A PRODUCTION LINE THAT BALANCES ITSELF

John J. Bartholdi, III [*]   Donald D. Eisenstein [†]

July 4, 1993; revised May 18, 1994, and September 7, 1995

To appear in *Operations Research*

Abstract

In "bucket brigade" manufacturing, such as recently introduced to the apparel industry, a production line has n workers moving among m stations, where each worker independently follows a simple rule that determines what to do next. Our analysis suggests and experiments confirm that if the workers are sequenced from slowest to fastest then, independently of the stations at which they begin, a stable partition of work will spontaneously emerge. Furthermore, the production rate will converge to a value that, for typical production lines, is the maximum possible among all ways of organizing the workers and stations.

Key words: *production line, line-balancing, self-organizing systems, discrete dynamical systems, fixed point*

[*]School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0205 USA. E-mail: john.bartholdi@isye.gatech.edu

[†]Graduate School of Business, The University of Chicago, Chicago, Illinois 60637 USA. E-mail: don.eisenstein@gsb.uchicago.edu

1 TSS production lines

Traditional means of organizing a production line, such as a classical assembly line, are inflexible. In a classical assembly line, workers are assigned fixed work stations and the station with the greatest work content determines the production rate. Realistically, there are only two ways to change the production rate: either change the number of shifts or else redistribute the tasks, tools, and parts over different stations. The first allows only coarse adjustments and the second is expensive and disruptive.

It is particularly important that production systems be flexible when products have extreme seasonalities or short life-cycles, such as in the apparel industry. To increase flexibility of production in the apparel industry, a variation of the assembly line has recently been introduced in which there are fewer workers than stations and workers walk to adjacent stations to continue work on an item. Control of the line is decentralized: Each worker independently follows a simple rule that determines what to do next. This idea has been commercialized by Aisin Seiki Co., Ltd., a subsidiary of Toyota, and named the "Toyota Sewn Products Management System", or TSS[1]. TSS is used in the manufacture of many types of sewn products, including apparel, furniture, shoes, hand bags, suitcases, and fish nets.

Here is how TSS works. Call each instance of the product an *item* and consider a flow line in which each of a set of items requires processing on the same sequence of m work stations, as in Figure 1. A station can process at most one item at a time, and exactly one worker is required to accomplish the processing. All items are identical and so each requires the same total processing time according to some work standard, which we normalize to one "time unit". Let the processing requirement at station j be p_j, a fixed percentage of the total standard work content of the product.

[1] "TSS" is a registered trademark of Aisin Seiki Co., Ltd. TSS is marketed in the western hemisphere by Americas 21st, Inc.

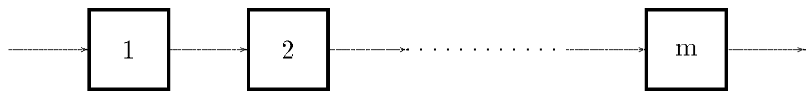

Figure 1: A simple flow line in which each item requires processing on the same sequence of work stations.

A TSS line functions as a sort of "bucket brigade" in which each worker carries an item from station to station, processing it at each station, until passing it off to a subsequent worker. This behavior can be realized by numbering the workers from 1 to n according to their sequence on the line (in the direction of product flow) and requiring each worker to independently follow this rule:

TSS Rule (forward part) Remain devoted to a single item, and process it on successive work stations (where at any station the worker of higher index has priority). If your item is taken over by your successor (or if you are the last worker and you complete processing the item), then relinquish the item and begin to follow the Backward Part.

TSS Rule (backward part) Walk back and take over the item of your predecessor (or, if you are the first worker, pick up raw materials to start a new item). Begin to follow the Forward Part.

Two points are worth emphasizing about the TSS rule. The first is that a worker can be *blocked* during the forward phase if she[2] is ready to work at a station but that station is occupied by another worker. Under the TSS rule each worker remains busy unless blocked, in which case she must wait for the station to become available.

The second point is that during the backward phase each worker interrupts her predecessor to take over her work. In effective TSS lines a great deal of effort

[2]Nearly all the TSS workers we have seen are female. This reflects the demographics of the apparel industry.

is invested to avoid introducing idle time during the “hand-off”. For example, the line is configured in a U-shape to reduce walking time, workers sew standing up at waist-high machines, and workers practice making the hand-off smoothly.

A typical TSS line that we observed was devoted to the production of women's slacks. It had seven stations and was staffed by three workers. The first worker began sewing cut cloth; and the last worker ironed the slacks, attached labels, folded and packaged the slacks, and put them in a box for shipping. The total work content of the slacks was about seven minutes, with 45–90 seconds of work at each station. It was no more than 2–3 seconds from the time the last worker finished a pair of slacks and began walking back to take over from her predecessor until the first worker began a new pair of slacks.

In fact none of the TSS lines we have observed formalized a TSS “rule” nor did they follow our rule to the letter. Instead there have always been local improvisations to account for particulars of the product, the equipment, or the team members. Moreover the standard implementation of TSS has changed since its introduction to the US apparel industry in 1989. The TSS “rule” is our abstraction, which captures the essential behavior of TSS lines.

Many questions leap to mind: How does a TSS line behave? Is TSS effective? How does one control or even predict the production rate? To be sure, factory managers are discovering answers to these questions on the shop floor. Here we begin a formal analysis of a model of TSS. Our results suggest and experiment confirms that if TSS workers are sequenced from slowest to fastest then, during the natural operation of the line, the work content of the product will be spontaneously reallocated among the workers to balance the line—without conscious intention by the workers and without intervention by management. This capacity for self-organization allows management to fine-tune the production rate by simply changing the number of workers on the line, which in turn elicits a spontaneous reallocation of work.

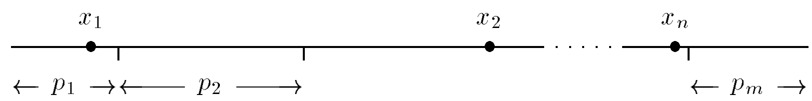

Figure 2: The standard work content of the product is represented as a line segment normalized to length 1, which is partitioned into intervals corresponding to the work stations. The position of worker i is given by x_i, the cumulative fraction of work content completed on her item.

2 A model of TSS

It is difficult to visualize the flow of items on a TSS line because the workers move asynchronously and are not explicitly limited to any particular set of stations. However, we can view the line as a dynamical system (Devaney, 1989) by expressing the position of worker i as the fraction x_i of work completed on her item, as illustrated in Figure 2; then the state of the system at any time can be summarized by the vector of worker positions $\mathbf{x} = (x_1, \ldots, x_n)$. The phase space of the system is a subset of the closed n-cell $\{ (x_1, \ldots, x_n) : 0 \leq x_1 \leq \cdots \leq x_n \leq 1 \}$ as illustrated in Figure 3. The inequalities $0 \leq x_1 \leq \cdots \leq x_n \leq 1$ arise because the TSS rule does not allow workers to pass one another.

We model each worker i by a velocity function $v_i(x)$ that gives her instantaneous work velocity at position $x \in [0, 1]$ (when not blocked by an occupied station). To avoid modeling pathologies we require that

- Each v_i is continuous almost everywhere on $[0, 1]$; and
- There exist numbers b and B such that, for each worker i, $0 < b < v_i(x) < B < \infty$ for all $x \in [0, 1]$.

Roughly speaking, these restrictions say that a worker cannot abruptly change speed at every instant; and workers are neither infinitely fast nor infinitessimally slow.

Our model of worker skills includes one used by the apparel industry, wherein each worker has a documented skill profile giving her velocities at different tasks

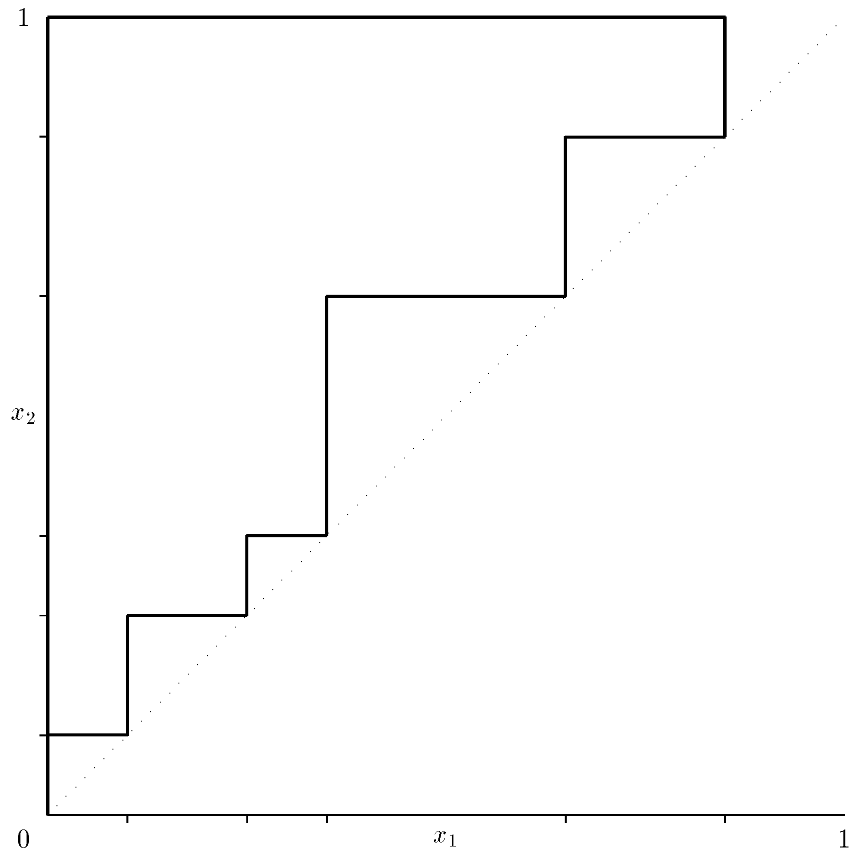

Figure 3: The phase space of a TSS line with two workers, whose positions are given by (x_1, x_2), is that portion of the upper triangle outlined in bold. The feasible region is above the dotted line because worker 1 can never pass worker 2. The tick marks on the axes correspond to the partition of work among the stations and the saw-toothed edge of the phase space arises because no more than one worker can use a station at a time.

as a percentage of work standards. This corresponds to a v_i that is a step function.

