacity planning to short-term capacity control [11–13]. Capacity control, the focus of this study, requires two elements: flexible capacity and a control mechanism that uses this flexibility. Most of the literature focuses on the latter [14]. To the best of our knowledge, no study to date has assessed the operational impact of the type of flexibility.

In fact, short-term capacity adjustments play an important role in manufacturing, and different mechanisms have emerged to decide when, where, and how to adjust capacity. But most of the literature has focused on when and where, for example, in the context of dual-resource constraint jobs (e.g., [15]). There are different methods to reach this decision, including simple when and where rules (e.g., [16]), heuristics solving mathematical models (e.g., [17]), game-theoretical approaches (e.g., [5,15]), and simulation (e.g., [18]). A recent review of resource scheduling can also be found in Geurtsen et al. [19]. This study focuses on the question of how capacity is adjusted. It seeks to evaluate whether this question matters, and if yes, what its operational impact is. The three main ways that capacity is adjusted will be introduced next.

2.1. Adjusting Processing Rates

Gomersall [20] has already observed the so-called backlog syndrome, which states that workers have a preferred backlog of work. They will consequently work slower if the current workload is below this preferred level, and they will work faster otherwise. Meanwhile, processing rates may also change due to hazard level [21] or fatigue [22]. There has consequently been significant research attention on state-dependent processing rates (e.g., [2,4,23–26]). There are different ways to adjust the processing rates. An example is discretionary task completion criteria, where a worker determines by their subjective standards how an operation is performed [3]. Batt and Terwiesch [27] further mention rushing, task reduction, and multi-tasking. Multi-tasking links into the second type of capacity flexibility considered in this study, which will be discussed next.

2.2. Parallel Work

Instead of a single worker working faster, the worker can also be supported by a second worker. Workers can then work in parallel. Putting more workers to work on a station increases the amount of processing work that can be performed per time-period. This multi-tasking can also be realized as part of bucket brigades [28,29], Baton Zone Bumping [30], or Variable Takt Time [7,31]. In a bucket brigade, the workers move between adjacent stations to continue working on an item. Workers work on a single item and pass it on to successive stations when they meet the next worker in the sequence. The worker then moves backward to take the predecessor's item. Baton zone bumping and Variable Takt Time create overlap zones along the production lines at which this passing can occur. If there is a connected production line in which products move at a certain speed, then the space allocated to a worker along the line represents the takt time. Creating zones where there is overlap allows for variability in the capacity allocation (i.e., takt times). Variable Takt Time additionally introduces a "utility operator". Utility workers handle the overload either by being dispatched to assist the regular workers during peak load situations, or by being stationed at various points along the assembly line to complete the unfinished operations [32,33]. Alternatively, workers could also help each other. Portioli-Staudacher et al. [6] assessed the impact of workers helping each other in flow shops with different labor flexibility configurations. Utility operators or workers helping each other assume that either the additional capacity can speed up the processing of the current item, or the additional capacity can avoid the propagation of the overload by starting another item in parallel. In the former case, the effect is equivalent to a reduction in processing times realized by a worker working faster. In the latter case, a different operational impact can be expected.

2.3. Overtime

Overtime is defined as an extension of the daily working time over the planning period. Using overtime allows firms to adjust their capacity while maintaining the same work system. Overtime is an easy-to-use tool for workforce flexibility that is widely applied, and few firms do not recur to overtime if there is an overload situation. Surprisingly, no study to date has assessed whether working overtime is the best solution for adjusting capacity in the short term. Overtime assumes a capacity buffer that can be used. An example is Toyota running two shifts instead of three to allow shifts to make up any shortfalls on their production quotas [8]. Even if overtime materializes only after the normal working time is over, a distinction between scheduled and unscheduled overtime is usually made. For example, authors such as Ingels and Maenhout [34] focus on different time-based strategies to include overtime in the personnel shift roster, while Akkan [35] proposes a scheduling approach to determine when and how much overtime is required to meet a requested due date. The previous two types of capacity flexibility can materialize at any moment during normal working time.

2.4. A Discussion of the Literature

There are different approaches to realize short-term capacity adjustments. These are summarized in Table 1, along with the key literature.

Table 1. A Summary of the Different Approaches to Adjusting Capacity.

