The division of labour under uncertainty

Article

Accepted Version

The Division of Labour under Uncertainty

By

Nigel Wadeson*

Date of First Submission: 19th May, 2011

Date of Second Submission: 29th May, 2012

Accepted for publication in *Journal of Institutional and Theoretical Economics*

Abstract

Reductions in the division of labour are a significant feature of modern developments in work organisation. It has been recognised that a reduced division of labour can have the advantages of job enrichment and lower coordination costs. In this paper it is shown how advantages from a lesser division of labour can stem from the flow of work between different sets of resources where the work rates of individual production stages are subject to uncertainties. Both process and project-based work are considered. Implications for the boundaries of the firm and for innovation processes are noted.

Keywords: Coordination, Division of Labour, Uncertainty, Production, Project

JEL: D20

* Department of Economics, University of Reading

1. Introduction

This paper explores implications for the division of labour that result from uncertainties in production processes. The firm has to balance the need for workers and other resources to be available to carry out tasks when needed against keeping them from idleness and from carrying out tasks which do not make full use of their skills and capabilities. Due to work rate uncertainties, availability has to be traded off against lower utilization rates if a rigid division of labour is maintained. Literature on the division of labour seems so far to have given such uncertainties little attention. In fact, with the notable exception of inframarginal analysis (Yang and Ng, 1998), there has been a relative dearth of modern theoretical work on the division of labour (Stigler, 1976, p. 1209; Cheng and Yang, 2004). Empirical work by management researchers and sociologists also suffered a substantial decline from the early 1970s (Carter and Keon, 1986).

The division of labour has been central to our understanding of the organisation of production and of economic progress since Adam Smith's (1776) *Wealth of Nations*. According to Smith, the division of labour is a powerful method of increasing productivity: it improves dexterity, eliminates time spent in switching between tasks, and leads to improvements in tools and machinery. Later, Babbage (1832) pointed out that the division of labour allows high-skilled workers to concentrate on high skill tasks rather than spending some of their time on tasks which do not require their skill level, while lower paid, lower-skilled workers perform the low skill tasks. This is known as the 'Babbage Principle' and can be considered as a key part of Taylorism. It has also been argued that the division of labour is advantageous to the firm in allowing it to exercise more control over workers and that a greater division of labour will develop where worker power to oppose management decisions is low (Reinstaller, 2007).

According to Adam Smith, the division of labour is limited by the extent of the market as, in order to be more specialised, workers need to face large enough markets for their specialist outputs. The division of labour is therefore increased when barriers to trade are reduced so that markets become less fragmented. Improvements in productivity generated through an increased division of labour are traded off against increases in both coordination and transportation costs (Houthakker, 1956). Becker and Murphy (1992) claimed that increasing the division of labour leads to greater agency costs and hold-up problems, the communication of misleading information, and breakdowns in production caused by poor coordination, also stressing that a growth in knowledge leads to increases in specialisation.

A reduction in the division of labour is sometimes termed 'job enlargement'. This refers to widening the number of tasks undertaken by a worker and has often been described in terms of the motivational advantages of 'job enrichment' (Parker et al., 2001). It has also been shown that job enlargement can reduce non-productive time, which includes balance delays (Conant and Kilbridge, 1965, 383-5; Kilbridge and Wester, 1961). Balance delays occur due to bottlenecks in

production systems (Kilbridge, 1960) where one stage processes units at a lower rate than other stages (Matanachat and Yano, 2001). Georgescu-Roegen (1970) saw the elimination of delays as central to explaining the success of the factory system of production (see also Morroni, 1999), as idleness can be eliminated where production is sufficiently large and a number of processes are arranged in a production line.

Note, however, that work rates are normally subject to uncertainties. Delays are therefore caused by variances in work rates and not just by differences in average or constant work rates. The term 'variance delays' can be used to differentiate delays caused by bottlenecks resulting from variances in the production rates of different stages of production from those caused by differences in average rates. Variance delays have been recognised in operations research (Schultz et al., 1998), in particular that greater inventory holding between stages reduces them, and that having a greater number of stations working in parallel in a stage of production increases the predictability of its output (Buxey, 1974).

The term 'task consolidation' refers to combining multiple tasks into one so that they are undertaken by the same workers or resource set. Task consolidation often accompanies a decentralization of decision making (Seidmann and Sundarajan, 1997). Rummel et al. (2005) have considered its advantages in eliminating hand-off delays between activities including delays in transferring knowledge and materials and in waiting for the resources to undertake the next stage to become available to start their work. Note that literature on project planning normally takes the activities as fixed, not considering the possibility of task consolidations. In addition, most literature on resource constrained project scheduling does not consider the effects of uncertainty (Ballestín and Leus, 2009).

Employee and machine flexibility and team working are stressed under lean production which involves teams of multi-skilled workers (Alony and Jones, 2008; Womack et al., 1990). One advantage of flexible equipment and workers is that they facilitate the rapid switching of a production line between different products which would otherwise be more disruptive to the production process. The importance of limits to the divisions of labour are also reflected in the stress that employers have placed more generally in industrial relations on increased employee flexibility and in the central part that task consolidation plays in business process re-engineering (Rummel et al., 2005). In addition, under bucket brigade manufacturing workers move from station to station with the product, which makes the line to some extent self-balancing (Bartholdi and Eisenstein, 1996).

As modern production techniques often involve reductions in the division of labour it is important that we are clear about the circumstances under which a greater or lesser division of labour is advantageous. Matters have changed significantly since the days of Adam Smith. For instance, IT systems, including the use of flexible, programmable production machinery (Milgrom et al, 1991),

can significantly affect the breadth of tasks that workers undertake (Lindbeck and Snower, 2000; Borghans and ter Weel, 2006). Also, the skills needed to operate different types of machines can be similar. In fact, taking Adam Smith's famous example of a pin factory, Pratten (1980) commented that, not only had pin making machines replaced many separate operations, but that each operative also controlled multiple machines. Modern production methods also often involve little or no in-process stock between production stages, so that the balance of work rates becomes of much greater importance (Piore, 1986: p. 7). High quality standards also make workflow more vulnerable to quality variations at different stages of production. Advanced manufacturing technologies increase interdependencies between different parts of firms (Zammuto and O'Connor, 1992: 704-10). Additionally, services make up a large part of many modern economies and delaying the completion of a service by queuing units of output during the production process can amount to a serious reduction in service quality.

The division of labour is also clearly important in understanding different forms of project-based organisation. Some projects involve a succession of different trades. In others the same team largely sees things through from start to end. The scheduling of resources is an important form of coordination, both for the firm's internal resources and in scheduling the work of other firms' resources involved in a project, which helps to ensure that they will be available at the times needed. So, for instance, a firm contracted to do work on the project will want to know when it will be able to start. It will need to schedule its resources to do other work before and after when they will be needed on the project. However, project schedules have to be repeatedly revised due to the uncertainties that exist when they are made. Such revisions disrupt resource coordination.