The TSS lines we have seen required no more than a few seconds for any worker to walk back and take over the work of her predecessor. Therefore we make the modeling assumption that, while the workers move forward with finite velocities, they move backward with effectively infinite velocity, so that when the last worker finishes an item, then—*at the same instant*—worker n takes over from worker $n-1$, who takes over from worker $n-2$, …, who takes over from worker 1, who introduces a new item into the system. We say that the line *resets* at such an instant. This simplification frees us from worry about the details of the continuous-time evolution of the TSS line; instead we can restrict our attention to the sequence $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(t)}, \ldots\}$ of worker positions at those instants immediately after the line resets[3]. Note that $x_1^{(t)} = 0$ and for convenience we define $x_{n+1}^{(t)} = 1$.

In the terminology of dynamical systems, $\mathbf{x}^{(t)}$ is the t-th *iterate* of the TSS system and the sequence of worker positions is the *orbit* beginning at $\mathbf{x}^{(0)}$. Let f be the function, defined implicitly by the TSS rule, that maps the vector of worker positions after one reset to that after the subsequent reset, so that $\mathbf{x}^{(t+1)} = f(\mathbf{x}^{(t)})$. Now we can study the behavior of a TSS line by studying its orbits, where each orbit $\{\mathbf{x}^{(t)} = f^t(\mathbf{x}^{(0)})\}_{t=0}^{\infty}$ is determined by the initial positions $\mathbf{x}^{(0)}$ of the workers.

3 Tracking the behavior of a TSS line

3.1 Repeatable behavior

A TSS production line is "balanced" if each worker repeats the same interval of work content on successive items. A balanced line produces at a steady rate;

[3] Such a subset of the phase space is called a *Poincaré section*. Restricting the phase space in this manner is a standard technique for analyzing dynamical systems (Morrison, 1991).

and each worker can concentrate on a subset of the work content. Because the TSS rule imposes no apparent restrictions on where a worker might move, it is not clear whether a line can maintain balance or indeed whether balance can be achieved at all.

Our first result shows that balance is always at least theoretically possible; that is, there exist worker positions such that, if the workers begin at these positions, then, after completion of each item, they will reset to exactly these same positions to begin work on the subsequent item.

Theorem 1 *For any TSS line, there exists a fixed point* $\mathbf{x}^* = f(\mathbf{x}^*)$*; that is, there exist worker positions* $\mathbf{x}^*$ *such that if the workers start at positions* $\mathbf{x}^*$*, then they will always reset to* $\mathbf{x}^*$*.*

Proof Appendix, section A.1. □

One of the main concerns of this paper will be to determine how a TSS line can be designed to operate at or near its point of balance (that is, its fixed point) so that the workers repeat the same portions of work.

4 Self-organizing behavior

We say that worker j is *faster* than worker i if she is discernibly faster at every portion of work content:

$$\sup_{x\in[0,1]} \left(\frac{v_i(x)}{v_j(x)}\right) < 1; \tag{1}$$

and we write $v_i \prec v_j$ to indicate this [4].

We first observe that, if the TSS workers are sequenced from slowest to fastest, then the fixed point is unique. In contrast, when workers are not sequenced from slowest to fastest, then there can be multiple fixed points.

[4] This condition is slightly stronger than the simpler requirement that $v_i(x) < v_j(x)$ but is necessary to avoid technical pathologies in which the velocity of one worker approaches that of the other in the limit.

Lemma 1 *If the workers are sequenced from slowest to fastest, then the fixed point of the TSS line is unique.*

Proof Appendix, section A.3. □

Our main result is that, if the workers are sequenced from slowest to fastest, then there exists a unique fixed point to which all orbits converge, so that the behavior of the line is independent of the starting positions of the workers. In addition, if there is no blocking, then each worker invests the same clock time in each item produced.

It makes sense to put slower workers before faster ones to avoid blocking (a local phenomenon); what may be surprising is that this is sufficient to elicit *global* organization: the spontaneous emergence of balance.

Theorem 2 *For any TSS line, if the workers are sequenced from slowest to fastest, then any orbit of worker positions* $\left\{\mathbf{x}^{(t)} = f^t\left(\mathbf{x}^{(0)}\right)\right\}$ *converges to the unique fixed point.*

Proof Appendix, section A.4. □

An interesting special case of our model illustrates the essential analysis (and probably provides an adequate description of most real TSS lines). Assume that each worker i has a velocity that is constant over the entire unit of work, so that $v_i(x) = v_i$. Since all velocities are constant, there are no advantages to specialization of labor and so the largest possible production rate under any organization of the workers is $\sum_{i=1}^{n} v_i$ items per unit time. If in addition no episode of blocking is "too long"—such as when $p_{\max}$ is sufficiently small—then the dynamics of the TSS line become linear and the following stronger results hold.

Theorem 3 *If worker velocities are constant with* $v_1 < \cdots < v_n$ *and if workers are never blocked, then the TSS line converges exponentially fast to a unique fixed point at which*

1. *Worker i repeatedly executes the interval of work content*

$$\left[\frac{\sum_{j=1}^{i-1} v_j}{\sum_{j=1}^{n} v_j}, \frac{\sum_{j=1}^{i} v_j}{\sum_{j=1}^{n} v_j}\right].$$

2. *The production rate is $\sum_{j=1}^{n} v_j$, the largest possible.*

Proof Because no episode of blocking is appreciably long, the worker positions change from one iteration to the next as follows:

$$\begin{cases} x_1^{(t+1)} &= 0; & \text{and} \\ x_i^{(t+1)} &= x_{i-1}^{(t)} + v_{i-1}\left(\frac{1-x_n^{(t)}}{v_n}\right) & i = 2, \ldots, n, \end{cases}$$

from which it follows by simple algebra that for $i = 2, \ldots, n$,

$$\begin{aligned} x_{i+1}^{(t+1)} - x_i^{(t+1)} &= x_i^{(t)} - x_{i-1}^{(t)} + (v_i - v_{i-1})\left(\frac{1-x_n^{(t)}}{v_n}\right) \\ \frac{x_{i+1}^{(t+1)} - x_i^{(t+1)}}{v_i} &= \frac{x_i^{(t)} - x_{i-1}^{(t)}}{v_i} + \left(1 - \frac{v_{i-1}}{v_i}\right)\left(\frac{1-x_n^{(t)}}{v_n}\right) \\ \frac{x_{i+1}^{(t+1)} - x_i^{(t+1)}}{v_i} &= \left(\frac{x_i^{(t)} - x_{i-1}^{(t)}}{v_{i-1}}\right)\left(\frac{v_{i-1}}{v_i}\right) + \left(1 - \frac{v_{i-1}}{v_i}\right)\left(\frac{1-x_n^{(t)}}{v_n}\right) \end{aligned}$$

Letting

$$a_i^{(t+1)} = \frac{x_{i+1}^{(t+1)} - x_i^{(t+1)}}{v_i},$$

which may be interpreted as the clock time separating workers i and $i+1$ at the start of iteration $t+1$, we can summarize the dynamics of the line as follows:

$$\begin{cases} a_1^{(t+1)} &= a_n^{(t)}; & \text{and} \\ a_i^{(t+1)} &= \left(\frac{v_{i-1}}{v_i}\right) a_{i-1}^{(t)} + \left(1 - \frac{v_{i-1}}{v_i}\right) a_n^{(t)} & \text{for } i = 2, \ldots, n. \end{cases}$$

Rewriting these equations as a linear system

$$\mathbf{a}^{(t+1)} = T\mathbf{a}^{(t)},$$

we observe that this sequence of iterates converges because T is the transition matrix of a finite state Markov chain that is irreducible and aperiodic (Resnick, 1992). After solving for the limiting "probabilities" $\pi_j = v_j / \sum_k v_k$, we get that the i-th component of the limit point $\mathbf{a}^*$ is $a_i^* = \sum \pi_j a_j^{(0)} = 1/\sum_k v_k$. The other claims follow by simple algebra. □

This may be interpreted as showing that, to configure a bucket brigade from well to fire, one should put the fastest people close to the fire; then the people will, without intention, space themselves to convey the greatest possible flow of water upon the fire. The system optimizes itself.

