AN INDIRECT WORKFORCE (RE)ALLOCATION MODEL FOR SEMICONDUCTOR MANUFACTURING

Chen-Fu Chien

Department of Industrial Engineering
and Engineering Management
National Tsing Hua University
Hsinchu 30013, TAIWAN, ROC

Wen-Chih Chen

Department of Industrial Engineering
and Management
National Chiao Tung University
Hsinchu 30013, TAIWAN, ROC

Shao-Chung Hsu

Department of Industrial Engineering
and Engineering Management
National Tsing Hua University
Hsinchu 30013, TAIWAN, ROC

ABSTRACT

Semiconductor industry is a capital intensive and knowledge intensive industry, in which human resource management and human capital enhancement is increasingly important. To maintain competitive human resource, it is critical to develop a decision framework for headcount planning and workforce allocation for indirect labors. Motivated by the needs in real setting, this study aims to develop a model for allocating indirect workforce among semiconductor fabrication facilities to meet expected outputs and labor productivity improvement. Workforce allocation and reallocation based on the overall corporate workforce level is essential so that the shortage or exceed workforce will be balanced among different production sites. The key to achieve this purpose is the proper understanding of real requirements of each production site according to its corresponding tasks assigned. Non-parametric activity analysis approach is used for the workforce requirement estimation given delegated tasks. The estimation is based on the best performance from the past with adjustments reflecting the expected productivity growth.

1 INTRODUCTION

Driven by Moore's law, semiconductor industry is knowledge and capital intensive. Thus, human capital enhancement and human resource management is getting important nowadays (Chien and Chen 2007; Leachman et al. 2007). In particular, workforce planning and headcount allocation have become important issues for both research and practice in semiconductor manufacturing. Research has been done on workforce planning decisions including staff scheduling or rostering, which determines work timetables for staff so that the demand can be satisfied while optimizing certain criteria. For example, Thompson and Goodale (2006) present a staff scheduling method for the cases where workforces have different productivity levels. Staff scheduling problem is an extension of conventional scheduling problems and comprehensive reviews can be found in (Aykin 2000; Burke et al. 2004; Ernst et al. 2004). These decisions are typically operational and all detailed information is assumed to be obtainable. Another research is related to job assignment or reallocation that determines the staff-job assignment. Assigning workforces to jobs can be modified and modeled as the classic assignment problems (e.g., Holder 2005) or by other means such as simulation (e.g., Zulch et al. 2004). In addition, long-term staffing deals with determine the optimal workforce requirements of each category hired at each period in light of production ramping and technology migration. A stream of this research is workforce planning optimization under deterministic conditions (e.g. Mundschenk and Drexl 2007). Another stream considers stochastic nature of the problem such as learning curve and turnover, which can be modeled as Markov decision processes and then solved using different techniques (e.g., Gans and Zhou 2002; Ahnet et al. 2005). The applications of workforce planning have been studied in various industries. For example, Mundschenk and Drexl (2007) propose an integer programming model for long-run staffing particularly for printing industry, and Pesch and Tetzlaff (2005) study the interactions between staffing and scheduling decisions in the automotive industry. Bard et al. (2007) investigate the workforces planning for United States Postal Service mail processing and distribution centers. Holder (2005) studies the optimizing process of assigning sailors to jobs for US Navy in an attempt to increase sailor satisfaction.

However, most of the existing studies on workforce decisions were based on simplified assumptions and thus can hardly solve real problems especially in knowledge intensive high-tech industries. This research was motivated by a semiconductor company in real setting in Taiwan. Firstly, with the scale and involved automation of semiconductor fab increases, the amounts and percentage of knowledge workers also raise hugely. Firms are facing the situation that the cost of automation and manpower is ascending year by year, and the engineers and technical operators play more and more important roles in the factories.

Secondly, semiconductor companies in Taiwan used to attract and retain talent by issuing stock dividends rather than giving high salaries. Beginning in 2008, the new accounting rules request expensing employee bonuses, which will affect the financial reports of most high-tech companies. Therefore, maintaining proper level of workforce and enhance people productivity become critical.

Thirdly, equipments are long-term investments that are lack of flexibility to any adjustments in response to demand variation once the decisions are made. On the

other hand, semiconductor companies usually have 10% to 20% of the workforce left annually. Thus, workforce planning and allocation can be applied for productivity improvement and people productivity also becomes a key performance indicator (KPI).

This study aims to propose a method for workforce planning and to allocate and reallocate indirect workforce for each production site so that the delegated tasked can be fulfilled using proper level of workforce. Indirect workforce is important and under investigation due to two primary reasons. First, different to direct workforce typically handling routine jobs, indirect labor deals with technical project oriented tasks, which are difficult to identify and quantify the required level. Second, indirect workforce is more technical and skilled employees, such as engineers, their salary and bonus are usually much higher than direct labors. Therefore, this paper focuses on indirect workforce.