The paper will proceed by first considering the implications of task consolidations for process-based work, such as a manufacturing line. It will then go on to consider project-based work. The term 'process-based' is used to refer to production that is on-going with each stage of production processing different units of output at any time. With project-based work it is assumed that there is only one unit of output. Here, any particular resource may only be needed for a single task or a restricted set of tasks within the project. The commencement of the work of further resources may then depend on a task done by the resource in question having been completed or having reached a certain stage. The initial consideration of project-based work is followed by a further section giving a more detailed model demonstrating the benefits of task consolidation.

2. Process-Based Production

Consider process-based production. Ideally each stage of production will work consistently at the same rate as the others. Once one stage has finished working on a unit of output, the next stage will have just become ready to work on it and the previous stage will have just finished working on the next unit of output. Hence, the production system is balanced across its different stages. However, assume instead that there is variation in the productivity of each stage of production. If, at

any time, some stages are working faster than others then the system is out of balance. The stages completing work at a lower rate then act as bottlenecks in the system, resulting in variance delays.

Short-lived imbalances can be buffered with stocks of part finished outputs, where applicable. A stage may then still be forced into idleness once stocks from the previous stage have run out. It may also be forced into idleness when available space in which to store stocks of its own output run out. Such space is itself costly. Buffering with inventory can be costly. It involves not only the costs of holding stock, but also, where units of output sit idle between stages of production, the costs of delaying supply in the case where individual units are produced to order. In the case of some services, the customer may be kept idle in passing from one stage of a service to another, such as when being passed from one telephone extension to another and being held in a queue of calls. Inventory buffering can also postpone the identification of a fault in the output of a stage, creating a delay before the next stage starts to work on each unit.

The consolidation of two or more consecutive stages of production, reducing the division of labour so that the same set of resources carries out each of them, is one method of addressing these problems. Workers can continue working on each unit of output until they have finished all of their stages of production. Balance across the stages concerned is thus in-built.

Say that there are two production stages that could be consolidated. Where there are multiple units of resources in a stage assume that they are working in parallel with each simultaneously carrying out the full stage on different units of output. When the production stages are unconsolidated, the rate of output of resource unit i in the first stage of production is r_{1di}. In the second stage the rate of output of resource unit j is r_{2dj}. Assume that the production rate of each resource unit has a random component, ε, so that $r_{1di} = \bar{r}_{1d} + \varepsilon_{1di}$ and $r_{2dj} = \bar{r}_{2d} + \varepsilon_{2dj}$. There are q_{1d} units of resources in the first stage and q_{2d} in the second. For simplicity, assume that there is no inventory between the stages. Each unit of output passes directly from one stage of production to the next. The overall rate of production of the non-consolidated stages at any time is then:

$$R_d = Min(q_{1d}\bar{r}_{1d} + \sum_{i=1}^{q_{1d}} \varepsilon_{1di}, q_{2d}\bar{r}_{2d} + \sum_{j=1}^{q_{2d}} \varepsilon_{2dj}) \qquad (1)$$

Hence, even in the case where the average rates of production are balanced, if the random components of the two stages have different signs or magnitudes then one stage of production is constrained to operate below its potential rate of output.

The resource quantities, q_{1d} and q_{2d}, are, of course, endogenous. For instance, if the resources used in one of the stages of production are very costly, then they can be kept operating more fully to their potential by employing extra resources in the

other stage. This is at the cost of incurring increased expected idleness in the other stage. Hence, production balancing based on average rates of output is not necessarily efficient.

Now consider the consolidated case. The total number of resource units used in the two stages is now q_c. The rate of output of each unit of resources, *k*, per unit of time spent in each stage is now r_{1ck} in Stage 1 and r_{2ck} in Stage 2, where $r_{1ck} = \bar{r}_{1c} + \varepsilon_{1ck}$ and $r_{2ck} = \bar{r}_{2c} + \varepsilon_{2ck}$. Assume that no resource time is now wasted. Each unit of output continues to be worked on by the same resource unit until it is finished.

The overall rate of output of each resource unit, *k*, now averages out the rates of the two stages being the inverse of the sum of the times taken to complete each of the two stages:

$$r_{ck} = \frac{r_{1ck} r_{2ck}}{r_{1ck} + r_{2ck}} \qquad (2)$$

If one of the stages slows down then the workers increase the output from the slowed-down stage by using some of the time they would normally use in the other stage of production. Hence, the rate of output across the consolidated stages as a whole is not as badly affected as if the stages had their own dedicated workers. The extent of the bottleneck in the wider production system is reduced.

In addition, the overall rate of output of the consolidated stages is likely to be more tightly distributed around the mean, relative to the mean, than for the unconsolidated tasks. This is inherent in the fact that the time taken to process a single unit involves the sum of the random elements involved in the time taken for a resource unit to complete each stage, so long as they are less than perfectly correlated. If work in one of the consolidated stages of production slows down, the other may progress at its normal rate or speed up. An extreme value of the sum of a number of random terms that are not highly correlated has a low probability as it requires that every term is of large magnitude and has the same sign. So, for example, the sum of two identical and independent uniform distributions has a triangular distribution.

Further, if we define $\bar{r}_c$ as the expected value of r_{ck}, ε_{ck} as the random component of r_{ck}, and R_c as total production summed over every resource unit, *k*, then.

$$R_c = q_c \bar{r}_c + \sum_{k=1}^{q_c} \varepsilon_{ck} \qquad (3)$$

A further advantage of consolidation is that it can lead to a greater degree of parallelism of production as q_c will often be greater than either q_{1d} or q_{2d}. This

again reduces the variability of production rates due to the summation of the random terms in Equation 3. For instance, say that there are two identical machines plus their operators in each of two separate stages of production. If one machine breaks down then the capacity of its stage of production is halved, so creating a substantial bottleneck in the overall production process. Now say that the two stages are consolidated and that there are now four identical machines and their operators which each carry out both stages. Now if one machine breaks down capacity falls by only a quarter. The bottleneck is much reduced. Consolidation could be described in terms of making the chain of production stages shorter and fatter. There are not only less interfaces between stages of production; there is also less dependence on individual workers and machines. There can therefore be significant advantage in machines and workers being capable of undertaking multiple production tasks.

The production rates of individual resources can also vary if the nature of the resources used depends on whether or not tasks are consolidated. For instance, a machine for consolidated production might be more complex and more likely to break down than a more specialised one. On the other hand, it might be a simpler, more general purpose item of equipment.

Note also a source of economies of scale inherent in the summations of the random terms in Equation 1. If larger scale results in a greater parallelism of activities in a stage of production then this reduces the variability of the overall output of that stage of production, so facilitating the division of labour. However, increased scale will also involve a discontinuity where a different production technology is employed that results in a much higher rate of production per unit of resources, or if the new technology is more automated and makes the production rate less susceptible to worker related variations.

The above discussion assumes that resources cannot be added or shed instantly and without cost. The firm cannot count on being able to buy on spot markets as any one of them may lack available supply. Searching and matching also takes time. So does the induction of new workers into the firm. Similarly, workers cannot be sure of being able to find alternative work at short notice. Workers who are not employed by the firm, and therefore given some degree of forward security of demand for their services, may contract with other firms and so become unavailable. Indeed, they have an incentive to seek forward contracts in order to avoid being unemployed in future time periods.