	Work Faster	Parallel Work	Overtime
Resource Reaction	Resources react to the current state of the system.	Resources react to the current state of the system given rules.	Resources react to the aggregated state of the system.
Processing Time Adjustment	Processing time adjustment is normally a continuous range.	Processing time adjustment is discrete.	Processing times remain equal, but the availability of capacity is increased.
Organizational Requirements	It can be encouraged by management, but it is not an organizational policy.	It is an organizational policy that requires rules and workplaces that allow for parallel work.	It is an organizational policy that requires rules and a capacity buffer.
Key Literature	Gomersall [20]; Schultz et al. [2]; Batt and Terwiesch [27].	Bratcu and Dolgui [29]; Portioli-Staudacher et al. [6].	Akkan [35]; Hopp and Spearman [8]; Ingels and Maenhout [34].

Until now, most of the literature has focused on when and where to adjust capacity. It was implicitly assumed that this is more important than the question of how capacity adjustment is realized. Previous research, therefore, did not compare the different types of capacity adjustments summarized in Table 1. In fact, authors use only one type of adjustment, and studies are not comparable, because they use different models and research methods. But the different types of flexibility needed to realize capacity adjustments are inherently different. State-dependent processing times produce the same job faster, parallel work produces two jobs at the same speed, and overtime means additional jobs are produced at the end of a shift. To answer the question of how capacity adjustment should be realized, discrete event simulation will be used. The applied model will be introduced next.

3. Simulation

3.1. Job and Shop Characteristics

A simulation model of a serial production line was implemented in SIMIO software, as illustrated in Figure 1. We used a generalized simulation model to avoid any factors that could interfere with the interpretation of our results. However, this kind of production system can be found in practice in several situations. One typical case is the assembly of heat shrink equipment for cable harness manufacturers observed by the authors. In this study, the simulation methodology of Banks [36] was used, while the model was validated by using graphical animation and running the model under extreme operational conditions.

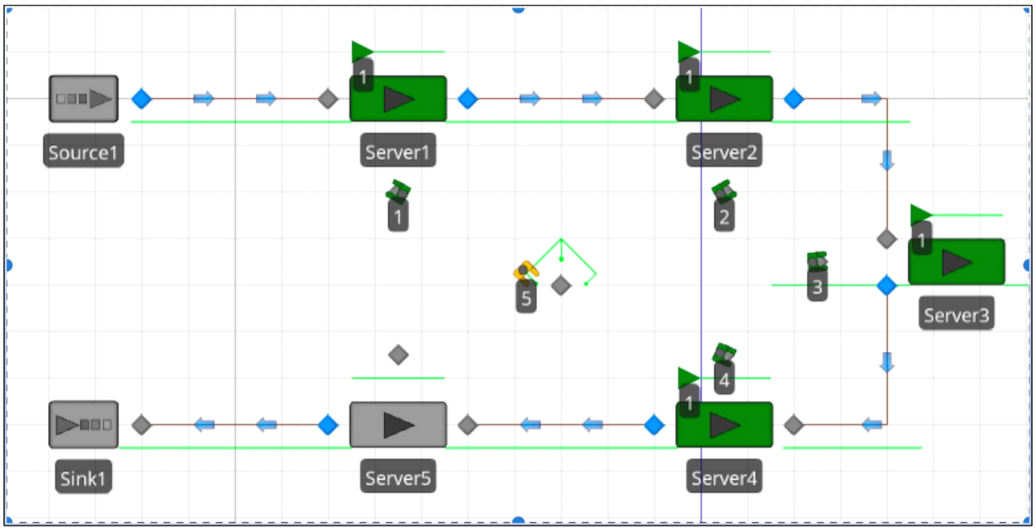

Figure 1. SIMIO model of a U-shaped line.