The remainder of this paper is organized as follows. Section 2 analyzes and structures the problem in the real case in fab operations and the decision modules are thus defined. Sections 3 and 4 addressed how to estimate workforce requirements incorporating with productivity improvement driving. Section 5 presents a workforce reallocation so that the shortage or exceed workforce will be balanced among different production sites, followed by the conclusion of this paper.

2 A DECISION FRAMEWORK

There are many factors and issues related to workforce decisions. Figure 1 presents the relationship among decision issues and the relevant information. Major decisions involved are as follows:

Workforce planning: is the decision on determining the proper corporate workforce level across certain time periods, such as two or three years, to meet the future product demand. In high technology industry, the only thing unchanged is change. The long term product demand usually has significant vibration, seasonality and uncertainty; these makes demand hard to predict. The workforce planning indeed decides the workforce supply

for the future. In addition, some external factors on human resource side, such as recruiting constraints, turnover rate, seasonality, and corporate strategic policy should be considered.

Workforce requirement estimation: is to understand the "real" workforce demand in response to actual task assignments. Typically, this is decomposed to the manufacturing site level and is site dependent to reflect the difference in production conditions at each site. The decision corresponds to a single coming period with all tasks assigned. Good understanding of the "real" workforce requirements to fulfill the expected tasks is the foundation for workforce decision. In particular, this issue states the workforce demand. It is noted that "real" means feasible and minimum, which indicates the requirements under ideal situation without any inefficiency.

Workforce allocation: deals with inconsistency, especially the shortage, between current available workforce supply and site demands. Given the workforce requirement and available supply, this decision tries to resolve the gap by reallocating workforce to each fab. It can be seen as a minor tuning based on up-to-date status and certain operational management objectives.

Driving productivity growth: is relatively ambiguous and not directly related to the workforce decision. It aims to drive the improvement of workforce-based performance to enhance the sustainability under competitive business environment. The decision may be made subjectively and more top-down oriented. Its role is more close to setting a target or constraints but not making a decision itself.

From the process aspect of decision making, one first needs to determine the supply and demand of workforce, identify the gap between supply and demand, and then take actions in response to the gap. The supply side of the workforce is a mid-term decisions based on a long term aggregate forecasting. The recruiting and training need longer lead time. The risk of uncertainty for the long term forecasting is high due to the competitive market and rapid technology growth. Therefore, the plan plays a role to smooth the fluctuant market to maximize the long term overall benefits. Robustness of the plan to
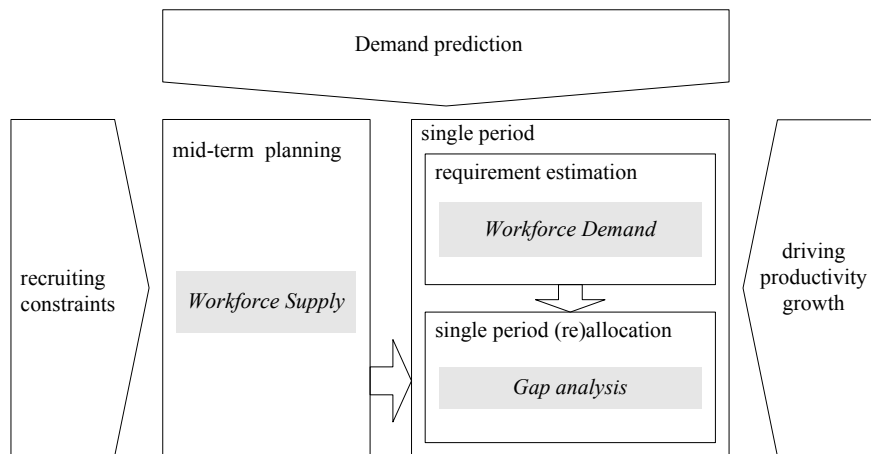

Figure 1: A framework for workforce decision

meet environmental uncertainty is more important than the forecasting accuracy and optimality of the plan. The demand side of the workforce is based on the assigned tasks which is known and obtained by arriving customer orders. The timing is right before the execution and after the monthly or quarterly production plan. That is to understand a minimal requirement of the workforce to meet the production plan.

Unfortunately, there is always a gap between the current workforce demand and the supply based by the much earlier plan due to business environment uncertainty. One needs to identify the gap and to resolve the issues. Some solutions such as production outsourcing are not under consideration in this study. We focus on reallocating workforce so that the infeasibility of meeting production plan can be minimized and balanced over manufacturing sites. In addition, besides "objective" requirement estimation and resource allocation, some subjective productivity enhancement concern, usually from top level, are considered when making the decision.