The details of this proof illustrate the general ideas of the full argument. The main trick is to simplify analysis by looking, not directly at the positions of the workers, but at the clock time $a_i^{(t)}$ required for each worker i to reach the position of her successor. We refer to this time as the *allocation* of work suggested by the partition $\mathbf{x}^{(t)}$. (Allocations are defined for the general model in the appendix, section A.2.) When workers are sequenced from slowest to fastest, the largest allocation—which would be the cycle time of the line if the allocations suggested by the current $\mathbf{x}^{(t)}$ were fixed—converges from above and is guaranteed to have decreased after each completion of n items. Thus the production rate increases to a limit whose value is independent of the starting positions of the workers.

Figures 4 and 5 show the convergence of a system from two complementary points of view. Figure 4 shows an example of how the movement of the workers stabilizes, with the faster workers eventually allocated more work; and Figure 5 shows the convergence of the system within the state space of worker positions. These simulations were generated by three workers of constant velocities $\mathbf{v} = (1, 2, 3)$.

5 Complicated behavior

In computational experiments with workers sequenced other than from slowest to fastest, (our model of) a TSS line can fail to balance itself. By Theorem 1

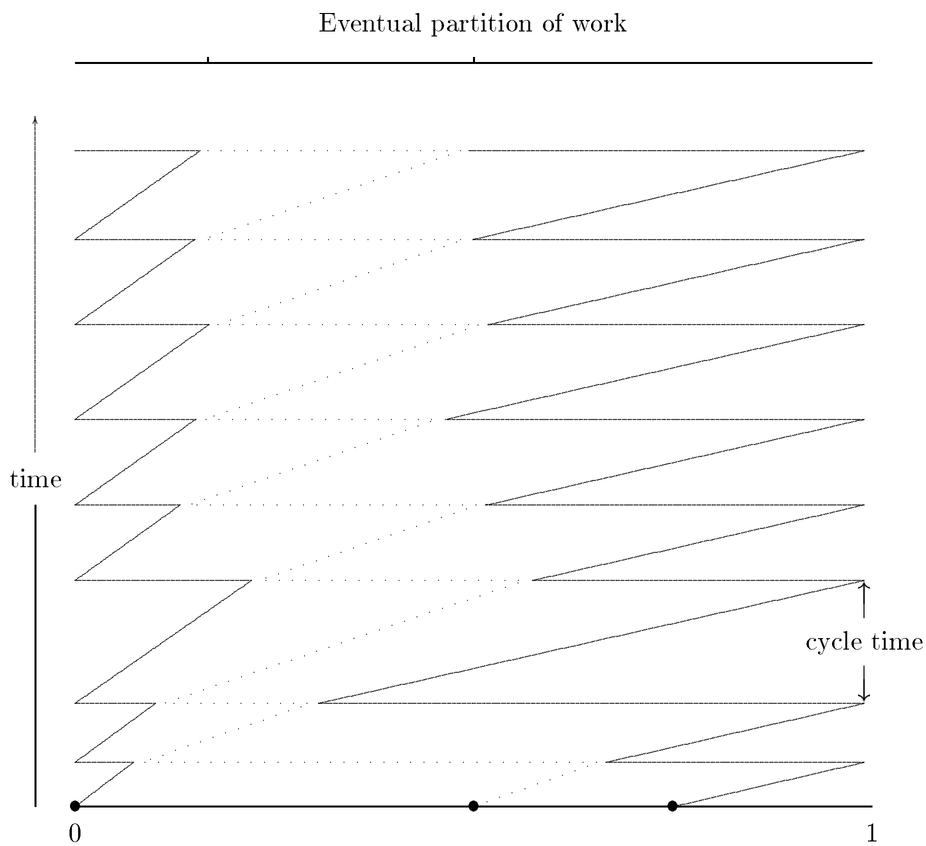

Figure 4: A time-expanded view of a TSS production line with three workers sequenced from slowest to fastest. The solid horizontal line represents the total work content of the product and the solid circles represent the initial positions of the workers. The zigzag vertical lines show how these positions change over time and the rightmost spikes correspond to completed items. The system quickly stabilized so that each worker repeatedly executes the same portion of work content of the product.

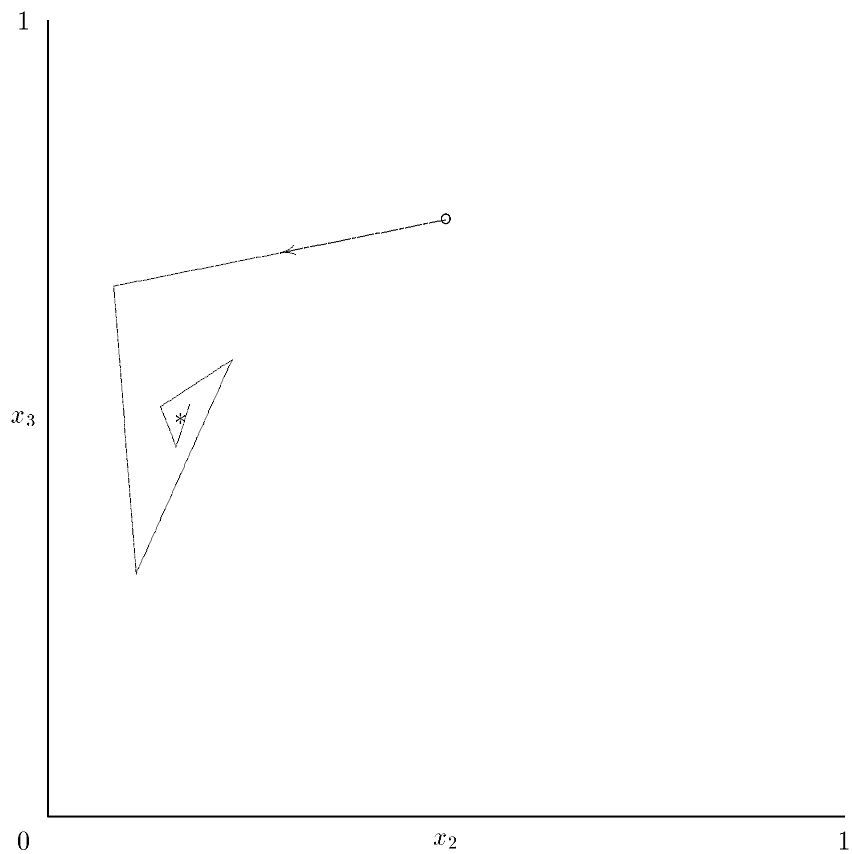

Figure 5: The positions of the workers on the production line at successive instants when it resets. (Since the position of the first worker is always 0 when the line resets, only (x_2, x_3), the positions of the second and third workers, are plotted here.) From any initial position the system converges to the fixed point $(1/6, 1/2)$. The production rate also converges to a unique value—in this instance, 6, the maximum possible—that is independent of where the workers start on the line.

there always exists a balance point so that workers always reset to the same position; the trouble is that the fixed point can be a *repeller*, so that if the system ever deviates, however slightly, from that point, then the system must inexorably diverge from it (Devaney, 1989). Typically the line becomes trapped in periodic behavior: It "sputters", producing erratically and at suboptimal production rate.

Figure 6 shows a simulation, generated by workers of constant velocities $\mathbf{v} = (3, 1, 2)$, in which the fixed point is a repeller and any orbit that strays must eventually be trapped by a limit cycle with production rate less than that at the fixed point. In the limit cycle a faster worker is repeatedly blocked by a slower worker, with consequent waste of productive capacity. In other simulations we have found instances of quite large cycles, some at the limits of the numerical resolution of our computer and of the patience of the observers. Our model is also capable of "quasi-periodic" behavior, which means that it is predictable but not periodic. For example, with constant velocities $\mathbf{v} = (2, 1, 2)$, all orbits converge to the periphery of an ellipse, the center of which is a fixed point—but worker positions never repeat.

For a fixed set and sequence of workers our simulations have shown such phenomena as multiple fixed points, both attractors and repellers, multiple limit cycles, and long-term behavior that depends on the starting positions of the workers. This suggests that, if workers are not sequenced from slowest to fastest, there can be a structural tendency toward persistent imbalance in a TSS line. This could be a practical problem if the imbalance is significant. (In a companion paper we have catalogued all possible asymptotic behavior of 2- and 3-worker lines and interpreted its significance for practice (Bartholdi, Bunimovich, and Eisenstein, 1995a).)