The production line is U-shaped and composed of five stations. There is one worker per station, so the shop is fully stuffed. Each station is modeled as a capacitated resource that allows for the simultaneous operation of two workers. The inter-arrival time of jobs and the operation processing times are stochastic. The inter-arrival time follows an exponential distribution with a mean of 0.8 time-units to implement stochastic processing with independent arrivals. The operation processing times follow a lognormal distribution with a mean of 0.5 time-units and a coefficient of variation of 0.8 time-units. The lognormal distribution was argued to better reflect real-life activity times than the exponential or 2-Erlang distribution [37]. The normal working time for the production line is 16 time-units, while demand occurs for 24 time-units. To keep our study focused, the following assumptions were also made:

- Stations have perfect yield.
- Setup times are sequence-independent and included in the operation processing time.
- Jobs' and workers' moving times between stations are assumed to be negligible.

To keep the due date assignment simple, a constant delivery time allowance of 56 time-units is added to the order arrival time. This results in about 24% of tardy jobs when no capacity adjustment is possible.

3.2. Short-Term Capacity Adjustments

The three different approaches summarized in Table 1 are considered. To ensure comparability, we model processing time adjustment and parallel work by workers helping each other similarly. There is one worker per station. A second cross-trained worker from the upstream station can help. When capacity adjustment is modeled using parallel work, then the second worker works on a different product unit. When processing time adjustment is modeled, then the processing time is reduced by 50%, and both workers work on the same product unit. For both options, the worker helps at a downstream station only if there are no waiting jobs in the queue of the station at which she/he is working (home station) and returns to her/his station after finishing the current job if one or more jobs are waiting at the home station.

For overtime, the production shift is extended. The overtime duration is determined by the reduction in machine processing hours for the above two options. As a result, we do have two settings for overtime. Finally, we also include the option of no capacity adjustment as a baseline.

3.3. Experimental Design and Performance Measures

Five experimental scenarios were considered: no capacity adjustment, processing time adjustment, parallel work, and overtime (2 settings). Each scenario was replicated 100 times, and the results were collected for over 10,000 time-units following a warm-up period of 2000 time-units. The number of replications and the length of each simulation run directly impact the confidence intervals of the performance measures. We therefore adjusted these two simulation parameters in the preliminary simulation experiments to result in confidence intervals that are sufficient for our analysis while keeping simulation run time at a reasonable level.

Four main performance measures were recorded: (i) time in system, i.e., the mean time between the arrival time and the completion time of a job; (ii) percentage of tardy jobs, i.e., the percentage of jobs completed after the due date; (iii) worker utilization, i.e., the percentage of time the worker is busy; and (iv) machine utilization, i.e., the percentage of time the machine is busy. The significance of the differences between the outcomes of individual experiments were, therefore, verified with paired t-tests which comply with the use of common random number streams to reduce variation across experiments. Whenever we discuss a difference in outcomes between two experiments, the significance can be proven by a paired t-test at a level of 95%.

4. Results

4.1. Assessment of Results

The results for the five different experimental scenarios are given in Table 2. The results indicate a similar worker utilization which ensures comparability. If two workers work on the same product unit, then worker utilization is not affected, because the 50% reduction means a 50% increase for the other worker. Only machine processing time is reduced. If there is parallel work, then the total amount of work that a worker has to execute is not affected. In both cases, the only effect is on the time a job spends in the machine, and, thus, machine utilization. As somewhat expected, the reduction in machine utilization for processing time adjustment is about 50% of the reduction for parallel work, since the helping time is sped up. Meanwhile, parallel work may lead to more discrete arrivals at downstream stations, while working faster may reduce inter-arrival variability. But this effect only affects time in system and percentage tardy results. For overtime, there is no effect for the first 16 time-units.

In terms of time in system and percentage of tardy jobs, the following can be observed:

- It is apparent that an increase in overtime improves performance. In fact, overtime appears to result in the best percentage tardy performance (1.5%).
- Parallel work leads to better performance than processing time adjustments (3.2% vs. 10% in terms of percentage tardy, and 23.2 vs. 33.4 time-units in terms of time in system).
- There are significant differences according to the type of flexibility used to realize capacity adjustments.

Table 2. Summary of results.

Option	Time In System (1)	Percentage Tardy (1)	Worker Utilization (16 Time-Units)	Machine Utilization (16 Time-Units)
None	44.5 ± 1.27	23.7% ± 2.07%	93.60%	93.60%
Adjust Processing Times	33.4 ± 0.89	10.0% ± 1.27%	94.30%	89.80%
Parallel Work	23.2 ± 0.66	3.2% ± 0.75%	93.90%	84.90%
Overtime (+29 min.)	32.4 ± 0.74	7.0% ± 1.05%	93.60%	93.60%
Overtime (+57 min.)	25.0 ± 0.44	1.5% ± 0.36%	93.60%	93.60%

(1) 95% confidence interval on the mean.