3 REQUIREMENT ESTIMATION

Clearly understanding the proper workforce requirement

$$\hat{T} \equiv \left\{ (\mathbf{x}, \mathbf{y}) : \sum_{r \in S} \mathbf{x}_r \lambda_r \leq \mathbf{x}; \sum_{r \in S} \mathbf{y}_r \lambda_r \geq \mathbf{y}; \lambda \geq 0; (\mathbf{x}, \mathbf{y}) \in \Gamma \right\} \quad (1)$$

so that specific tasks can be accomplished on time is important but not easy. This is because of the complex production process and rapid changes in product scale and mix. Traditional approaches use labor standard and are bottom-up approaches. These approaches rely on very detailed understanding of labors and tasks under the stable processes. Instead of engineering bottom-up approaches, we utilize the nonparametric frontier models to determine minimal requirement based on past experiences. In the following sections, notions of nonparametric frontier models will first be introduced, and followed the workforce estimation model.

3.1 Resource-output Transformation

Broadly speaking, any organization is a transformation process that consumes resources (inputs) to provide valuable outputs including products or services (Chien et al. 2003; Chien et al. 2007). It is always demanded to answer "What is the ideal level of resources to provide required outputs?" If the underlying relationship of the transformation from resources to outputs is provided, the question can be answered easily. However, this input-output transformation process is unknown in practical applications. Moreover, it is difficult to describe and specify the processes since many resources and outputs are involved and their interactions are too complex to understand. The major methodology used in this work is the non-parametric frontier analysis. Unlike bottom-up engineering approach, this method does not try to specify all explicit input-output relations in detail. It looks the process as a "black box" with all unknown detailed relations; focuses on the overall results observed from the process, and estimates real the transformation

process by observations. Therefore, it can handle multiple inputs and multiple outputs without *a priori* functional form or weight assignment.

Consider an input set I and an output set J. Let $\mathbf{x}$ be the $|I| \times 1$ positive input vector and $\mathbf{y}$ be the $|J| \times 1$ positive output vector. The production possibility set (PPS), T, is defined as $T \equiv \{ (\mathbf{x}, \mathbf{y}) : \mathbf{y}$ can be produced by $\mathbf{x} \}$. PPS is a way of describing the resource-output transformation process. In practice, T is unknown. Given S as a set of observations with input-output vectors $\{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \ldots, (\mathbf{x}_{|S|}, \mathbf{y}_{|S|})\}$, the empirical production possibility set (EPPS) $\hat{T}$ can estimate the real T. The estimation is based on free disposal, convexity and constant returns to scale (CRS) (Charnes et al. 1978; Banker et al. 1984). Free disposal assumes that $(\mathbf{x'}, \mathbf{y'}) \in T$ if $(\mathbf{x}, \mathbf{y}) \in T$ and $\mathbf{x'} \geq \mathbf{x}$, $\mathbf{y'} \leq \mathbf{y}$. Convexity says $(\mathbf{x'}, \mathbf{y'}) \in T$ and $(\mathbf{x}, \mathbf{y}) \in T$ $\rightarrow$ $\alpha(\mathbf{x'}, \mathbf{y'}) + (1 - \alpha)(\mathbf{x}, \mathbf{y}) \in T$ for $0 \leq \alpha \leq 1$. Constant returns to scale assumes that any change in input size results in the same rate of change in output size, i.e., $(\mathbf{x}, \mathbf{y}) \in T \rightarrow (k\mathbf{x}, k\mathbf{y}) \in T$ for $k \geq 0$. The EPPS can consequently be expressed as a set of linear inequalities in $|S|$ nonnegative variables and denoted as:

where Γ represents a set of additional technical constraints, which is already well known or is required to achieve, for a workforce-output transformation process.

It is reasonable to assume, and thus to require, that the productivity is always achievable as long as it ever achieved. The philosophy of the proposed method is to identify the "best performance" in the past as the estimation the future capability. However, the "best performance" requires detailed clarification first; it is related to the trade-offs among performance indices that will be addressed here.

In practices, single factor productivity indices, ratios of one output to one input, are tracked to evaluate and monitor the performance of the production unit, such as fab operations. As one may often observe, among a set of productivity values, no consistent conclusion can be made. For example, among all historical records, fab A may be the best according to Y1 per unit resource but performs poorly in Y2 per unit resource where Y1 and Y2 representing different outputs. Does fab A perform well or badly? How well does fab A perform, overall? Is fab A a "best practice"? These are hard questions to be answered, and they are the main points in the fights and the "this is because" excuses when evaluating performance.