More troubling than complicated behavior is anomalous behavior. The simplest manifestation of this is that adding a worker to the line can *decrease* the production rate if the slowest-to-fastest sequence is not respected. For example,

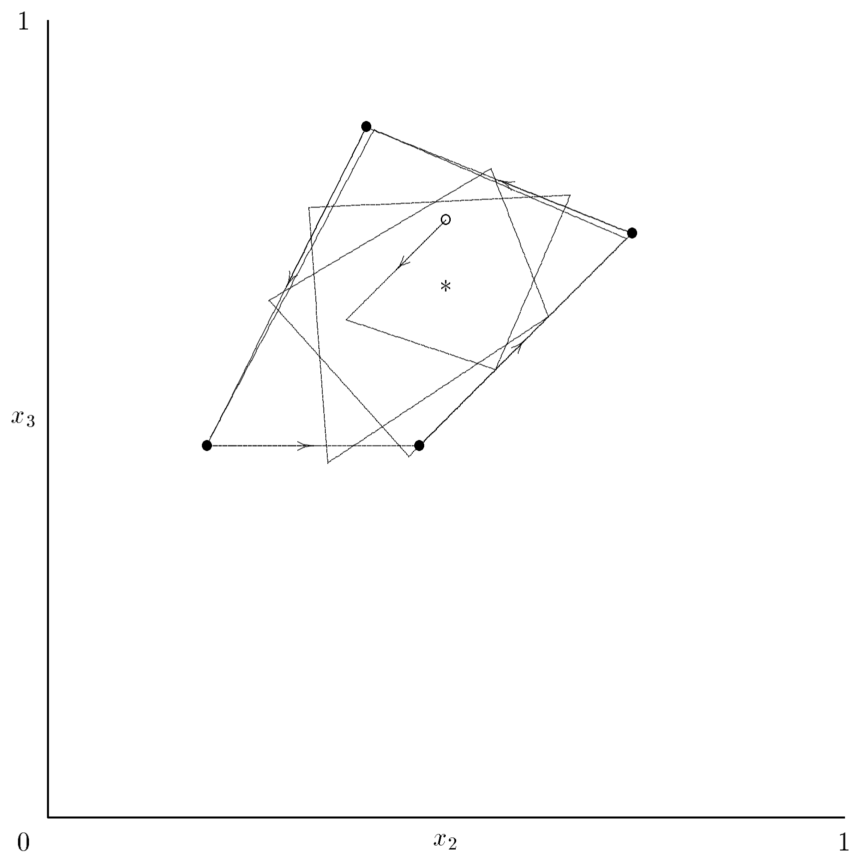

Figure 6: The positions of the workers on the production line at successive instants when it resets. In this instance the fixed point $(1/2, 2/3)$, indicated by $*$, achieves the largest possible rate of production, but is a repeller, and any orbit that strays from it will be trapped by the attracting but suboptimal limit cycle consisting of the points $(3/15, 7/15)$, $(7/15, 7/15)$, $(11/15, 11/15)$, and $(2/5, 13/15)$, indicated by $\bullet$'s. As the system is trapped by the limit cycle, the cycle time of the line oscillates and the average production rate converges to $60/11$, which is less than the optimum value of 6.

consider an m station line with processing times $\mathbf{p} = (1/m, \ldots, 1/m)$ staffed by a single worker of velocity k, where $1 < k \ll m$. The production rate of this line will plummet from k items per time unit to only a little more than 2 items per time unit if a worker of velocity 1 is added to the end of the line, where she repeatedly blocks the faster worker. Thus one can induce an arbitrarily large gap between production capacity and realized production rate.

It is less obvious that increasing the velocity of a worker can decrease the production rate. For example consider a TSS line with processing times $\mathbf{p} = (1/2, 1/4, 1/4)$. On this line workers of constant velocity $\mathbf{v} = (2, 1, 1)$ will achieve a production rate of 4 items per unit time; but if worker 3 doubles her velocity, the first worker will always be blocked and the production rate will decrease from 4 to 8/3. Therefore, increasing production capacity by 25% causes a 33% *decrease* in realized production rate!

We emphasize that system behavior can be neither complicated nor anomalous when workers are sequenced from slowest to fastest.

Theorem 4 *If workers on a TSS line are maintained in sequence from slowest to fastest, then adding or speeding up a worker will never decrease the production rate.*

Proof Appendix, section A.3. □

6 The production rate of TSS lines

Here we demonstrate that the logic of TSS cannot by itself guarantee the best production rate if there is a pathological mismatch between the sequence of workers and the assignment of work content to stations. Nevertheless a TSS line in which workers are sequenced from slowest to fastest will always achieve a production rate that is good in the following sense: Other sequences can perform much worse but not too much better.

The following example shows that a sequence other than slowest-to-fastest can be arbitrarily less productive than the slowest-to-fastest sequence of workers. Consider a TSS line with $\mathbf{p} = (\epsilon, 1-\epsilon)$ and two workers, one of constant velocity ϵ and the other of constant velocity $1-\epsilon$. In this sequence the workers achieve a production rate of one item per time unit; but reversing the sequence gives a production rate of $\epsilon/(1-\epsilon)$ items per time unit, which can be made arbitrarily small. Thus the worst-case ratio of production rates is unbounded above.

On the other hand, a sequence of workers other than slowest-to-fastest cannot achieve a production rate that is "too much" better than that of the slowest-to-fastest sequence. More specifically, when the fastest worker is last, the production rate of a TSS line at its fixed point is always within a factor n of the best achievable by any other sequence of the workers. This follows because the fastest, last worker is never blocked. One can construct examples that achieve this bound asymptotically.

We emphasize that, although this worst-case behavior is possible within our model, it is not a practical problem because neither persistent blocking nor pathological mismatch of workers to stations will be found. First, persistent blocking is not tolerated. If some station is recognized as a bottleneck, then either workers are removed from the line or else the station is duplicated (and all of our results can be shown to hold for a line with parallel stations). Second, a pathological mismatch of workers to stations is generally not possible where, as in the apparel industry, most tasks are variations of a single skill such as simple dexterity. Theorem 3 therefore suggests that real TSS lines will achieve a (nearly) maximum rate of production if the workers are sequenced from slowest to fastest.

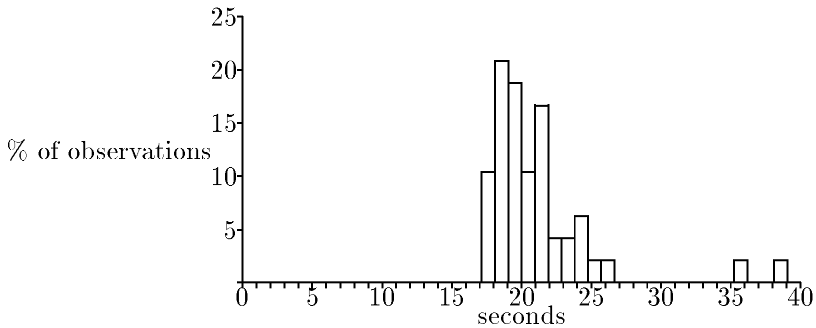

Figure 7: The distribution of actual processing times for a single worker at a TSS station. The small bump to the right is due to regular pauses to position raw materials.

7 Validating the model

The details of our model are confirmed by evidence, both anecdotal and experimental, and by industry practice. More importantly, the predictions of our model are confirmed by the operation of real lines, in both the laboratory and at commercial sites.

Here are the distinctive features of our model and its predictions.

7.1 Deterministic processing times

We timed several workers over hundreds of task executions on the shop floor at a commercial site. Figure 7 shows the distribution of actual times for a specific worker to complete a specific task. It is representative of all the workers we timed and it is consistent with measurements by factory personnel.

In analyzing actual processing times measured on the shop floor, we discerned two sources of variance in the execution of a given task by a given worker. The first source was the inevitable small "noise". This was generally because small variations in the cut cloth resulted in small variations in the time to position the cloth under the needle. While this component of task time seems properly described as "random", its variance was insignificant. (It is reflected in the first, major peak in the distribution of task times in Figure 7.)

The second source of variance was that due to occasional interruptions to which sewing is subject (and which form the second, small peak in Figure 7). Most of the long delays were *not* randomly occurring: Instead, these were the regular pauses to position new bundles of raw materials. Significant "random" interruptions, such as a dropped garment, a quality problem, or a thread break, occurred much less frequently.

We judged the first type of variance to be insignificant because, ignoring documented pauses to position raw materials, the coefficient of variation of the processing time was less than 10% for an average worker and task. This is to be expected because the workers perform the same set of tasks several hundred times each day and so become quite consistent at it. Accordingly we modeled processing times as deterministic. Interruptions are not intrinsic to our model; instead, we consider them to be extraneous shocks to the system. Our model explains how the dynamics compensate for such shocks to reëstablish balance.

(We have found contexts in which a stochastic model seems appropriate. For example, Bartholdi, Bunimovitch, and Eisenstein (1995b) analyze TSS-style picking from a flowrack in a warehouse supporting retail operations. In this case the composition of orders to be picked can vary significantly and seems best modeled as a random variable. Suffice it to say here that under mild assumptions the vector of worker positions converges to a random variable, the distribution of which is independent of the starting positions of the workers.)