4.2. Discussion of Results

Firms are finding it more and more difficult to meet increasingly personalized consumer demand by using inventory alone. Short-term capacity adjustments are, therefore, required to deal with consumer preference changes over time. Our study makes an informed decision on how this can be realized. Managers in practice can use our operational measures and adapt them to their idiosyncratic company characteristics and cost structures to make better decisions.

Our results on capacity adjustments impact the previous literature. The question of when and where to adjust capacity may be impacted by the question of how capacity is adjusted. The literature that focuses on scheduling capacity adjustment [14], event-driven capacity control [6,38], or factory planning [39] should consider potential two-way interactions. In fact, the results may be different according to the type of capacity flexibility used to actually materialize the adjustment.

Parallel work is better than speeding up an item, because the time of the adjustment is longer. But it may incur higher costs during set-up of the workspace. It may also require a parallel flow of material, which, in many practical situations, is not easy to realize. Overtime leads to the best results since it clears the backlog. In a two-shift system, non-working hours also contribute to the time in system. But overtime incurs additional labor costs. In contrast, parallel work and working faster use existing worker capacity. But parallel work requires more worker training since workers must be able to work on more than one station. The same holds for working faster if it is realized, as in our simulation. These considerations highlight that the choice of the type of capacity flexibility to use is strongly dependent on idiosyncratic company characteristics. Our study provides an important means to better balance the benefits of the different types with the incurred costs for practicing managers. Future research could further explore this multi-dimensional decision problem.

5. Conclusions

Short-term capacity adjustments are a major means to realize output control. But the existing literature has focused mainly on when and where to adjust capacity. It was implicitly assumed that how capacity adjustments are realized is less important. Our results question this assumption, showing that there are significant differences in operational performance across realizing capacity adjustments through adjusting processing rates, working in parallel, or working overtime. This not only has implications for future research but also helps managers in practice to make better decisions when balancing output control and costs. In general, the decision on how to adjust capacity is constrained by the organizational structure of the organization. Working faster requires motivation, working in parallel requires assignment rules and workplaces that allow for working in parallel, and working overtime requires a capacity buffer and the availability of workers to do so. There is consequently a strong link between managerial decisions on organizational structure, capacity adjustments, and the different methods used to decide when and where to adjust capacity that needs to be considered when making decisions. Our study contributes by informing these decisions.

In this study, we assumed cross-trained workers can realize different tasks with an identical proficiency level to keep the three different adjustment types comparable. Therefore, this study did not consider the potential degradation of service times when multiple tasks are performed by the same worker. Future research work on this, together with the impact of fatigue for working faster and overtime, is needed to address this limitation of our study. Moreover, our study used a relatively limited experimental setting to keep it focused. Future research could explore performance differences while considering different decision rules and different environmental factors. Future research is also needed to provide better decision models that balance the decisions of when and where to adjust capacity, the decision of which type of capacity flexibility to use to actual realize the adjustments, and costs. Our study highlights the need for these models.