The observations in fact come from the nature of substitution and allocation among resources and outputs. There are always trade-offs among different performance indices. The concept of the "best performance" bases on the final winner of pairwise comparisons. More precisely, if a unit loses to another, then it has no chance to be the best. In the situations with multiple productivity indices, losing to another unit means that it is worse in all indices. It is also called being dominated and the one being

superior in all indices refers as dominant. After all possible pariwise comparisons, the remaining non-dominated units are different to draw any conclusion since all of them are better off in some aspects and worse-off in the rest criteria. This is the concept of Pareto (in)efficiency in economics (McGuigan et al. 2002). It should be noted that the definition of "best" performance here is conservative in the sense of disqualifying some input-output to be the best.

Figures 2 and 3 are simple examples with two productivity indices $Y1/X$ and $Y2/X$ where $Y1$ and $Y2$ are two outputs generated by resource X. In Figure 2, E is no better than B and C in both $Y1/X$ and $Y2/X$ (comparing with C, they have the same performance in $Y2/X$, but C is better than E in $Y1/X$). Hence, E is dominated by B and/or C. C also dominates D since it is better in $Y2/X$ although the same in $Y1/X$. No conclusions of being dominated can be drawn for the pairwise comparisons related to A, B and C. Therefore, $\{A, B, C\}$ are all the best practices according to the definitions. Figure 3 is another case where all units are dominated by F in both $Y1/X$ and $Y2/X$. The remaining non-dominated unit is F and therefore
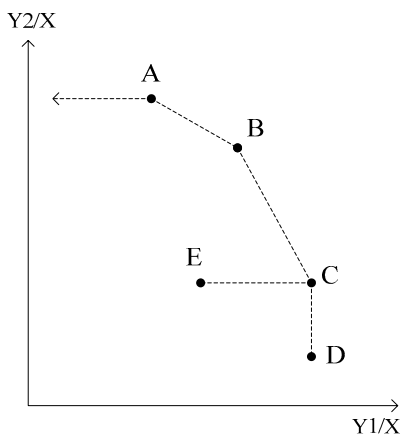

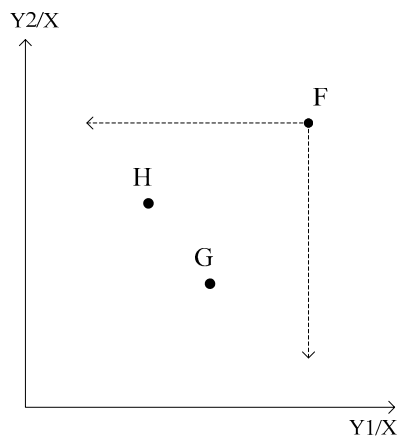Figure 2: An example with two productivity indices

Figure 3: An example with two productivity indices

only F is the best performance. The proposed method can handle both cases and identify the proper best practices based on the definition.

3.2 Estimating Total Indirect Labor Requirement

A fab can be considered as a process transforming resources (inputs) to products or services (outputs). This section particularly considers the fab-wise workforce-outputs transformation. Additional notations are: F represents a set of fabs, S^k denotes the historical observed production records for fab k, $k \in F$. For fab k, based on the historical experience $\left\{ \left(\mathbf{x}_r^k, \mathbf{y}_r^k \right), r \in S^k \right\}$ where $\mathbf{x}_r^k \in \Re_+^{|I|}$ and $\mathbf{y}_r^k \in \Re_+^{|J|}$ are the resource and output vectors.

Figure 4 represents the workforce-output transformation capability of a single fab. This model tries to estimate the indirect labor requirements for mature fabs. It is difficulty to estimate indirect labor needed than directly labor. This is because the work content is not regular and routine, and thus bottom-up approaches may not be appropriate here. In this model, all input/output records in S^k are measured in a three-month period (quarter) with different starting month. It is a quarterly estimation and planning. We have the input set $I = \{PE, PIE, EE\}$ representing three different indirect labor categories (process engineer, process integration engineer, and equipment engineer). The output set is $O = \{Q, SL, TD\}$ where Q is the total output volume, SL represents the customer service loading and TD represents the total technology difficulty. The detailed definitions are as follows:

Process engineers (PE) and **equipment engineers** (EE): are in charge of period maintenance and trouble shooting. They are measured in headcount and take the average in the period.