7.2 Distinguishable workers

Workers vary significantly in speed, as shown in Figure 8, which gives the average velocities of sixty-one apparel workers at a commercial site. Furthermore, this large variation in worker speeds persists because the turnover rate among employees in the US apparel industry is over 40% per year, with many new

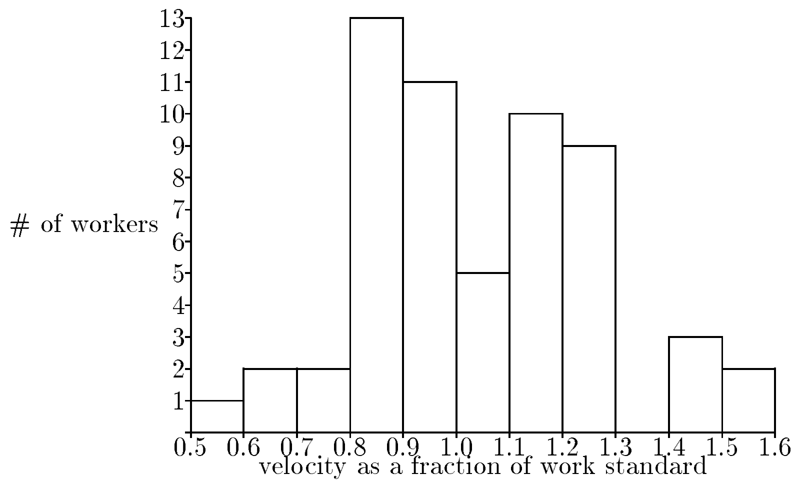

Figure 8: Distribution of average velocities of sixty-one workers.

employees entering the workforce ("Apparel Plant Wages Survey", Human Resources Committee of the American Apparel Manufacturers Association, 1993).

7.3 Well-defined ranking

The usefulness of our results depends on whether workers can be ranked by velocity. All the people we interviewed, both managers and workers, said this could be done. Furthermore, they suggested that people who had worked together would generally agree on the ranking. We tested this at the Apparel Manufacturing Technology Center of the Southern College of Technology. We asked three experienced workers ("A", "B", and "C") and a fellow-worker to rank each of A, B, and C by speed. These rankings were gathered by secret ballot and are as follows:

$$\begin{array}{ccccccc} A & < & B & < & C & & \\ A & < & B & < & C & & \\ A & = & B & < & C & & \\ & & B & < & C & = & A, \end{array}$$

where the last ballot was cast by A. The near unanimity of the rankings supports our assumption that workers can in fact be sequenced from slowest to fastest. (This probably reflects the fact that work in the apparel industry involves mostly variations on a single skill, sewing. Such agreement may be less likely on production lines that require a mix of quite different skills.)

In further support of our model, an employment test that is widely-used in the apparel industry is based exactly on our hypothesis: that workers can be ranked according to a single measure that will predict their productivity (Trego, 1981 and 1989). Furthermore, this test is required by federal statute to be statistically significant at the 0.05 level (Volume 29 of the Code of Federal Regulations (7-1-93 Edition), Chapter XIV "Equal Opportunity Employment Commission", §1607.14.B "Uniform Guidelines on Employee Selection Procedures; Technical Standards for Validity Studies"). Therefore, either our model is correct that workers can be ranked or else many factory managers are in potential violation of the law.

7.4 Predictions of the model

The most important confirmation of our model was provided by comparing its predictions to observed behavior. We ran a TSS line with workers A and C from section 7.3. The line manufactured two batches of six items each, with both workers beginning each run at position $(0, 0)$ (the beginning of the line). As shown in Figure 9, when the workers were sequenced from slower to faster the clocktime contributed by the last worker to successive items appeared to be converging. Furthermore, it is evident that the contribution of the last worker was on average smaller when the workers were sequenced from slower to faster; and because this contribution is also the time between completions of successive items, one can see that the slower-to-faster line had a higher average rate of production.

The correctness of our model was further confirmed by visits to industrial

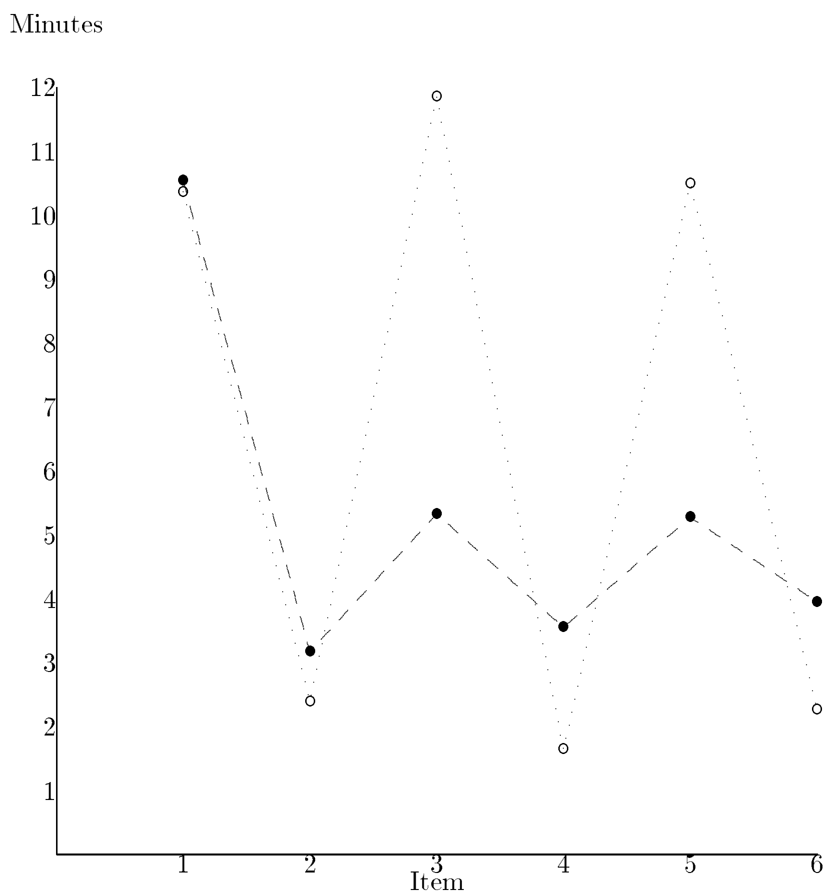

Figure 9: Clocktime invested in each item by the last worker for each of two production runs. When two workers were sequenced from slower to faster (marked by •), the contribution of the last worker was apparently converging as predicted by our model. When the workers were sequenced from faster to slower (marked by ∘), no convergence was evident and the production rate slowed perceptibly.

sites. Each site we visited had given some thought to sequencing the workers. Though none sequenced workers from slowest to fastest, two sites had nearly reached that conclusion by trial and error. At one site management initially put the oldest, slowest workers at the last position in each line, thinking the work there would be less strenuous because it was mostly inspecting and packaging finished items. When this resulted in a substandard rate of production management reversed their initial policy and began placing the fastest worker at the end of the line and observed an immediate and significant increase in production rate.

At another site management initially assigned any new, slow worker to the middle of the line, thinking that, with a faster worker on either side, the new worker would be forced to improve quickly. This idea was subsequently abandoned in favor of assigning new workers to the first position to avoid blocking more experienced, faster workers.

Sometimes we found that workers were sequenced by criteria other than speed. For example, on one unusual TSS line there was a station that required special skills and the single qualified worker had to be assigned to that station. We also found, to our consternation and delight, a line in which the workers were organized according to the principle tallest-to-shortest! This came about by special circumstance. The team included both a very tall worker and a very short worker; and the heights of the tables at which they usually sewed were set accordingly. Management wanted to avoid a short person trying to sew at a high table.

8 Related work

The first paper on TSS was apparently that of Schroer, Wang, and Ziemke (1991), who built a simulation model of a particular TSS line they observed at a trade show. Because the point of their work was to demonstrate capabilities in object-oriented simulation, they gathered statistics on the single instance they simulated but did not pursue analysis of the TSS system and reached no general conclusions about TSS lines.

Unfortunately, their paper contains some inaccuracies that have mislead subsequent researchers. In particular, it is mistaken in its description of a TSS line (private communication, Len Egan, President, Americas 21st, Inc.). Contrary to the description of Schroer et al. there are *no* buffers for work-in-process inventory. It is the firm opinion of TSS practitioners that this is disadvantageous.

Our model of TSS is different from others in two main respects. Most have followed Schroer et al. (1991) in assuming that all workers are identical but that task times are random (Bischak, 1993; Zavadlav, McClain, and Thomas, 1994). The assumption of identical workers fails to hold in our experience, as described in section 7.2. The assumption of stochastic processing times may be a plausible modeling decision, but in our experience the variance of task times was small, certainly much smaller than the variance of velocities among workers.

Assuming deterministic processing times is not only arguable from real data as explained in section 7.1 but it also confers a clarity to the model. Complex modes of behavior, such as very long limit cycles, become more readily apparent while in a stochastic model they may be hidden by the randomness. For example, by assuming deterministic processing times we can gain new insight into the models of Schroer et al. (1991), Bischak (1993), and Zavadlav et al. (1994). First we observe that when all workers are identical a TSS line is indistinguishable from one in which each worker circles back upon completion of an item to start a new item; and so the line is equivalent to a cyclic queue (Bischak, 1993). Choosing dimensions so that all $v_i = 1$, one can show that in a line with n workers the production rate converges to the maximum possible, $\min\{n, 1/p_{\max}\}$ items per time unit (Bartholdi, Eisenstein, Jacobs-Blecha, and Ratliff, 1995). In this case the line does balance itself, but only on the average, in the sense of Zavadlav et al. (1994). Workers will in general be required to perform different tasks on successive items so the assignment of work will not be stable. The problem is that the system converges to a region with an infinite number of n-cycles, each achieving the same (optimal) production rate, and each neutral (neither attracting nor repelling). Furthermore, the eventual limit cycle is determined by the starting positions of the workers. If processing times are allowed to be stochastic then the system simply jumps among limit cycles randomly, which is why it is hard to see the structure of behavior in a simulated line with identical workers and random processing times.