However, under certain circumstances, the firm can instead use the flexibility that it has over the coordination of its internal resources. For instance, it could use resources that are capable of working across multiple stages of production and redeploy some of them between stages as and when imbalances occur. If one stage is processing units of output at a higher rate than another, then resources will be transferred in order to rebalance the work rates. Evidence on the benefits of this is cited by Daniels, Mazzola and Shi (2004). Note that it may be possible to deal

with a significant proportion of the variance by only making a fraction of the resources flexible in this way. Such redirection could be achieved through hierarchical management control or through a degree of self-management of production teams working under appropriate incentives. The latter could facilitate quick decisions based on knowledge of current conditions held within the teams.

3. Project-Based Production

Now consider a project such as a new product development or construction project. Assume that different specialist resources could be used in each of its stages. Each resource has to be brought into the project to perform its tasks and, once they are complete, is no longer needed. The project schedule has uncertain timings. Here, the uncertainty is not about the number of units of output processed in each time period. Rather, it is how long each task within the project will take to be completed, or at least to reach a point which allows others to start work.

As above, the firm cannot rely simply on spot contracting for resources. Whole sets of resources need to be available simultaneously and in sequence so that coordinated availability across multiple markets is needed. The firm can instead forward book the resources. However, given the uncertainty in the project's schedule, the firm does not know exactly when they will actually be needed. If the divisions of labour are rigidly maintained, therefore, then where a delay in a resource starting and completing its work will be costly, there is an incentive to insert extra time buffers into the schedule (the use of buffers is common in projects), so that a resource is scheduled to be available earlier than it might actually be needed and is scheduled to finish later than its work might actually be completed. Widening the time periods of the resource bookings increases the chances that they will be available when it will turn out that they will actually be needed. In addition, there may be uncertainty as to whether the resources will actually become available when booked, as they may not complete previous work to schedule and so be late in starting work on the project. These problems apply whether the firm is contracting for resources in the market or whether it is using internal resources; there are competing demands for resources within the firm and their use needs to be coordinated through scheduling.

The costs can be split into four components which are traded off against each other in determining the optimal start time and duration of the resource booking. Assume that the task in question can only be started once the preceding project task has been completed. Firstly, z_{s1} is the expected cost due to the possibility that the preceding task ends after the resources for the task have been scheduled to start their work. Assume that this keeps them idle while waiting for the preceding task to be finished. Note that another possibility in reality is that they would move to another project which could then keep them unavailable for some time. Secondly, z_{s2} is the expected cost due to the possibility that the preceding project task is completed before the resources for the task become available to start their work. These are costs of the project being unnecessarily delayed while awaiting the resources. z_{s1} and z_{s2} depend both on the uncertainty over the duration of the

preceding task and on the endogenously determined time, S, at which the resources are scheduled to start work on the task. The minimum possible value of S is at the earliest time that the preceding task might be completed and its maximum value is at the latest possible time that the preceding task might be completed.

A later start-time of the resource booking, S, decreases z_{s1} while increasing z_{s2}: $\delta z_{s1}/\delta S<0$ and $\delta z_{s2}/\delta S>0$. As the start time becomes later it becomes progressively less likely that the resources will be left idle while waiting for the previous task to end and more likely that there will be a delay between the end of the previous task and the start of the resource booking. Hence, as S increases, z_{s1} falls at a decreasing rate and z_{s2} rises at an increasing rate: $\delta^2 z_{s1}/\delta S^2>0$, $\delta^2 z_{s2}/\delta S^2>0$. The marginal benefit of a later start time, in terms of a reduction in z_{s1}, is therefore downward sloping and the marginal cost in terms of an increase in z_{s2} is upward sloping. The optimal start time is earlier the higher are the costs of delay and the lower are the costs of the resources per time period.

Where a task's duration is fairly long it may be quite likely that it will be possible to lengthen the booked resource time after a delayed start. Where there is some likelihood of not being able to do this then for a given duration of the resource booking, due to the increased expectation of some idleness before starting work, an earlier start time increases z_{e1} and decreases z_{e2}, where z_{e1} is the expected cost resulting from the possibility that the scheduled resource time ends before the task is finished and z_{e2} is the expected cost of the resource booking ending after the task has been completed. Assume that the latter involves expected costs of idleness. In reality, however, workers might actually slow down to fill the time available. However, the effects of an earlier start time on these two expected costs can be countered by an increase in the duration, T, of the scheduled resource time, $\delta T/\delta S<0$, in order to make up for the expected idleness before starting work.

The task end time is made more uncertain by an earlier start time, S. If there were further tasks that could only start when the task had been finished then this would have knock-on effects on their scheduling. This therefore gives an increased incentive to schedule later start times for the resources undertaking each task, so delaying the expected project completion date.

The longer the duration of the resource booking, the lower is z_{e1} and the greater is z_{e2}: $\delta z_{e1}/\delta T<0$ and $\delta z_{e2}/\delta T>0$. Increases in the duration are progressively more likely to result in extra idle time rather than avoided delays: $\delta^2 z_{e1}/\delta T^2>0$ and $\delta^2 z_{e2}/\delta T^2>0$. The marginal benefit of a longer duration, in terms of a lower z_{e1}, is therefore downward sloping and the marginal cost in terms of a higher z_{e2} is upward sloping. The optimal duration will be longer the more costly are delays per time period, resulting in greater expected idleness. It will be shorter the higher are the costs of the resources per time period, resulting in higher expected delays. The total of the expected costs, over and above those that would be borne if the resources could be scheduled with certainty over how long each task will take, or

if they could be obtained instantly as an when needed, is:

$$z = z_{s1} + z_{s2} + z_{e1} + z_{e2} \quad (4)$$

Scheduling uncertainty therefore results in both expected costs of idleness and of resources not being available when needed. The expected cost, *z*, will be high when uncertainty over the end time of the previous task and over the duration of the task itself are high and where the costs of both the delays (and disruptions) caused by non-availability and the costs of resources per time period are high. Non-availability will be more likely where the resource type is fairly scarce in the firm and in the market, such as may be the case where specific knowledge is involved. Note that the duration, *T*, will be set close to the latest possible completion time for the task and the start time close to its earliest possible value if delay costs dominate resource costs. However, if this reflects very high delay costs then the expected costs of idle resources may still be large in their own right. Similarly, if resource costs dominate delay costs then the start time will be set at near the latest possible time and the duration near its shortest possible value, so resulting in significant expected delays.

However, the firm does not simply have to accept these expected costs, even after minimising them in terms of selecting the scheduled start times and durations. Instead they provide an incentive for task consolidations which reduce them in the following ways. Firstly, if the same resources undertake multiple tasks then their overall work duration is made more predictable relative to the mean than the durations of the individual tasks. Hence, the expected costs due to the possibilities that a scheduled resource duration will turn out to be insufficient or that it will be too long (represented by z_{e1} and z_{e2} above) are both reduced. The overall scheduled resource time can therefore be both shorter and more effective.