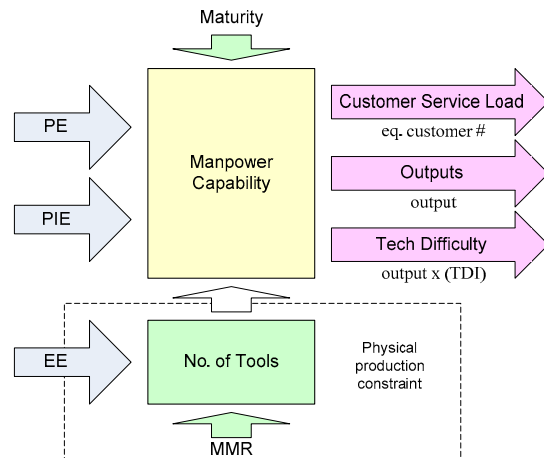**Process integration engineers** (PIE): is

Figure 4: Fab-wise workforce-output transformation (indirect labor)

responsible for the product yield and customer services such as answering questions related to process recipe and experiments. It is also measured as average headcount in quarterly basis.

Total output (*Q*): is the quarterly equivalent 8 inch wafer output. It is the key valuable product generated by the fab.

Customer service loading (*SL*): uses equivalent service customer number directly. A fab indeed provides customized products based on customers' needs. More customers require more efforts in serving and setup.

Technology difficulty (*T D*): represents the *extra* loading due to technology complexity. It can be measured in *Q*, i.e. equivalent output $TD = Q \times (1 - TDI)$, where *TDI* is the *technical difficulty index* that is similar to the proposed technical difficulty weight (*TDW*). However, *TDW* is the changing percentage comparing to $t-1$, a relative metric not an absolute metric which has constant meaning over time. *TDI* is a relative metric comparing to a very early base period ($t = 0$) and thus can be seen as an absolute metric. For example, if TDW_{Y99}^{Y00} = 1.2 and TDW_{Y00}^{Y01} = 1.1, TDI^{Y01} = 1.2 × 1.1 for year 2001 using year 1999 as the base period.

As shown in Figure 4, there are many other factors affecting the workforce-output transformation of a fab. For example, the maturity of the product represents the level of experience to produce current product mix and has influence on the transformation efficiency. In fact, rather than having direct impact on output level, equipment engineers (*EE*) affect outputs through the operation of the tools which include number of tools, man-machine ratio (*MMR*), and other physical production constraints. However, most interactions among functions to outputs, functions to functions and outputs to outputs are not understood and are very difficult to understand. For example, the computation of *MMR* depends on actual workforce and tool numbers, but this ratio does not reflect the real operational condition when the man and machine utilization is low. Therefore, the macro viewpoint is adopted to describe the process and let historical data speak themselves.

For fab *k* at time period *t*, one can assume that the estimated production capability of fab *k*, is $\hat{T}^k$:

$$\hat{T}^k \equiv \left\{ (\mathbf{x}^k, \mathbf{y}^k) : \sum_{r \in S^k} x_{ir}^k \lambda_r^k \le x_i^k, \forall i \in \{PE, PIE, EE\}; \right.$$
$$\sum_{r \in S^k} y_{jr}^k \lambda_r^k \ge y_j^k, \forall j \in \{Q, SL, TD\}; \quad (2)$$
$$\left. \lambda_r^k \ge 0, \ \forall r \in S^k \right\}$$

S^k here is the set of adjusted historical records of fab *k*.

The key concept and assumption to determine the requirement is that something can be done today and tomorrow if it was done before. Therefore, a "similar" best practice, as addressed in Section 4.1, from historical experience can be identified. There are many possible best practices with different indirect labor combination (**x**)

based on the definition. We try to estimate total workforce requirement for time period *t* with delegated output tasks (*Q, SL, TD*). Therefore, the estimated **x** not only is the best practice but also is minimal in total amount. Given delegated tasks, $\widetilde{y}_j^k \quad \forall j \in \{Q, SL, TD\}$, the required total workforce is computed as follows:

$$THC^k(\widetilde{\mathbf{y}}^k, S^k) = \min\left\{ \sum_{i \in I} \hat{x}_i^k : (\hat{\mathbf{x}}^k, \widetilde{\mathbf{y}}) \in \hat{T}^k \right\}$$
$$= \min_{\lambda, \hat{x}^k} \sum_{i \in I} \hat{x}_i^k$$
$$s.t. \ \sum_{r \in S^k} x_{ir}^k \lambda_r^k \le \hat{x}_i^k, \forall i \in I = \{PE, PIE, EE\}; \quad (3)$$
$$\sum_{r \in S^k} y_{jr}^k \lambda_r^k \ge \widetilde{y}_j^k, \forall j \in J = \{Q, SL, TD\};$$
$$\lambda_r^k \ge 0, \ \forall r \in S^k.$$

$THC^k(\widetilde{\mathbf{y}}^k, S^k)$ is the minimum total workforce requirement based on the experience of S^k. In practices, the total amount is preferred than detailed quantities in each category. This is because that the indirect labor is flexible to be responded to different types of tasks through training, and fab managers also can gain some reallocation flexibility.