To avoid confusion it is worth pointing out that our definition of balance is stricter than that of Zavadlav et al. (1994) and Ostolaza, McClain, and Thomas (1991). For us a line is balanced when a *stable* partition of work has emerged, so that each worker performs the same portion of work content from item to item. In contrast, the balance of Zavadlav et al. (1994) and Ostolaza et al. (1991) shares work only *on average* over all items. This distinction can matter in practice because a line that is balanced only on average can require the workers to move all about, with no persistent assignment of work. This means the workers must be trained at more or all stations and are therefore likely to have lower velocity. The net effect is a lower realized production rate than if workers were able to sustain a stable sharing of work.

Another possible source of confusion is that both Zavadlav et al. (1994) and the advertising literature for TSS use the term "self-balance" to mean *local* adjustments between adjacent workers. However, they produce no guarantee that such local adjustments cumulatively lead to *global* balance among all the workers of the line. It is exactly this guarantee that is our concern here: the spontaneous emergence of global organization from local interaction.

9 Conclusions

When a production line is laid out, tasks are first assigned to stations, which results in a partition of work among stations that is static, unchanging, and generally imperfect. Then if workers, sequenced from slowest to fastest, follow the TSS rule, a second partition of work emerges, this time among the workers. This second allocation arises spontaneously and, because it is self-adjusting, it can, without management intervention, smooth over imperfections in the underlying static partition. Furthermore, the partition can adapt; for example, when a worker takes a break, the work content will be spontaneously reallocated among the remaining workers.

Our model suggests that a TSS line is easy to manage. The production rate can be fine-tuned by adjusting the number of workers; and because the line does not exhibit anomalous behavior, adding workers never reduces the production rate and removing workers never increases it. In addition, the line is parsimonious in its data requirements, which are only the relative speeds of the workers (not even their values); and it does not require knowledge of task times, and thus might reduce the expense of time-motion studies.

The main weakness of our model is that it treats workers simplistically in describing each as merely a velocity function $v_i(x)$. This fails to capture one of the key features of TSS, which is the emphasis on teamwork[5]. For example, skilled TSS workers may accelerate in spurts to smooth the production rate when required. Nevertheless we believe that our model correctly describes the qualitative behavior of real TSS lines even if it might not predict the exact positions of the workers over time.

Finally we observe that TSS may be seen as part of a more general approach that we call "bucket brigade manufacturing". As in many types of work cells,

[5] We have also ignored other issues that are important to the effectiveness of actual implementations, such as the assignment of work content to stations, the detailed choreography of worker movement, and the strategic training and motivation of workers.

there are fewer workers than stations; but the distinctive feature of the bucket-brigade is that the workers maintain their sequence on the line while sharing stations. It would be interesting to explore this style of manufacturing in other environments.

Acknowledgements

We thank Leonid Bunimovich and Robert Foley for helpful technical discussions; Michael Trick and John Vande Vate for comments on earlier drafts; Len Egan and Frank Guest of Americas 21st, Inc., Ken Williams of Riverside Fashions, Inc., David Zimmerman of The Coach Factory, Mitchell Johnson and Tom Tanner of Champion Products, Inc. for showing us TSS in operation; Charlotte Jacobs-Blecha of the Georgia Tech Research Institute and Carolyn Ring of the Apparel Manufacturing Technology Center at the Southern College of Technology for help in conducting experiments on a TSS line. We also appreciate support provided by the Defense Logistics Agency, grant #DLA900-87-D-0018-0010; the Office of Naval Research, grant #N00014-89-J-1571 (Bartholdi); IBM (Eisenstein); the Graduate School of Business at the University of Chicago (Eisenstein); and the National Science Foundation, grant #DDM-9215564 (Bartholdi and Eisenstein).

Portions of this work have been presented at the ORSA/TIMS Joint National Meeting, Anaheim, November 1991; Purdue University, January 1992; Carnegie Mellon University, January 1992; Auburn University, February 1992; University of Chicago, February 1992; Columbia University, March 1992; Massachusetts Institute of Technology, March 1992; Georgia Institute of Technology, March 1992; Universitè Laval, April 1992; Dynamics Days, Phoenix, January 1993; the TIMS/ORSA Joint National Meeting, Chicago, May 1993; Cornell University, October 1993; and the University of Michigan, March 1994. We thank the audiences for many stimulating questions.

References

[1] J. J. Bartholdi, III, L. Bunimovich, D. D. Eisenstein. "Dynamics of 2- and 3-worker bucket brigade production lines", submitted (1995a).

[2] J. J. Bartholdi, III, L. Bunimovich, D. D. Eisenstein. "Bucket brigades: A self-balancing order-picking system for a warehouse", forthcoming (1995b).

[3] J. J. Bartholdi, III, D. D. Eisenstein, C. Jacobs-Blecha, H. D. Ratliff. "Design of bucket brigade production lines", submitted (1995).

[4] D. Bischak. "Performance of a manufacturing module with moving workers", to appear in *IIE Transactions on IE Research* (1993).

[5] B. Bollobás. *Linear Analysis*, Cambridge University Press (1990).

[6] R. L. Devaney. *An Introduction to Chaotic Dynamical Systems*, 2nd ed., Addison-Wesley, Reading, Mass. (1989).

[7] F. Morrison. *The Art of Modelling Dynamic Systems: Forecasting for Chaos, Randomness, and Determinism*, John Wiley & Sons, New York (1991).

[8] J. Ostolaza, J. McClain, and J. Thomas. "The use of dynamic (state-dependent) assembly line balancing to improve throughput", *J. Mfg. Oper. Mgt.* **3**:105–133 (1990).

[9] S. I. Resnick. *Adventures in Stochastic Processes*, Birkhäuser Boston (1992).

[10] B. J. Schroer, J. Wang, and M. C. Ziemke. "A look at TSS through simulation", *Bobbin* magazine, July, 114–119 (1991).

[11] R. Trego. "Skil-Sew: A predictor of operator productivity", *Bobbin Magazine* 138–144 (March 1981).

[12] R. Trego. "Skil-Sew and flexible manufacturing", *Apparel Manufacturer* 73–78 (August 1989).

[13] E. Zavadlav, J. McClain, J. Thomas, "Self-buffering, self-balancing, self-flushing production lines", manuscript, Graduate School of Business, Cornell University, Ithaca, NY 14853 (1994).

A Technical details

Because of excessive length many of the proofs have been summarized or omitted. However, all of the proofs were refereed and are available from the authors in their entirety.

A.1 Existence of a fixed point

Proof sketch Extend f so that its domain is the entire closed n-cell defined by $0 \leq x_1 \leq \cdots \leq x_n \leq 1$. (The TSS rule does this implicitly by implying that, if more than one worker is at a station, then that worker closest to completion has priority and the others must wait.) One can now show that the extended TSS function g is continuous on its domain and therefore, by Brouwer's Fixed Point Theorem, has a fixed point (Bollobás, 1990). Furthermore, this fixed point must lie within the natural domain of f because, by the logic of the TSS rule, no point in the extended domain can remain fixed under g. Therefore, since f agrees with g on the domain of f, the fixed point with respect to g must also be fixed with respect to f. $\square$

A.2 Technical preliminaries

In the same way that some physics problems become simpler under a suitable change of coordinates, it is easier to analyze the behavior of a TSS line if we introduce a new coordinate system to keep track not just of the positions of the workers but also of their *relative* positions. Before formalizing this, we must introduce some notation.

Define $p_0 = 0$, and let $P_k = \sum_{j=0}^{k} p_j$ be the cumulative amount of work invested in an item when it has just completed processing at station k. The work at station k corresponds to the open interval (P_{k-1}, P_k); and, because workers cannot use the same station at the same time, no two x_i's can assume values within the same interval (P_{k-1}, P_k).

Sometimes we will need to know the endpoints of the station (interval) containing position x; accordingly we define

$$\underline{x} = P_{k-1} \quad \text{if } x \in [P_{k-1}, P_k);$$

and

$$\overline{x} = P_k \quad \text{if } x \in (P_{k-1}, P_k].$$

Define $\tau_i(x, x')$ to be the time it would take for worker i, if not blocked, to travel from position x to position x', where $0 \leq x \leq x' \leq 1$. By the properties of v_i, the function $\tau_i(x, x')$ is well-defined and may be expressed as

$$\tau_i(x, x') = \int_x^{x'} \frac{dz}{v_i(z)}.$$

Note that $\tau_i(x, x')$ is strictly increasing in x', strictly decreasing in x, and continuous in both.