Secondly, some tasks will be very unlikely to run out of forward booked resource time. For instance, if a number of tasks are consolidated, undertaken one after another by the same resources, then the earlier ones will be very unlikely to take so long as to exhaust the total booked resource time. Hence, the risk of disruption costs being incurred when those tasks run out of scheduled resource time (represented by z_{e1} above) is reduced or eliminated by consolidation. In addition, if the work begins to slip behind schedule then managers have more time to react in order to gain extra resource time before the scheduled time runs out.

Thirdly, resources used to carry out a consolidated sequence of tasks are available to undertake their next task once they have completed their current task. Hence, the expected costs z_{s1} and z_{s2} are both eliminated if the task is consolidated with the previous task. There will often also be advantages in workers being able to utilise project specific knowledge gained in earlier tasks while undertaking later tasks and also in avoiding the costs of transferring materials.

Finally, sometimes projects will require some reworking of tasks. This can result

from information gained during the performance of later tasks. If a worker is used across a range of project tasks then it is more likely that they will still be working on the project when the need for reworking is discovered. This is particularly advantageous when the individual concerned has significant project specific knowledge.

A further strategy, as with process-based work, is to move resources between tasks within a project, or from other activities within the firm, as and when required. Again, resources must be available that are capable of undertaking the task types in question. The more that the firm's resources are capable of working across multiple task types, the greater the flexibility there will be. However, there will also be clashes of resource needs within the firm. A resource will not always be available to transfer immediately into the project even when it is internal to the firm. There will be more freedom to move resources out of other tasks if those tasks are relatively non time-critical and if the disruption costs involved are low.

It is significant that different tasks within a project have different levels of time criticality. For instance, if the laying of the foundations of a house is completed too late then other tasks, such as building the walls, will be delayed while if turf laying in the garden is delayed somewhat then it may have no effect on other tasks. One strategy would be to consolidate critical path tasks, giving the advantages set out above. However, critical path tasks might also be consolidated with non-critical path tasks. A critical path task could then take scheduled resource time from non-critical path tasks as and when needed. It could also release resource time to them once completed. Hence, overall resource bookings could be made which would ensure that the critical path task would not become short of resource time. Note that a similar strategy could also be applied to process-based work. While work on the production line itself will often be time-critical, if some of a worker's time is allocated to other tasks that are not then the worker can be switched into them when not needed on the production line and also back from them when necessary.

Masten et al. (1991, p. 12) found evidence of skilled workers in a large naval construction project being kept busy in tasks that were not time-critical in order to utilise them for more significant periods of time, keeping them occupied outside times when they were needed to carry out their primary tasks. Love (2010, pp. 487-8) reports evidence of internalisation due to time criticality and a high cost of delay even in the absence of opportunistic hold-up. Hameri and Heikkila (2002) give case study evidence showing major delays between a project task being completed and the commencement of the subsequent task. They also present evidence demonstrating the importance of good communications on task progress and reallocations of resources in leading to the improved interfacing between different tasks. Hammer and Champy (1993) provide further evidence of delays between tasks. Serpell et al. (1997) give evidence of labour and equipment lying idle in Chilean construction projects observing c. 53% of work time being spent on non-productive activities. Eden et al. (2000) discuss how small delays and

disruptions can have serious knock-on effects on a project.

4. A Model of Project Task Consolidation

The model that follows demonstrates how task consolidations, under which the same resources are made to carry out more than one task one after another, can ease the effects of a combination of project schedule uncertainty and resource constraints. The model demonstrates the first two of the above sources of advantage from task consolidation. Firstly, task duration uncertainties are consolidated, so resulting in a more favourable probability distribution. As explained above, the summation of independent random variables, in this case task durations, results in a variable more tightly distributed around the mean, relative to the mean. Secondly, where tasks are consolidated, an earlier task is less likely to be disrupted through the resource booking not being long enough to complete it. Indeed, the overall resource booking may well be longer than the maximum time that might be taken to complete the first task. The model involves two optimisations. Firstly, the optimal resource booking for a non-consolidated task is derived. Then the same is done for consolidated tasks. This then allows the expected costs of consolidated and non-consolidated tasks to be compared.

First consider the case where resources are to be allocated to a single task in isolation. They are to be booked to carry out only that task and then leave the project. Forward booking reserves resources to work on the project for a specified number of time periods. The duration of the task, t, is uniformly distributed $(0, h]$. Hence, the probability density of t is $1/h$. The initial resource booking is for a duration of T ($T \leq h$) time periods. The cost of the resources per time period is w ($w>0$). For simplicity, assume that the resources actually do become available on the date for which they are booked to start work and that the state of the project is such that they can start work on that date rather than sitting idle. This places the focus of the model on the uncertainty over how long the task itself will take. If the initial resource booking turns out not to be long enough for the task to be completed then an extra expected cost of D ($D>0$) is incurred due to disruption of the work. This might actually involve either a delay to the next task that the resources will work on so that they can complete their current task before moving on (in which case D might be a penalty charge) or a delay while waiting for further resource time. This has to be weighed against the chance that the initial resource booking turns out to be longer than required to complete the task, in which case, for simplicity, it is assumed that the resources stand idle from the time of completion of the task until the end of the booking (rather than being expected to be used in some less than ideal way).

The expected extra costs, Y, resulting from possible disruption or resource idleness over and above the costs of the resource time that will actually turn out to be needed to carry out the task, tw, are the disruption cost, D, multiplied by the probability that the task takes longer than the duration of the resource booking, T, plus the per period resource cost multiplied by the expected value of the duration

of booked resource time that might be left over following the completion of the task:

$$Y = \int_T^h \frac{D}{h} dt + \int_0^T \frac{w}{h}(T-t)dt \quad (5.1)$$

$$= D - \frac{TD}{h} + \frac{T^2 w}{2h} \quad (5.2)$$

Differentiating this with respect to T gives:

$$\frac{dY}{dT} = \frac{Tw - D}{h} \quad (6)$$

Setting this equal to zero gives the optimal resource booking:

$$T_{01}^{*} = \frac{D}{w} \quad (7)$$

The corresponding value of Y, obtained by substituting the value of T_{01}^{*} into Equation 5.2, is:

$$Y_1^{*} = D - \frac{D^2}{2hw} \quad (8)$$

This illustrates the trade-off between saving on the costs of a longer resource booking and avoiding the disruption costs that result when the booking turns out to be too short. If resource time is cheap relative to the potential disruption costs then the optimal booking is relatively long. These values apply so long as $D \leq hw$, otherwise T has reached its maximum value (i.e. the maximum time that the task could possibly take):

$$T_{02}^{*} = h \quad (9)$$

The value of Y corresponding to $T = T_{02}^{*}$ is:

$$Y_2^{*} = \frac{hw}{2} \quad (10)$$

Now consider the case where there are two such tasks, A and B, and a strategy of task consolidation is applied so that Task B is carried out immediately after the completion of Task A by the same resources. This means that a single resource booking is made to cover both tasks. This consolidates the uncertainties over the

completion times of the tasks. If one of the tasks is completed slowly compared to expectations then the other may be completed relatively quickly. The initial resource booking for the consolidated tasks is again T ($T \leq 2h$).