4 PRODUCTIVITY IMPROVEMENT

It is reasonable to assume, and thus to require, that the productivity is always achievable as long as it ever achieved. This spirit is the foundation of using (3) as an estimation for the workforce transformation. However, composition of S^k with raw historical records would lead to a conservative estimation. It is that S^k does not consider the real or target growth of productivity indices for each fab. In fact, the growth exists, it may come from some well known production properties, such as increasing returns to scale and economies of scale, or the technology improvement and experience accumulation. On the other hand, the target growth is set to drive productivity improvement as shown in Figure 1. Suppose the database has (x_i^k, y_j^k) for $i \in I$, $j \in J$ and $k \in S^k$, and (x_i^k, y_j^k) is the records occurs at given base time *b*. Moreover, suppose we need to use this information to predict the workforce requirement for any particular time period *t*, and a predetermined annual discount rate α_j is given. More precisely, it is one plus the growth rate and should be no less than one. One should adjust y_j^k to $y_j^k(t)$ using α_j as:

$$y_j^k(t) = y_j^k \times (\alpha_j)^{t-b}. \quad (4)$$

It should be noted that the value of α_j is output type dependent and assumed to be constant over time. α_j

also depends on the decision horizon of the analysis. For example, in this study, the records are quarterly with different beginning month, e.g., one record is between Jan. and March in 1997 and the next is for Feb., March and April 1997, and the monthly, not quarterly, discount rate α_j should be used.

For example, suppose the annual growth rate is 5%, and thus monthly rate $\alpha_j = (1.05)^{1/12}$. If $y_j^k = 100$ is the output j usage for record k occurring at time $b = 5$, e.g., the initial month is May. 1997, one should adjust using (4) and its adjusted level for the initial month Dec. 1997, $t = 12$, is $y_j^k (t = 12) = 100 \times (1.004)^{12-5} = 102.887$ where $\alpha_j = (1.05)^{1/12} = 1.004$. It means that using the same input level and providing the annual growth rate 5%, output j should be 102.887, i.e., 102.887 is achievable if it is found to be 100 seven months ago for this particular record k. Moreover, the discount rate is not necessary to be constant; an indexing system similar to consumer price index can be developed.

All records in S^k are modified to S^{k*} using (4) with respect to their time stamps. The total workforce considering productivity growth thus is computed by (3) given substituting S^k by S^{k*}.

In fact, saving resources can also lead to productivity growth rate α, at least mathematically; only the output-oriented approach is adopted in this case. This is because that there is possibility of low demand while the workforce cannot be layoff due to enterprise culture in Taiwan. The short term bottom line of workforce reduction action is usually recruiting frozen and by its nature labor turnover.

5   WORKFORCE ALLOCATION

This section addresses single-period workforce (re)allocation problem. That is to allocate the up-to-date total available workforce (*GHC*) to each fab so that its assigned tasks can be fulfilled. The *GHC* is determined by long-term workforce planning and also the subjective top-down enforcement to drive the productivity growth. Namely the actual target is less than available workforce.

5.1   Suffering Index

One of the most important issues to allocate available workforce to each units is the fairness no matter whether the workforce is sufficient or not. This is particularly critical when the shortage of workforce exists because one hopes to share the shortage loading evenly. Suffering index (*SI*) measures the magnitude of suffering the workforce shortage by normalizing the current workforce level. It is defined as

$$SI^k \equiv \frac{required\ workforce}{allocated\ workforce} \qquad (5)$$

$SI^k > 1$ indicates fab k is suffering labor shortage. The larger SI^k is, more serious workforce shortage fab k is suffering. $SI^k < 1$ indicates fab k has more labor than it needs.

The objective of the allocation is to minimize the largest (worst) suffering index of all fabs. It is $\min \max_{k \in F} SI^k$ and can be rewritten as

$$\begin{aligned} &\min \nu \\ &s.t. \quad \nu \geq SI^k, \forall k \in F. \end{aligned} \qquad (6)$$

5.2   Allocation Model

It is not a good idea to have significant change of the workforce level for each fab. The associated problems are costs of the learning experience and morale. Therefore, it is preferred that the new allocated workforce has minimal change in terms of volume and mix or is within the tolerance.