Now we can introduce our new coordinate system. The vector $\mathbf{x}^{(t)}$ can be interpreted as suggesting a partition of the work, with the interval $\left[x_i^{(t)}, x_{i+1}^{(t)}\right]$ assigned to worker i during iteration t. The *allocation* $a_i^{(t)}$ is the clock time that would be required for worker i to complete her suggested share of work, including both work time and possible delays due to blocking. More precisely,

$$\left\{ \begin{array}{lcl} a_n^{(t)} & = & \tau_n\left(x_n^{(t)}, 1\right);\ \text{else} \\ a_i^{(t)} & = & \tau_i\left(x_i^{(t)}, x_{i+1}^{(t)}\right) + \max\left\{0,\ \tau_{i+1}\left(x_{i+1}^{(t)}, \overline{x_{i+1}^{(t)}}\right) - \tau_i\left(x_i^{(t)}, \underline{x_{i+1}^{(t)}}\right)\right\}. \end{array} \right. \tag{2}$$

The preceding expression consists of two terms: the first is the travel time required for worker i to reach the starting position $x_{i+1}^{(t)}$ of her successor; and the second term is the delay if worker i is blocked en route, in which case i reaches $i+1$'s station at time $\tau_i\left(x_i^{(t)}, \underline{x_{i+1}^{(t)}}\right)$, but $i+1$ does not finish at that station until time $\tau_{i+1}\left(x_{i+1}^{(t)}, \overline{x_{i+1}^{(t)}}\right)$.

We distinguish between two types of allocations:

- $a_i^{(t)}$ is a *simple* allocation if $a_i^{(t)} = \tau_i\left(x_i^{(t)}, x_{i+1}^{(t)}\right)$, so that the allocation is pure work time;

- $a_i^{(t)}$ is a *delay allocation* if $a_i^{(t)} = \tau_i\left(\underline{x_{i+1}^{(t)}}, x_{i+1}^{(t)}\right) + \tau_{i+1}\left(x_{i+1}^{(t)}, \overline{x_{i+1}^{(t)}}\right)$, so that the allocation includes both work time and idle time spent waiting for an occupied station.

Note that all the time of a simple allocation is productive while a delay allocation includes time during which the worker is blocked and therefore idle.

Finally we state without proof some simple results about allocations.

- Allocation $a_i^{(t)}$ is continuous in $x_i^{(t)}$ and continuous in $x_{i+1}^{(t)}$ everywhere except possibly at P_k $(k = 1, \ldots, m-1)$.

- When workers are ordered from slowest to fastest then $a_i^{(t)}$ is non-increasing in $x_i^{(t)}$ and strictly increasing in $x_{i+1}^{(t)}$.

- The allocation $a_n^{(t)}$ of the last worker is always a simple allocation; and it represents the time between completions of the t-th and $(t+1)$-st items.

- At a fixed point $\mathbf{a}^*$, $a_i^* \leq a_n^*$, with strict inequality only if worker $i+1$ is blocked at the beginning of the iteration.

Now we can study the evolution of a TSS line given either by its orbit $\left\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \ldots\right\}$ of worker positions or by its orbit $\left\{\mathbf{a}^{(0)}, \mathbf{a}^{(1)}, \ldots\right\}$ of corresponding allocations.

A.3 Uniqueness of fixed point and freedom from anomalies

This section is based on the definitions and notation of section A.2.

We will need the following result, which says that when workers are sequenced from slowest to fastest then a slower team of workers cannot sustain a faster production rate.

Lemma 2 *Let $v_1 \prec \cdots \prec v_n$ and $v_1' \prec \cdots \prec v_n'$ be two teams of workers for a given TSS line, with $v_i \leq v_i'$. Let $\mathbf{a}$ and $\mathbf{a}'$ be vectors of allocations that are fixed points for the respective teams. Then $a_n \geq a_n'$.*

Proof Assume the opposite, that $a_n < a'_n$, so that $x_n > x'_n$. We claim this implies that $x_i > x'_i$ for all i, which yields the contradiction that $0 = x_0 > x'_0 = 0$.

If, given the hypothesis, there exists a TSS line and two teams for which $x_i > x'_i$ fails, let j be the largest index for which $x_j \leq x'_j \leq x'_{j+1} < x_{j+1}$. Worker $j+1$ of the faster team must be blocked at the beginning of each iteration because $a'_j < a_j \leq a_n < a'_n$. Therefore it must be that $x'_{j+1} = P_{k-1}$ for some station k and some worker $j+c$ has position $P_{k-1} < x'_{j+c} < P_k$. Consequently worker $j+c$ of the faster team is never blocked at the start of an iteration and so $a'_{j+c-1} = a'_n$.

If $x'_{j+c} < x_{j+1}$ then $x_j \leq x'_{j+c-1} < x'_{j+c} < x_{j+1}$ and so $a'_{j+c-1} < a_j \leq a_n < a'_n$, which contradicts $a'_{j+c-1} = a'_n$. Assume, on the other hand, $x'_{j+c} \geq x_{j+1}$; then a'_{j+c-1} is strictly smaller than the time required for worker $j+c-1$ of the faster team to travel from P_{k-1} to P_k, which is no larger than a_j. But then $a'_{j+c-1} < a_j \leq a_n < a'_n$, again contradicting $a'_{j+c-1} = a'_n$. □

Now we can prove Lemma 1, which claims uniqueness of the fixed point:

Proof Assume there exist two distinct fixed points $\mathbf{x}$ and $\mathbf{x}'$ and corresponding allocations $\mathbf{a}$ and $\mathbf{a}'$. By Lemma 2 it must be that $a_n = a'_n$. Let j be the first index for which $x_j \neq x'_j$ and assume without loss of generality that $x'_j < x_j$. Now the argument proceeds as in Lemma 2: Because allocation is strictly increasing in its second argument $a'_{j-1} < a_{j-1} \leq a'_n$ and so for the fixed point $\mathbf{x}'$ worker j must be blocked at the beginning of each iteration. Therefore it must be that $x'_j = P_{k-1}$ for some station k and some worker $j+c$ has position $P_{k-1} < x'_{j+c} < P_k$. Consequently worker $j+c$ is never blocked at the start of an iteration and so $a'_{j+c-1} = a'_n$.

If $x'_{j+c} < x_j$ then $x_{j-1} \leq x'_{j+c-1} < x'_{j+c} < x_j$ and so $a'_{j+c-1} < a_{j-1} \leq a'_n$, which contradicts $a'_{j+c-1} = a'_n$. Assume, on the other hand, $x'_{j+c-1} \geq x_j$; then a'_{j+c-1} is strictly smaller than the time required for worker $j+c-1$ to travel from P_{k-1} to P_k, which is no larger than a_{j-1}. But then $a'_{j+c-1} < a_{j-1} \leq a'_n$, again contradicting $a'_{j+c-1} = a'_n$. □

Similarly we can show Theorem 4, which claims freedom from anomalies:

Proof Let $\mathbf{a}$ be the fixed point for the original team. If the velocities of some workers are increased then by the preceding lemma a_n cannot increase and so the production rate $1/a_n$ cannot decrease. If a worker is added to the line, then we may compare the new team with the former team augmented by a worker of zero velocity. The latter has the fixed point $(0, a_1, \ldots, a_n)$ and so has the same production rate as the original team; and by Lemma 2 this cannot exceed the production rate of the new team. □

A.4 Convergence to fixed point

This section is based on the definitions and notation of section A.2.

Our main result is to establish Theorem 2, that when workers are ordered from slowest to fastest then there exists a fixed point $\mathbf{x}^*$ to which all orbits $\{\mathbf{x}^{(t)}\}_{t=0}^{\infty}$ converge. Our proof of convergence is a greatly elaborated version of that establishing Theorem 3. What complicates the argument is that no more than one worker can use a station at one time.

An outline of the argument is as follows.

1. Observe that each successive allocation for worker i is a mixture of allocations from the previous iteration.

2. Model this mixing as the action of a finite Markov chain for which the transition probabilities are not stationary, but have special structure.

3. Use the Markov chain to show convergence of $\{a_n^{(t)}\}_{t=0}^{\infty}$.

4. Show convergence of the $\{x_i^{(t)}\}_{t=0}^{\infty}$.

A.4.1 Evolution of the allocations

Let $\left\{\mathbf{x}^{(t)}\right\}_{t=0}^{\infty}$ be any orbit of worker positions and let $\left\{\mathbf{a}^{(t)}\right\}_{t=0}^{\infty}$ be the corresponding orbit of allocations.

Define

$$\rho = \max_i \sup_{x\in[0,1]} \left\{ \frac{v_i(x)}{v_{i+1}(x)} \right\}.$$

Recall that $\rho < 1$ because the workers are sequenced from slowest to fastest. As will be seen, the value of ρ will determine the rate at which the line achieves balance.

First we show that the line cannot become more imbalanced over time: More specifically, the largest allocation $a_{\max}^{(t)} = \max_i\{a_i^{(t)}\}$ is non-increasing. This will follow from the next two lemmata.