First consider the case where the resource booking is at least as great as the maximum time taken to undertake Task A, $h \leq T \leq 2h$. This is equivalent to assuming that $D \geq hw/2$, as is shown later. The expected extra costs due to the possibility that the tasks will be completed before the end of the initial resource booking are as follows. Note that this expression allows both for the case where Task A is completed early enough for there to be more than the maximum time to complete Task B left, $t_A < T-h$, and the case where it does not, $t_A \geq T-h$. In the latter case, for both tasks to be completed before the end of the resource booking requires that $t_B \leq T - t_A$.

$$z_{11} = \int_0^{T-h} \int_0^h \frac{w}{h^2}(T - t_A - t_B) dt_B dt_A + \int_{T-h}^h \int_0^{T-t_A} \frac{w}{h^2}(T - t_A - t_B) dt_B dt_A \qquad (11.1)$$

$$= \frac{w}{6h^2}\left(2h^3 - 6Th^2 + 6T^2 h - T^3\right) \qquad (11.2)$$

This result can be interpreted by noting that it can also be obtained by utilising the fact that the sum of the two durations with independent uniform distributions has a duration, t, with a triangular distribution. Note that the apex of the triangle of the distribution is at $t=h$, the probability density to the left of the apex ($0 < t \leq h$) is t/h^2, and to the right of the apex up until the maximum possible value of t is $2/h - t/h^2$, giving the following expression:

$$z_{11} = \int_0^h \frac{t}{h^2}(T - t) w dt + \int_h^T \left(\frac{2}{h} - \frac{t}{h^2}\right)(T - t) w dt \qquad (12)$$

The expected cost resulting from the possibility that the initial booking of resources will turn out not to be long enough to complete Task B, which requires that there are less than h time periods of the booking left after the completion of Task A ($t_A > T-h$) and that Task B takes more than the remaining duration of the booking ($t_B > T - t_A$), is:

$$z_{12} = \int_{T-h}^h \int_{T-t_A}^h \frac{D}{h^2} dt_B dt_A \qquad (13.1)$$

$$= \frac{D}{2h^2}(T - 2h)^2 \qquad (13.2)$$

The overall expected extra costs, over and above those that would be incurred if the resources were booked for the exact actual durations of Tasks A and B, is Z_1:

$$Z_1 = z_{11} + z_{12} \tag{14}$$

Differentiating Z_1 with respect to the duration, T, gives:

$$\frac{dZ_1}{dT} = \frac{1}{2h^2}\left(4wTh + 2DT - wT^2 - 2wh^2 - 4Dh\right) \tag{15}$$

The optimal value of T is thus:

$$T_1^* = \frac{1}{w}\left(D + 2hw - \sqrt{D^2 + 2h^2w^2}\right) \tag{16}$$

Note that this value is increasing in D and approaches $2h$ as D rises to values that are very high relative to hw. Hence, where disruption costs are high relative to resource costs the duration of the resource booking for the consolidated tasks, being close to the maximum duration of the tasks in order to largely eliminate the possibility of disruption costs being incurred, is close in value to the sum of the resource bookings made when they are not consolidated.

Substituting this value into Equation (14) gives the associated value of Z_1:

$$Z_1^* = \frac{1}{6h^2w^2}\left(2D^3 + (2h^2w^2 + D^2)^{3/2} - 3D^2\sqrt{2h^2w^2 + D^2} + 6h^3w^3 + 6h^2w^2(D - \sqrt{2h^2w^2 + D^2})\right) \tag{17}$$

For convenience, define r as a measure of the strength of potential disruption costs relative to resource time costs, such that $r=D/hw$. Now consider the superiority of consolidation. This is measured in terms of the focus of the model which is the scheduling problem. It should therefore be interpreted relative to other factors not included in the model, particularly whether using resources specialised in individual tasks has advantages over the same resources undertaking both tasks. Other factors that might favour consolidation should also be noted, such as project specific knowledge gained in the first task being useful in the second and any incentive advantages that might result.

The superiority, S_1, of consolidation over the two tasks being carried out by separate resource sets, where $r \leq 1$ (i.e. $D \leq hw$) so that the value of Y is Y_1^* as defined in Equation 8, is:

$$S_1 = 2Y_1^* - Z_1^* \tag{18}$$

Expanding this gives:

$$S_1 = hw\left(r + \frac{(r^2+2)^{\frac{3}{2}}}{3} - r^2 - \frac{r^3}{3} - 1 \right) \tag{19}$$

Where $r \geq 1$, so that the value of Y is Y_2^* as defined in Equation 10, the superiority of consolidation is instead given by S_2:

$$S_2 = 2Y_2^* - Z_1^* = hw\left(\frac{(r^2+2)^{\frac{3}{2}}}{3} - \frac{r^3}{3} - r \right) \tag{20}$$

Now consider the case where the duration of the booked resource time, T, is for less than the maximum time that it will take to complete Task A, $T<h$. Note that this puts T to the left of the apex of the triangular distribution of the duration of the consolidated tasks. Assume now that each task has a minimum duration, m, and that T, t_A, t_B, and h are therefore measured as times over and above m. Hence the overall booking is $2m+T$, and the overall duration of a task is $m+t$. The significance of m comes from the further assumption, for simplicity, that $2m>h$, so that Task A is never interrupted when the overall resource booking is exhausted. It is always Task B that is interrupted when this happens. Hence, we do not need to consider the case where Task A runs out of resource time and then later Task B does as well. Note that, even if it were assumed that $2m<h$, the time booked for Task B would still give extra security that Task A could be completed within the overall booked resource time, so reducing expected disruption costs.

The expected cost of resources becoming idle due to the two tasks being completed within T is now:

$$z_{21} = \int_0^T \int_0^{T-t_A} \frac{w}{h^2}(T - t_A - t_B)dt_B dt_A = \frac{T^3 w}{6h^2} \tag{21}$$

The expected cost resulting from Task B not being completed within the initial resource booking, allowing both for the case where Task A takes less than T ($t_A<T$) and where it does not ($T \leq t_A \leq h$) is now:

$$z_{22} = \int_0^T \int_{T-t_A}^h \frac{D}{h^2} dt_B dt_A + \int_T^h \frac{D}{h} dt_A = \frac{D}{2h^2}(2h^2 - T^2) \tag{22}$$

The overall extra expected cost is Z_2:

$$Z_2 = z_{21} + z_{22} \tag{23}$$

Differentiating with respect to T gives:

$$\frac{dZ_2}{dT} = \frac{1}{2h^2}(T^2 w - 2TD) \tag{24}$$

Setting this equal to zero gives the optimal value of T, which is now the same as the optimal overall resource booking if the tasks are not consolidated ($2T_{01}^*$):

$$T_3^* = \frac{2D}{w} \tag{25}$$

Hence, the condition $T \leq h$ is equivalent to:

$$D \leq \frac{hw}{2} \qquad \text{(or, equivalently, } r \leq 0.5\text{)} \tag{26}$$

The corresponding value of Z_2 is:

$$Z_2^* = D - \frac{2D^3}{3h^2 w^2} \tag{27}$$

The superiority of consolidation over the two tasks being carried out by separate sets of resources ($2Y_1^* - Z_2^*$) is now:

$$S_3 = hw\left(r - r^2 + \frac{2}{3}r^3\right) \tag{28}$$

The figure below plots the superiority of consolidation (S_1, S_2, and S_3) for the case where hw=1. S_3 is applicable in the range $0 \leq r \leq 0.5$. S_3 meets S_1 at r=0.5, and S_1 is then applicable in the range $0.5 \leq r \leq 1$. S_1 meets S_2 at r=1, with S_2 being applicable to the right of this point.