Given the current workforce CHC^k for fab k, the allocated amount HC^k should results a percentage change within the predetermined tolerance ε_k, e.g., $\varepsilon_k = 20\%$. That is

$$\frac{HC^k - CHC^k}{CHC^k} \leq \varepsilon^k \qquad (7)$$

Combining the factors addressed above, the complete allocation of workforce to each fab can be determine by (HC-Allocation) as follow:

$$\min_{\nu, HC^k} \nu \qquad \text{(HC-Allocation)}$$

$s.t.$ minmax constraints: $\qquad (8)$
$$\nu \geq \frac{THC^k(\tilde{\mathbf{y}}^k, S^{k*})}{HC^k}, \forall k \in F ;$$
Individual capability constraints:
$$HC^k \leq (1+\varepsilon^k)CHC^k, \quad \forall k \in F ; \qquad (9)$$
$$HC^k \geq (1-\varepsilon^k)CHC^k, \quad \forall k \in F ; \qquad (10)$$
Overall requirement constraints:
$$\sum_{k \in F} HC^k = GHC \qquad (11)$$
$$HC^k \geq 0.$$

(HC-Allocation) is a non-linear programming problem, but its optimal solutions can be obtained without solving (HC-Allocation). That is, without considering (9) and (10), one has optimal allocated workforce to fab k:

$$HC^{k*} = GHC \times \frac{THC^k(\tilde{y}, S^{k*})}{\sum_{k \in F} THC^k(\tilde{y}, S^{k*})} \qquad (12)$$

With (9) and (10) under consideration, if HC^{k*} in (12) violates (9) or (10), HC^{k*} to be $(1+\varepsilon^k)CHC^k$ or $(1-\varepsilon^k)CHC^k$, respectively. Eq. (12) is

used to reallocate rest of the workforce to remaining fabs. The pseudocode of the procedure *allocateHC* can be summarized as the following algorithm:

procedure $allocateHC\left(GHC, \varepsilon^k, CHC^k, THC^k(\tilde{y}, S^{k*}), \forall k \in F\right)$

while $GHC > 0$

begin

$\qquad HC^k := GHC \times \dfrac{THC^k(\tilde{y}, S^{k*})}{\sum_{k \in F} THC^k(\tilde{y}, S^{k*})}, \forall k \in F$

$\qquad$ for $k \in F$

$\qquad$ begin

$\qquad\qquad$ if $HC^k > (1 + \varepsilon^k)CHC^k$ then

$\qquad\qquad\qquad HC^k := (1 + \varepsilon^k)CHC^k, F := F - k, GHC := GHC - HC^k$

$\qquad\qquad$ if $HC^k < (1 - \varepsilon^k)CHC^k$ then

$\qquad\qquad\qquad HC^k := (1 - \varepsilon^k)CHC^k, F := F - k, GHC := GHC - HC^k$

$\qquad$ end

end

return $HC^k, \forall k \in F$

6   CONCLUSION

This work is motivated by a real semiconductor industry case in Taiwan. The objective is to provide a method to (re)allocation indirect workforce for each production site so that the delegated tasked can be fulfilled using proper level of workforce. Indirect workforce planning is important because indirect workforce is more technical and skilled employees, such as engineers, their salary and bonus are usually much higher than direct labors. Moreover, different to direct workforce typically handling routine jobs, indirect labor deals with technical project oriented tasks, which are difficult to identify and quantify the required level. Therefore, this paper focuses on indirect workforce. The proposed method is a two-stage approach. A workforce requirement in response to assigned tasks is estimated based on historical experience. Best performance from the past is identified as the "ideal" requirement. Given the total available workforce at cooperate level and requirement for each fab, the workforce is allocated or reallocated to each of the fabs so that the workload can be balanced. Productivity growth was also incorporated in the proposed model.

ACKNOWLEDGEMENTS

REFERENCES

Ahn, H.-S., R. Righter, and J.G. Shanthikumar. 2005. Staffing decisions for heterogeneous workers with turnover. *Mathematical Methods of Operations Research* 62:499–514.

Aykin, T. 2000. A comparative evaluation of modeling approaches to the labor shift scheduling problem. *European Journal of Operational Research* 125:381–397.

Banker, R. D., A. Charnes, and W. W. Cooper. 1984. Some models for estimating -technical and scale inefficiency in data envelopment analysis. *Management Science* 30(9):1078–1092.

Bard, J. F., D. P. Morton, and Y. M. Wang. 2007. Workforce planning at USPS mail processing and distribution centers using stochastic optimization. *Annals Operations Research* 155:51–78.

Burke, E., P. Cowling, P. de Causmaecker, and G.V. Berghe. 2004. The state of the art of nurse rostering. *Journal of Scheduling* 7:441–499.