Author Contributions: Conceptualization, N.O.F., M.T. and F.C.; methodology, N.O.F., M.T. and F.C.; software, N.O.F.; validation, N.O.F., M.T. and F.C.; writing—original draft preparation, M.T.; writing—review and editing, N.O.F., M.T. and F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FCT–Fundação para a Ciência e Tecnologia grant number UIDB/00319/2020.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, [N.F.], upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, C.-S.; Mestry, S.; Damodaran, P.; Wang, C. The capacity planning problem in make-to-order enterprises. *Math. Comput. Model.* **2009**, *50*, 1461–1473. [CrossRef]
2. Schultz, K.L.; Juran, D.C.; Boudreau, J.W.; McClain, J.O.; Thomas, L.J. Modeling and Worker Motivation in JIT Production Systems. *Manag. Sci.* **1998**, *44*, 1595–1607. [CrossRef]
3. Hopp, W.J.; Iravani, S.; Yuen, G.Y. Operations systems with discretionary task completion. *Manag. Sci.* **2007**, *53*, 61–77. [CrossRef]
4. Bertrand, J.W.M.; van Ooijen, H.P.G. Workload Control order release and productivity: A missing link. *Prod. Plan. Control* **2002**, *13*, 665–678. [CrossRef]
5. Kose, Y.; Cevikcan, E.; Ertemel, S.; Murat, M. Game theory-oriented approach for disassembly line worker assignment and balancing problem with multi-manned workstations. *Comput. Ind. Eng.* **2023**, *181*, 109294. [CrossRef]
6. Portioli-Staudacher, A.; Costa, F.; Thürer, M. The use of labour flexibility for output control in workload controlled flow shops: A simulation analysis. *Int. J. Ind. Eng. Comput.* **2020**, *11*, 429–442. [CrossRef]
7. Mönch, T.; Huchzermeier, A.; Bebersdorf, P. Variable takt times in mixed-model assembly line balancing with random customization. *Int. J. Prod. Res.* **2021**, *59*, 4670–4689. [CrossRef]
8. Hopp, W.J.; Spearman, M.L. To pull or not to pull: What is the question? *Manuf. Serv. Oper. Manag.* **2004**, *6*, 133–148. [CrossRef]
9. Nelson, B.L. Rebooting simulation. *IISE Trans.* **2023**, *56*, 385–397. [CrossRef]
10. Bertrand, J.W.M.; Wortmann, J.C. *Production Control and Information Systems for Component-Manufacturing Shops*; Elsevier Scientific Publishing Company: Amsterdam, The Netherlands, 1981.
11. Wortman, J.C.; Euwe, M.J.; Taal, M.; Wiers, V.C.S. A review of capacity planning techniques within standard software packages. *Prod. Plan. Control* **1996**, *7*, 117–128. [CrossRef]
12. Olhager, J.; Rudberg, M.; Wikner, J. Long-term capacity management: Linking the perspectives from manufacturing strategy and sales and operations planning. *Int. J. Prod. Econ.* **2001**, *69*, 215–225. [CrossRef]
13. Denkena, B.; Dittrich, M.A.; Wilmsmeier, S.; Stamm, S. Optimization of delivery adherence based on capacity planning and bid pricing. *Prod. Eng.* **2020**, *14*, 309–318. [CrossRef]
14. Altendorfer, K.; Hübl, A.; Jodlbauer, H. Periodical capacity setting methods for make-to-order multi-machine production systems. *Int. J. Prod. Res.* **2014**, *52*, 4768–4784. [CrossRef]
15. Renna, P.; Thürer, M.; Stevenson, M. A game theory model based on Gale-Shapley for dual-resource constrained (DRC) flexible job shop scheduling. *Int. J. Ind. Eng. Comput.* **2020**, *11*, 173–184. [CrossRef]
16. Thürer, M.; Zhang, H.; Stevenson, M.; Costa, F.; Ma, L. Worker assignment in dual resource constrained assembly job shops with worker heterogeneity: An assessment by simulation. *Int. J. Prod. Res.* **2020**, *58*, 6336–6349. [CrossRef]
17. ElMaraghy, H.; Patel, V.; Abdallah, I.B. Scheduling of manufacturing systems under dual-resource constraints using genetic algorithms. *J. Manuf. Syst.* **2000**, *19*, 186–201. [CrossRef]