[Insert Figure around here]

A point illustrated by the figure is that the superiority of consolidation must decline after some point as the optimal resource booking, T, under consolidation becomes closer to its maximum of $2h$ with increasing values of r. As the disruption cost increases the duration of the resource booking for an unconsolidated task climbs more quickly towards its maximum value than does the duration of the booking made if the tasks are consolidated. In fact, for an unconsolidated task, the resource booking reaches its maximum duration of h at

r=1 (i.e. D=hw). Once it has reached this value any further increases in the disruption cost have no impact as the task will then never run out of booked resource time before it is finished, the booking having been made to cover the maximum time that the task can take. However, for consolidated tasks increases in the disruption cost still continue to increase the expected costs and so the superiority of consolidation declines. It is at values of r close to unity, where D and hw are close in value, that the superiority of task consolidation is greatest. Where, on the other hand, one these costs dominates the other then the overall resource bookings for consolidated and non-consolidated tasks are similar and consolidation is less advantageous.

Note that this result is based on the consolidation of two similar tasks. However, consolidation is clearly also valuable if the disruption and resource costs are both high for the first task but the disruption cost is significantly lower for the second task, so long as the resources are not replacing ones in the second task that are so much less costly that the benefits are wiped out. The trade-off involved in scheduling the first task when not consolidated is eased by consolidation, under which the disruption cost becomes the lower cost involved with the second task. However, consolidation is still valuable if one of the two types of cost is significantly higher than the other if the task can be consolidated with a second with a lower disruption cost. For instance, if disruption costs dominate resource costs in the first task, so that the resource booking for the task if not consolidated would be towards the top end of the possible value of the task duration, then consolidation with a second task with a low disruption cost can lead to a significant saving in resource costs.

High disruption costs are particularly associated with critical path tasks and with resources that are difficult to replace or to secure extra time for quickly. An example of a situation where the disruption cost and resource cost would both be high would therefore be a critical path task needing high-cost specialists using expensive equipment who gain significant specific knowledge during their work. If they are not booked for long enough then they are difficult and costly to replace. Note that disruption costs can potentially be reduced somewhat by prompt information sharing so that managers can react earlier when tasks are not running to schedule.

5. Conclusion

The effects of production uncertainties on the division of labour have been considered. Dividing labour into narrower sets of tasks has the consequence of making the production process more vulnerable to the performance of individual workers and machines. In addition, it creates extra interfaces between production stages. A unit of production can be delayed at an interface between production stages, awaiting the attention of the resources that will undertake the next stage of production. Knowledge may have to be transferred across it in order to process each unit of output or such knowledge may be lost at it. A production stage may

be forced into idleness while waiting for the subsequent stage to catch up or while awaiting the output of the previous stage. Dividing tasks can make production rates more uncertain. This increases expected costs of both idleness and delays for a given level of resources. Although it is not a focus of the paper, it should also be recognised that the division of labour is important for incentives within firms and for contracting issues between firms. For instance, if two tasks are interdependent then it is likely to be easier to identify who is responsible for performance if they are performed within the same firm and by the same team.

A model of project-based work was presented in order to show formally how a consolidation of tasks, so that they are undertaken one after another by the same resources, can reduce both resource costs and the expected value of disruption costs that result when a resource booking turns out not to be long enough to complete its tasks. The model demonstrated two advantages to a reduced division of labour. Firstly, the consolidation of tasks results in a task duration probability that is more tightly distributed around its mean, relative to the mean. For instance, if one task takes longer than expected then another may be completed more quickly than expected. The consolidation of tasks therefore eases the problem of scheduling resources to undertake tasks with uncertain durations. Secondly, the model demonstrated how the consolidation of tasks reduces expected disruption costs. This is because the earlier of a set of consolidated tasks is less likely to run out of booked resource time as the resource booking is made to cover a full sequence of tasks rather than just the first task alone. While the model focused on advantages of task consolidations, these have to be weighed against disadvantages such as losing advantages of more specialised knowledge.

The model further demonstrated that the advantages of consolidating similar tasks are greatest when both the disruption cost and resource costs are high but neither type of cost is so great that one dominates the other. At very low values of the disruption cost relative to resource costs very short resource bookings are made whether the tasks are consolidated or not. Hence, there is very little expected idle resource time but a high expectation that the very low disruption cost will be incurred. The overall resource costs incurred in both cases are then similar and the disruption cost has little impact on overall expected costs, being very low in value. Hence, the expected costs are similar whether or not the tasks are consolidated. As the disruption cost rises from low values consolidation becomes increasingly superior. As it rises further to intermediate values relative to resource costs, the resource bookings for unconsolidated tasks rise more quickly towards their maximum durations than is the case for consolidated tasks. Because of this, further rises in the disruption cost, having no further impact on unconsolidated tasks, then start to cause a decline in the superiority of consolidation. At high values of the disruption cost relative to resource costs the resource bookings are at or near the maximum times that the tasks might take, whether they are consolidated or not, and there is little or no chance of the disruption cost being incurred. In this case, therefore, the superiority of consolidation is again low.

In addition to the gains that can be made by consolidating similar tasks, significant gains can also be made from consolidating a pair of tasks where the first task performed has high values of both disruption and resource costs and the second task has a significantly lower disruption cost. An example could be where a critical path task is consolidated with a second task that is not time critical.

Note that the model did not make the assumption that there is a dependency between the tasks involved such that one must always be done after the other, whether or not they are undertaken by the same resources. There might be such a dependency or it might be that the tasks can be arranged in that way in order to gain the advantages of task consolidation. However, for the latter case it should be recognised that interdependencies between tasks mean that they cannot always be resequenced without costs (Simon, 2002; Langlois, 2002). Hence, it is more likely that some tasks will be chosen for consolidation with any particular task than others. A task with only weak interdependencies with other tasks could be useful for task consolidation, if its sequencing in the work schedule could be easily changed to allow it, or alternatively two tasks with strong interdependencies might be consolidated with each other. The consolidation of two tasks that in any case have to be completed one after the other has additional advantages to those explored in the model because, as explained in Section 3, consolidation avoids a possible delay between the first task ending and resources becoming available to start the second task and also a possible overlap where resources booked to undertake the second task are left idle while waiting for the first task to be finished.

While tasks may be fully consolidated so that they are always undertaken by the same workers, there is often, in reality, also the alternative of moving resources between tasks as the need arises. For instance, if one stage of process-based production is moving slowly then it may be possible to speed it up by moving in resources from other stages. Such dynamic transfers of resources can also reduce idleness by redirecting otherwise idle (or underutilised) resources into other tasks, among which could be quality assurance activities and machine maintenance.