Lemma 3 *If $a_i^{(t+1)}$ is a simple allocation*

1. *and $x_i^{(t)} \leq x_i^{(t+1)}$, then $a_i^{(t+1)} \leq a_n^{(t)}$.*
2. *and $x_i^{(t)} > x_i^{(t+1)}$, then $a_i^{(t+1)} \leq a_{i-1}^{(t)}$.*

Proof If $x_i^{(t)} \leq x_i^{(t+1)}$, then

$$a_i^{(t+1)} = \tau_i\left(x_i^{(t+1)}, x_{i+1}^{(t+1)}\right) \leq \tau_i\left(x_i^{(t)}, x_{i+1}^{(t+1)}\right) \leq a_n^{(t)}.$$

If, on the other hand, $x_i^{(t+1)} < x_i^{(t)}$, then

$$\begin{aligned} a_i^{(t+1)} &= \tau_i\left(x_i^{(t+1)}, x_i^{(t)}\right) + \tau_i\left(x_i^{(t)}, x_{i+1}^{(t+1)}\right) \\ &\leq \tau_i\left(x_i^{(t+1)}, x_i^{(t)}\right) + a_n^{(t)} \\ &< \tau_{i-1}\left(x_i^{(t+1)}, x_i^{(t)}\right) + a_n^{(t)} \\ &\leq a_{i-1}^{(t)}. \end{aligned}$$

□

If $a_i^{(t+1)}$ is a delay allocation, then there must exist k such that

$$P_{k-1} = \underline{x_{i+1}^{(t+1)}} < x_{i+1}^{(t+1)} < \overline{x_{i+1}^{(t+1)}} = P_k;$$

and for this station k:

Lemma 4 *If $a_i^{(t+1)}$ is a delay allocation*

1. *and $x_i^{(t)} \leq P_{k-1}$*
   (a) *and $x_{i+1}^{(t)} \leq x_{i+1}^{(t+1)}$, then $a_i^{(t+1)} \leq a_n^{(t)}$.*
   (b) *and $x_{i+1}^{(t)} > x_{i+1}^{(t+1)}$, then $a_i^{(t+1)} \leq \rho a_i^{(t)} + (1-\rho)a_n^{(t)}$.*
2. *Otherwise $a_i^{(t+1)} \leq \rho a_{i-1}^{(t)} + (1-\rho)a_n^{(t)}$.*

Proof sketch The argument here is similar to that of Lemma 3 with the added complexity of handling delay allocations. □

In addition, we have the following, more explicit bound on $a_n^{(t+1)}$ which follows since worker n is never delayed (proof omitted).

Lemma 5 *If $x_n^{(t)} > x_n^{(t+1)}$ then there exists a $\lambda \in (0, \rho]$ such that*

$$a_n^{(t+1)} = \lambda a_{n-1}^{(t)} + (1-\lambda)a_n^{(t)}.$$

A.4.2 The corresponding Markov chain

For any orbit $\{\mathbf{a}^{(t)}\}_{t=0}^{\infty}$ we will define a specially-structured Markov chain for which the transition probabilities are not stationary. First augment each $\mathbf{a}^{(t)}$ by prepending the dummy allocation $a_0^{(t)} = 0$. Now we define a Markov chain on states $0, 1, \ldots, n$, where state i at step t corresponds to the allocation of worker i at a certain iteration (to be explained shortly). By this correspondence, we may speak of state i at iteration t as being simple or delay according to whether $a_i^{(t)}$ is a simple or delay allocation.

The transition probabilities of this chain model how the values of the allocations change from iteration to iteration. We define $T^{(t+1)}$ to be the matrix of transition probabilities such that $\mathbf{a}^{(t+1)} = T^{(t+1)}\mathbf{a}^{(t)}$.

There are many cases but only a few simple patterns of transitions for each $T^{(t+1)}$, which are as follows.

- From state 0, transit only back to state 0 (so that state 0 is an absorbing state).
- From state i transit to state n or to state 0.
- From state i transit to state $i-1$ or to state 0.
- From state i transit back to state i or to state n or to state 0.
- From state i transit to state $i-1$ or to state n or to state 0.
- From state n transit to state $n-1$ or back to state n.

The values for each $T^{(t+1)}$ are taken from the details of the previous three lemmata. We omit the details since our main result depends only on the patterns and relative sizes of the transition probabilities, not on their exact values; however, we provide the following example:

- If state $i = 1, \ldots, n-1$ is a simple state and $x_i^{(t)} \leq x_i^{(t+1)}$, then from Lemma 3.1 we transit to state n with probability $a_i^{(t+1)}/a_n^{(t)}$ and to state 0 with the complementary probability.

Now, from Lemmata 3, 4, and 5, we have the following.

Lemma 6 *The matrix $T^{(t+1)}$ of transition probabilities is well-defined, stochastic, and satisfies*

$$\mathbf{a}^{(t+1)} = T^{(t+1)}\mathbf{a}^{(t)} = T^{(t+1)}T^{(t)}\cdots T^{(1)}\mathbf{a}^{(0)}.$$

Note that the Markov chain actually makes its transitions in opposite order from TSS iterations; for example, the Markov chain makes its first transition according to matrix $T^{(t+1)}$, the second according to $T^{(t)}$, and so on. Consequently we relabel the transition matrices so that

$$\mathbf{a}^{(t+1)} = A^{(1)}A^{(2)}\cdots A^{(t+1)}\mathbf{a}^{(0)},$$

and therefore, the Markov chain makes its t-th transition according to matrix $A^{(t)}$.

A.4.3 Convergence of $a_n^{(t)}$

Lemma 7 *For any random process in state i at step t the probability of transition to state n or absorption by state 0 within the next i transitions is at least $(1-\rho)^i$.*

Proof sketch This follows from the details of the construction of transition matrices $A^{(t)}$. □

The next lemma shows that the time between successive product completions converges to a constant.

Lemma 8 *The sequence $\{a_n^{(t)}\}_{t=0}^{\infty}$ converges to a positive constant.*

Proof sketch Assume the sequence fails to converge. Then there exists some $\delta > 0$ for which $2\delta < \left|a_n^{(t)} - a_n^{(t')}\right|$ for an infinite number of non-overlapping intervals of indices $[t, t']$. Because the sequence $\{a_n^{(t)}\}_{t=0}^{\infty}$ is bounded above and below, it must be that $\delta < a_n^{(t)} - a_n^{(t')}$ for infinitely many non-overlapping subsequences of indices $[t, t']$ (which will in general lie between the aforementioned intervals). We collapse each of these latter transitions into one, from $\mathbf{a}^{(t)}$ directly to $\mathbf{a}^{(t')}$ via the transition matrix $T^{(t')} \cdots T^{(t+2)} T^{(t+1)}$, and reindex the steps so that t' becomes $t+1$. Now, if the sequence $\{a_n^{(t)}\}_{t=0}^{\infty}$ fails to converge, there must exist some $\delta > 0$ for which $\delta < a_n^{(t)} - a_n^{(t+1)}$ for an infinite number of steps t. For each such step let $\epsilon = \max\{1 - \delta/\tau_n(0,1), 1-\rho\}$. Then for each step t:

- If $\delta < a_n^{(t-1)} - a_n^{(t)}$, then the probability of transition from state n to state 0 is at least $1-\epsilon$ (because $\epsilon \geq 1 - \delta/\tau_n(0,1) \geq a_n^{(t)}/a_n^{(t-1)}$); otherwise,
- the probability of transition from state n back to state n is at least $1-\rho$ or the transition from state n to state 0 is at least $1-\epsilon$ (by Lemma 5 and Lemma 3.1).

Consider a step q at which $\delta < a_n^{(q-1)} - a_n^{(q)}$; there must be probability of at least $1 - \epsilon$ of transition from state n to 0. From Lemma 7 it follows that a process in any state at step $q - n$ has probability at least $(1-\epsilon)(1-\rho)^{2n}$ of being absorbed by state 0 by step q. Because there are infinitely many such q, all processes must eventually be absorbed by state 0, and so $A^{(1)}A^{(2)} \cdots A^{(t)}$ must converge to the zero matrix (with a column of ones for the absorbing state), so that

$$\lim_{t\to\infty} \mathbf{a}^{(t)} = \lim_{t\to\infty} A^{(1)}A^{(2)} \cdots A^{(t)}\mathbf{a}^{(0)} = \mathbf{0}.$$

This, however, is a contradiction because $\sum a_i^{(t)} \geq \tau_n\,(0,1) \geq 1/B > 0$. Therefore the allocation for worker n must converge to some constant; and this constant must be positive because the long-term average production rate is bounded above by nB. □

A.4.4 Convergence of $\{x_i^{(t)}\}_{t=0}^{\infty}$

We have established that $\lim_{t\to\infty} a_n^{(t)} = a_n^*$ exists and so the TSS line eventually produces metronomically. Now we show that the positions of the workers converge.

Theorem 5 $\lim_{t\to\infty} x_i^{(t)}$ *exists for each* $i = 1, \ldots, n$.

Proof sketch $\lim_{t\to\infty} x_2^{(t)}$ exists because the sequence $\{a_n^{(t)}\}_{t=0}^{\infty}$ converges. Then by induction $\lim_{t\to\infty} x_i^{(t)}$ exists for $i = 3, \ldots, n$. □

Having established that $lim_{t\to\infty}\mathbf{x}^{(t)} = \mathbf{x}^*$ we can conclude that, because the TSS dynamics function f is continuous,

$$f(\mathbf{x}^*) = f(lim_{t\to\infty}\mathbf{x}^{(t)}) = lim_{t\to\infty} f(\mathbf{x}^{(t)}) = lim_{t\to\infty}\mathbf{x}^{(t+1)} = \mathbf{x}^*.$$

Therefore the limit point $\mathbf{x}^*$ is the fixed point of f.