Charnes, A., W. W. Cooper, and E. Rhodes. 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2:429–444.

Chien, C.-F., F.Y. Lo and J.T. Lin. 2003. Using DEA to Measure the Relative Efficiency of the Service Center and Improve Operation Efficiency through Reorganization. *IEEE Transactions on Power Systems* 18 (1): 366-373.

Chien, C.-F., and L. Chen. Using Rough Set Theory to Recruit and Retain High-Potential Talents for Semiconductor Manufacturing. 2007. *IEEE Transactions on Semiconductor Manufacturing* 20(4):528-541.

Chien, C.-F., W. Chen, F.Y. Lo and Y. Lin. 2007. A Case Study to Evaluate the Productivity Changes of the Thermal Power Plants of the Taiwan Power Company. *IEEE Transactions on Energy Conversion* 22(3): 680-688.

Ernst, A.T., H. Jiang, M. Krishnamoorthy, and D. Sier. 2004. Staff scheduling and rostering: a review of applications, methods and models. *European Journal of Operational Research* 153:23–27.

Gans, N., and Y.-P. Zhou. 2002. Management learning and turnover in employee staffing. *Operations Research* 50(6):991–1006.

Holder, A. 2005. Navy personnel planning and the optimal partition. *Operations Research* 53(1): 77–89.

Leachman, R., S. Ding, and C.-F. Chien. 2007. Economic Efficiency Analysis of Wafer Fabrication. *IEEE Transactions on Automation Science and Engineering* 4(4):501–512.

McGuigan, J. R., R. C. Moyer, and F. H. deB Harris. 2002. *Managerial Economics: Applications, strategy, and tactics*. 9th ed. South-Western College Publishing, Cincinnati, OH.

Mundschenk, M., and A. Drexl. 2007. Workforce planning in the printing industry. *International Journal of Production Research* 45(20):4849–4872.

Pesch, E., and U. Tetzlaff. 2005. Scheduling personnel for press machines in the automotive industry. *Pacific Journal of Optimization* 1:545–564.

Thompson, G. M., and J. C. Goodale. 2006. Variable employee productivity in workforce scheduling. *European Journal of Operational Research*

170:376–390.

Zulch, G., S. Rottinger, and T. Vollstedt. 2004. A simulation approach for planning and re-assigning of personnel in manufacturing. *International Journal of Production Economic* 90:265–277.

AUTHOR BIOGRAPHIES

CHEN-FU CHIEN is a Professor in the Department of Industrial Engineering and Engineering Management and EMBA, National Tsing Hua University (NTHU). Since 2005, he has been on-leave to serve as the Deputy Director of Industrial Engineering Division in Taiwan Semiconductor Manufacturing Company. He received B.S. degree with double majors (with Phi Tao Phi Honor) in Industrial Engineering and Electrical Engineering from NTHU in 1990. He received M.S. in Industrial Engineering and Ph.D. of Operations Research and Decision Sciences from the University of Wisconsin-Madison in 1994 and 1996, respectively. He was a Fulbright Scholar in the Department of Industrial Engineering and Operations Research, UC Berkeley from 2002 to 2003 and received Executive Training in Harvard Business School in 2007. Dr. Chien is a Steering Committee Member of Industrial Engineering Division in National Science Council, Taiwan. He is Board Member of the Chinese Institute of Industrial Engineers (CIIE) and Chinese Institute of Decision Sciences. He is an Associate Editor of *IEEE Transactions on Automation Science and Engineering* and an Advisory Board Member of *OR Spectrum*. His research and development efforts focus on decision analysis, modeling and analysis for semiconductor manufacturing, dada mining, and manufacturing strategy. He received Outstanding Research Award and Tier 1 Principal Investigator (Top 3%) from National Science Council, Distinguished University-Industry Collaborative Research Award from Ministry of Education, Distinguished Young Industrial Engineer Award, Distinguished Young Faculty Research Award from NTHU, Best Paper Award from CIIE, Best Engineering Paper Award by Chinese Institute of Engineers, Taiwan. Email: <cfchien@mx.nthu.edu.tw>

WEN-CHIH CHEN received the B.S. degree in Industrial Engineering from the National Tsing Hua University, Taiwan in 1995. He received his Ph.D. from School of Industrial and Systems Engineering at Georgia Institute of Technology, USA, in 2003. Dr. Chen is an Assistant Professor at Department of Industrial Engineering and Management in National Chiao Tung University, Taiwan. His current research interests focus on performance evaluation and decision making analysis in semiconductor manufacturing.

SHAO-CHUNG HSU is a Ph.D. Candidate in the Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan. His research interests include data mining, statistical process control, and yield enhancement.