The arguments made in this paper may seem less relevant to production on a larger scale. For instance, if there are many workers and machines carrying out the same stage of production in parallel then this will make the overall output of that stage more predictable, although sometimes large scale production will involve individual machines with very high capacities. Also, worker related uncertainties can be reduced through automation. However, it should be remembered in respect to process-based production that, firstly, modern production methods often involve the holding of little stock between different stages of production. Secondly, high quality standards can make the system more vulnerable to disruptions caused by variations in output quality that occur in individual stages. Thirdly, the introduction of advanced manufacturing technologies tends to result in close integration and hence interdependencies between different parts of the firm. Fourthly, the flexibility of some modern production technologies allows for

easier switching between products. However, this can result in much more frequent switching which can increase the frequency of disruptions to the production system. Variations in work flow can therefore be more important than might otherwise be expected. Instead of an ever higher division of labour, multi-skilled work teams are often used instead.

With regards to the scale of project-based work, consider a firm which repetitively undertakes similar large projects and has a number under way at any one time. If the carrying out of the work of individual resources is of uncertain duration and is dependent on each project reaching a certain stage in a schedule with uncertain timings then it will still be a difficult problem to ensure that they are able to move between projects without delays and periods of idleness. Reducing the division of labour, as has been shown, reduces the severity of this problem, even though the firm might be considered to be operating at a large scale. At the extreme, a single team carries out the whole of a project and then starts a new one.

The arguments made in this paper are relevant to the boundaries of the firm. The need for the consolidation of tasks and for the flexibility to reallocate resources between tasks as needed can lead to more tasks and resources being brought within the same firm. However, the flexible reallocation of resources may also be facilitated by relational contracting. There is more scope for flexible reallocations if a supplier is responsible for undertaking a greater amount of work for the same customer, creating options for resources to be shifted between different tasks being undertaken for that customer as urgency dictates. Such coordination issues have been relatively ignored in the theory of the firm in favour of those relating to opportunism and incentives and so represent a significant opportunity for future work (Foss, 1999; Love, 2010).

Recent history has shown a significant movement in some industries from large, vertically integrated firms to less vertically integrated supply chains (Brusoni and Prencipe, 2001) and so it might be asked whether the division of labour between firms is actually steadily increasing. Such changes have been facilitated by factors such as modular product architectures, standardisation, and modern information and production technologies. Individual firms concentrate on one or more modules, except for the lead firm which has an important systems integration function. A modular architecture significantly reduces interdependencies between components, so that the design and production of each module is less interconnected to those of others. The parallel development of different modules can then significantly speed up product development. While there is some vertical disintegration of the supply chain, there may be multi-skilled teams within individual firms, as noted above.

With such an arrangement of production among firms there is a risk of disruptions caused by delays in the supply of individual modular components, at stages where they need to be assembled together, but these are reduced in some ways. One of these ways is through scale. Some component manufacturers are very large and

utilise flexible manufacturing technologies which allow production lines to swiftly switch between products. Hence, demand variations can be well diversified across different customers and markets. Another is there often being a high level of sharing of information on demand forecasts and production schedules. Yet another is that the customer may commit to buy from a supplier for a significant period of time and may also smooth its own production in order to have a more predictable demand for supplies. Hence, the costs of a greater division of labour between firms are reduced. However, in some respects the extent of the division of labour between firms has actually fallen as individual suppliers now tend to offer wider ranges of services. This led Sturgeon (2002) to refer to them as 'turn-key' suppliers. This suggests that there are still important reasons for consolidating multiple activities within each firm.

The arguments of the paper are also relevant to innovation. Bringing knowledge to bear in an innovation project can be seen, not just in terms of having the knowledge within the firm or within a network of firms, but also in terms of having the ability to access the right individuals at the right times without holding up the project. One implication is that there is an increased incentive to resolve uncertainties early on, such as through prototyping, so that resources can then be scheduled with greater certainty. Uncertainties may also be reduced by limiting the scope of the project so that it is less radically innovative. Modularisation can also be used, as the technology allows, reducing the extent of interdependence between different components so that there is less need for interactions between those working on different components and less need for the reworking of tasks done previously on other components. The division of labour can also be reduced, as explained in this paper, though this may mean sacrificing advantages of more specialised knowledge

Being located in an industrial cluster could improve resource availability, rather than simply search costs, and so ameliorate such problems. The use of information technology could also increase the pool of potentially usable resources by making distance less important, particularly where regular face to face communication is relatively unimportant and where the resources do not have to work on a particular site. Clusters and information technology also help to facilitate multitasking in the sense of a resource working concurrently on more than one project. For instance, a worker may spend part of a day working on a project for one customer and the rest working on a project for another.

Resource coordination problems can be exacerbated by greater competition where this increases the importance of a short time to market, so making product development project delays more costly. In addition, technological progress can require the use of a greater range of technical specialists. Hence, it can be argued that globalisation and technological progress have led to more use of flexible relationships between firms collocated in clusters or closely linked through modern information technologies partly in order to facilitate the division of labour in the face of uncertainties within production and product development processes.

References

Alony, I. and Jones, M. (2008), "Lean Supply Chains, JIT and Cellular Manufacturing – The Human Side," *Issues in Informing Science and Information Technology*, 5, 165-175.

Babbage, C. (1832), *On the Economy of Machinery and Manufactures*. R. Clay, London.

Ballestín, F. and Leus, R. (2009), "Resource-Constrained Project Scheduling for Timely Project Completion with Stochastic Activity Durations," *Production and Operations Management*, 8, 4, 459-474.

Bartholdi III, J. J. and Eisenstein, D. D. (1996), "A Production Line that Balances Itself," *Operations Research*, 44, 1, 21-34.

Baumgardner, J.R. (1988), "The Division of Labor, Local Markets, and Worker Organisation," *The Journal of Political Economy*, 96, 3, 509-527.

Becker, G. and Murphy, K. (1992), "The Division of Labor, Coordination Costs, and Knowledge," *Quarterly Journal of Economics*, 107, 4, 1137-1160.

Borghans, L., and Ter Weel, B. (2006), "The Division of Labour, Worker Organisation, and Technological Change," *Economic Journal*, 116, 509, F45-F72.

Brusoni, S., and Prencipe, A. (2001), "Unpacking the Black Box of Modularity: Technologies, Products and Organizations," *Industrial & Corporate Change*, 10, 1, 179-205.

Buxey, G.M. (1974), "Assembly Line Balancing with Multiple Stations," *Management Science*, 20, 6, 1010-1021.

Carter, N. M. and Keon, T. L. (1986), "The Rise and Fall of the Division of Labour, the Past 25 Years," *Organization Studies*, 7, 1, 57-74.

Cheng, W. and Yang, X. (2004), "Inframarginal Analysis of Division of Labor: A Survey," *Journal of Economic Behavior & Organization*, 55, 2, 137-174.