18. Arakawa, M.; Fuyuki, M.; Inoue, I. An optimization-oriented method for simulation-based job shop scheduling incorporating capacity adjustment function. *Int. J. Prod. Econ.* **2003**, *85*, 359–369. [CrossRef]
19. Geurtsen, M.; Didden, J.B.; Adan, J.; Atan, Z.; Adan, I.J.B.F. Production, maintenance and resource scheduling: A review. *Eur. J. Oper. Res.* **2023**, *305*, 501–529. [CrossRef]
20. Gomersall, E.R. The Backlog Syndrome. *Harv. Bus. Rev.* **1964**, *92*, 105–115.
21. Zhu, L.; Zhang, Z.; Wang, Y. A Pareto firefly algorithm for multi-objective disassembly line balancing problems with hazard evaluation. *Int. J. Prod. Res.* **2018**, *56*, 7354–7374. [CrossRef]
22. Delasay, M.; Ingolfsson, A.; Kolfal, B. Modeling Load and Overwork Effects in Queueing Systems with Adaptive Service Rates. *Oper. Res.* **2016**, *64*, 867–885. [CrossRef]
23. Van Ooijen, H.P.G.; Bertrand, J.W.M. The effects of a simple arrival rate control policy on throughput and work-in-process in production systems with workload dependent processing rates. *Int. J. Prod. Econ.* **2003**, *85*, 61–68. [CrossRef]
24. Öner-Közen, M.; Minner, S.; Steinthaler, F. Efficiency of paced and unpaced assembly lines under consideration of worker variability—A simulation study. *Comput. Ind. Eng.* **2017**, *111*, 516–526. [CrossRef]
25. Shunko, M.; Niederhoff, J.; Rosokha, Y. Humans Are Not Machines: The Behavioral Impact of Queueing Design on Service Time. *Manag. Sci.* **2018**, *64*, 453–473. [CrossRef]
26. Thürer, M.; Stevenson, M.; Aitken, J.; Silva, C. State-Dependent Service Rates in Make-to-Order Shops: An Assessment by Simulation. *Oper. Manag. Res.* **2020**, *13*, 70–84. [CrossRef]
27. Batt, R.J.; Terwiesch, C. Early Task Initiation and Other Load-Adaptive Mechanisms in the Emergency Department. *Manag. Sci.* **2017**, *63*, 3531–3551. [CrossRef]
28. Bartholdi, J.J., III; Bunimovich, L.A.; Eisenstein, D.D. Dynamics of two-and three-worker "bucket brigade" production lines. *Oper. Res.* **1999**, *47*, 488–491. [CrossRef]
29. Bratcu, A.I.; Dolgui, A. A survey of the self-balancing production lines ("bucket brigades"). *J. Intell. Manuf.* **2005**, *16*, 139–158. [CrossRef]
30. Protzman, C.; Whiton, F.; Kerpchar, J.; Lewandowski, C.; Stenberg, S.; Grounds, P. *The Lean Practitioner's Field Book: Proven, Practical, Profitable and Powerful Techniques for Making Lean Really Work*; Productivity Press: New York, NY, USA, 2018.
31. Huchzermeier, A.; Mönch, T. Mixed-model assembly lines with variable takt and open stations. *Prod. Oper. Manag.* **2023**, *32*, 704–722. [CrossRef]
32. Yano, C.A.; Rachamadugu, R. Sequencing to minimize work overload in assembly lines with product options. *Manag. Sci.* **1991**, *37*, 572–586. [CrossRef]
33. Cevikcan, E.; Durmusoglu, M.B. Minimising utility work and utility worker transfers for a mixed-model assembly line. *Int. J. Prod. Res.* **2011**, *49*, 7293–7314. [CrossRef]
34. Ingels, J.; Maenhout, B. The impact of overtime as a time-based proactive scheduling and reactive allocation strategy on the robustness of a personnel shift roster. *J. Sched.* **2018**, *21*, 143–165.
35. Akkan, C. Overtime Scheduling: An Application in Finite-capacity Real-time Scheduling. *J. Oper. Res. Soc.* **1996**, *47*, 1137–1149. [CrossRef]
36. Banks, J. *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*; Wiley: Hoboken, NJ, USA, 1998.
37. Trietsch, D.; Mazmanyan, L.; Gevorgyan, L.; Baker, K.R. Modeling activity times by the Parkinson distribution with a lognormal core: Theory and validation. *Eur. J. Oper. Res.* **2012**, *216*, 386–396.
38. Land, M.J.; Stevenson, M.; Thürer, M.; Gaalman, G.J.C. Job Shop Control: In Search of the Key to Delivery Improvements. *Int. J. Prod. Econ.* **2015**, *168*, 257–266.
39. Schäfer, L.; Klenk, F.; Maier, T.; Zehner, M.; Peukert, S.; Linzbach, R.; Treiber, T.; Lanza, G. A systematic approach for simulation-based dimensioning of production systems during the concept phase of factory planning. *Prod. Eng.* **2024**, 1–13. [CrossRef]