Conant, E.H. and Kilbridge, M.D. (1965), "An Interdisciplinary Analysis of Job Enlargement: Technology, Costs, and Behavioral Implications," *Industrial and Labor Relations Review*, 18, 3, 377-395.

Daniels, R., Mazzola, J., and Shi, D. (2004), "Flow Shop Scheduling with Partial Resource Flexibility," *Management Science*, 50, 5, 658-669.

Eden, C., Williams., T., Ackermann, F., and Howick, S. (2000), "The Role of

Feedback Dynamics in Disruption and Delay on the Nature of Disruption and Delay (D&D) in Major Projects," *Journal of the Operational Research Society*, 51, 3, 291-300.

Foss, N.J. (1999), "The Use of Knowledge in Firms," *Journal of Institutional and Theoretical Economics*, 155, 458-486.

Georgescu-Roegen, N. (1970), "The Economics of Production," *The American Economic Review*, 60, 2, 1-9.

Hameri, A.-P. and Heikkilä, J. (2002), "Improving Efficiency: Time-critical Interfacing of Project Tasks," *International Journal of Project Management*, 20, 143-153.

Hammer, M. and Champy, J. (1993), *Reengineering the Corporation: A Manifesto for Business Revolution*, Harper Business, New York.

Houthakker, H. S. (1956), "Economics and Biology: Specialization and Speciation," *Kyklos*, 9, 181–189.

Kilbridge, M.D. (1960), "Reduced Costs through Job Enlargement: a Case," *The Journal of Business*, 33, 4, 357-362.

Kilbridge, M. and Wester, L. (1961), "The Balance Delay Problem," *Management Science*, 8, 1, 69-84.

Langlois, R.N. (2002), "Modularity in Technology and Organization," *Journal of Economic Behavior & Organization*, 49, 1, 19-37.

Lindbeck, A. and Snower, D. (2000), "Multi-task Learning and the Reorganization of Work," *Journal of Labor Economics*, 18, 3, 353-376.

Love, J.H. (2010), "Opportunism, Hold-up and the (Contractual) Theory of the Firm," *Journal of Institutional and Theoretical Economics*, 166, 3, 479-501.

Masten, S.E., Meehan, J.W. and Snyder, E.A. (1991), "The Costs of Organization," *Journal of Law Economics and Organization*, 7, 1–25.

Matanachat, S. and Yano, C. (2001), "Balancing Mixed-Model Assembly Lines to Reduce Work Overload," *IIE Transactions*, 33, 1, 29-42.

Milgrom, P.R., Qian, Y. and Roberts, J. (1991), "Complementarities, Momentum and the Evolution of Modern Manufacturing," *American Economic Review*, 81 (2), 84-88.

Morroni, M. (1999), "Production and Time," in K. Mayumi and J.M. Gowdy

(eds), *Bioeconomics and Sustainability*, Cheltenham: Edward Elgar.

Parker, S., Wall, T., and Cordery, J. (2001), "Future Work Design Research and Practice: Towards an Elaborated Model of Work Design," *Journal of Occupational & Organizational Psychology*, 74, 4, 413-440.

Piore, M.J. (1986), "Corporate Reform in American Manufacturing and the Challenge to Economic Theory," mimeo., Sloan School of Management, MIT, Massachusetts.

Pratten C.F. (1980), "The Manufacture of Pins," *Journal of Economic Literature*, 18, 1, 93-96.

Reinstaller, A. (2007), "The Division of Labor in the Firm: Agency, Near-Decomposability and the Babbage Principle," *Journal of Institutional Economics*, 3, 293-322.

Rummel, J.L., Walter, Z., Dewan, R. and Seidmann, A. (2005), "Activity Consolidation to Improve Responsiveness," *European Journal of Operational Research*, 161, 3, 683-703.

Schultz, K., Juran, D., Boudreau, J.W., McClain, J.O. and Thomas, L.J. (1998), "Modeling and Worker Motivation in JIT Production Systems," *Management Science*, 44, 12, 1595-1607.

Seidmann A. and Sundararajan A. (1997), "Competing in Information Intensive Services: Analyzing the Economic Impact of Task Consolidation and Employee Empowerment," *Journal of Management Information Systems*, 14, 2, 33-56.

Serpell, A., Venturi, A., and Contreras, J. (1997), "Characterization of Waste in Building Construction Projects", in: Alarcon, L. (ed.), *Lean Construction*, A.A. Balkema, Rotterdam, pp. 67-77.

Simon, H. A. (2002), "Near Decomposability and the Speed of Evolution," *Industrial & Corporate Change*, 11, 3, 587-599.

Smith, A. (1776), *The Wealth of Nations*, Strahan, London.

Stigler, G.J. (1976), "The Successes and Failures of Professor Smith," *The Journal of Political Economy*, 84, 6, 1199-1213.

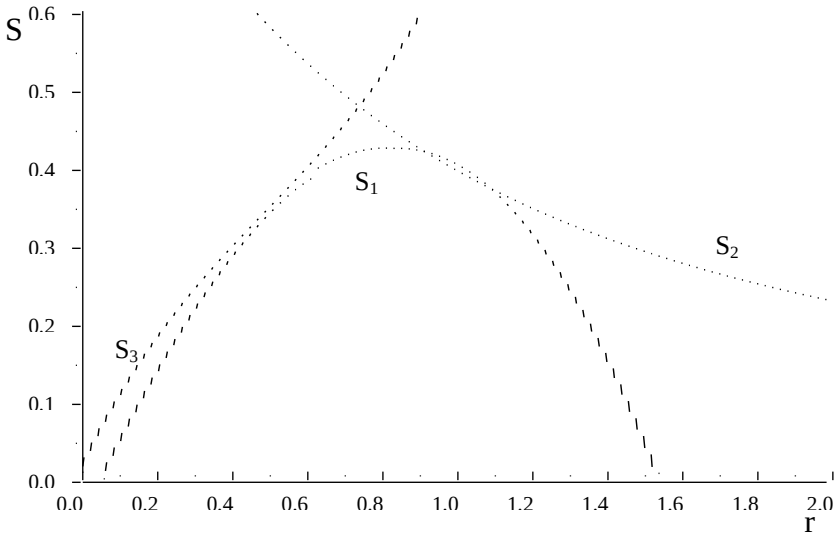
Sturgeon, T. J. (2002), "Modular Production Networks: a New American Model of Industrial Organization," *Industrial & Corporate Change*, 11, 3, 451-496.

Womack, J.P., Jones, D.T. and Roos, D. (1990), *The Machine that Changed the World*, Rawson Associates, New York.

Yang, X. and Ng, S. (1998), "Specialization and Division of Labor: a Survey", in K. Arrow et al (eds.), *Increasing Returns and Economic Analysis*, Macmillan, London.

Zammuto, R.F., and O'Connor, E.J. (1992), "Gaining Advanced Manufacturing Technologies' Benefits: The Roles of Organization Design and Culture," *Academy Of Management Review*, 17, 4, 701-728.

The Superiority of consolidation (*hw*=1)



Nigel Wadeson
Department of Economics
University of Reading
PO Box 218
Reading
RG6 6AA
UK
Email: n.s.wadeson@rdg.ac.uk