

The Impact of Process Noise on VLSI Process Improvement

Roger E. Bohn, *Member, IEEE*

Abstract—Process improvement is critical to commercial success in VLSI fabrication, especially during ramp-up. This paper investigates one of the factors—process noise—that drives the success of process improvement. Split-lot controlled experiments are vulnerable to confounding by experimental noise, caused by process variability. Fabs with low noise levels have a higher potential for learning (and hence improving their production processes) than high noise fabs.

Detailed probe yield data from five semiconductor fabs were examined to estimate process noise levels. A bootstrap simulation was used to estimate the error rates of identical controlled experiments conducted in each fab. Absolute noise levels were high for all but the best fabs, leading to lost learning. The magnitude of lost learning is estimated numerically; it ranges from ten percent to above one hundred percent of the theoretically possible learning in an experiment. In some cases, experiments are little better than coin flipping. Standard statistical methods are either expensive or ineffective for dealing with these high noise levels. Some alternative nonstatistical countermeasures are recommended.

I. INTRODUCTION

PROCESS IMPROVEMENT, critical for success in most VLSI manufacturing, depends on rapid technological learning. This is especially true during new product/process ramp up, when production volume must be increased rapidly. The speed and success of the ramp to high volume is determined by the rate at which problems and opportunities on the line are detected, diagnosed, and solved.

Process variability obscures true cause and effect relationships in the manufacturing process and makes process improvement and learning more difficult. If two fabs make the same product but have different process variability, one will be able to improve with less effort than the other. They will also have learning curves of different slopes [1], [2].

This paper analyzes process variability and concludes that it significantly affects semiconductor fabrication. It considers the impact of process variability on experimental error, and the consequences of those errors for process improvement. Experiments to improve yield can be confounded by experimental error, which can arise due to inherent process variability, measurement error, sampling error, analytical error, or careless mistakes [3].

Because of the complexity of processes and the potential for unforeseen side effects, most process changes are made carefully and systematically. Proposed changes are tested

through *engineering trials*. Typically an engineering trial is conducted as a *split lot experiment*, which compares two production methods. A regular production lot is split in half just before the step where the change is to be made. Half the wafers in the lot are processed in the conventional way, and half according to the proposed new recipe. The split halves are recombined and processed normally through the rest of the fab. At the end, the individual wafers are measured and the average measurements for each of the split lots calculated. Differences in the averages are due to the different recipes, plus the effects of process variability. This split lot procedure blocks (neutralizes) most of the between-lot variation, but does not block any of the within-lot variation. The within-lot variation is a major cause of experimental error and is the subject of this paper. Other situations in which process variability is a problem include acceptance sampling [4] (higher noise leads to larger sample sizes required for a given degree of statistical assurance) and statistical process control (real but small process disturbances are hidden by everyday high noise, leading to wide control limits) [5].

This paper is written for fab engineers, who have to cope with the effects of noise, and for researchers who are studying the improvement process. It contains three kinds of material, as delineated below.

- *A methodology* for measuring the effects of noise. This methodology makes no assumptions about the underlying probability distribution of defects; instead it uses live data from normal production.
- *Exploratory results* which quantify the magnitude of the problem in five fabs.
- *Practical implications and suggestions* for fab engineers.

Section II of the paper presents the empirical data used, showing the raw variability levels in fab yield for the five fabs. Section III shows how this variability translates into noise in experiments, and therefore into incorrect experimental outcomes. Sections IV and V translate these errors into a quantitative measure of lost process improvement due to noise. The final section of the paper discusses the inadequacy of conventional statistics for dealing with this problem, and presents some nonstatistical methods.

A. Prior Work on Yield Variability and Related Topics

This section reviews literature relating to yield variability in semiconductor manufacturing, and touches briefly on literature in related fields. Various authors have analyzed the nature of yields in VLSI integrated circuit manufacturing. An important

Manuscript received October 26, 1993; revised August 24, 1994.
The author is with the School of International Relations and Pacific Studies, University of California, San Diego, La Jolla, CA 92093-0519 USA.
IEEE Log Number 9412778.

observation in that literature is that the number of defective dice on a wafer does not follow a Poisson distribution, due to spatial clustering of defects. For example, the variance of defects may be ten times the mean, in contrast to the Poisson, which has the variance equal to the mean number of defects [6]. In consequence, standard formulas for probabilistic calculations involving yields can be quite erroneous [7]. Albin and Friedman propose the use of a Neyman Type-A distribution. They show that it leads to very different acceptance sampling plans [8] and control charts for detecting out-of-control processes [4].

Wein and various co-authors investigate the issue of yield variability and its impact on normal fab operations. A fab with constant yield (no matter how low) can be balanced and scheduled with a known ratio of machine capacity at different process stages. In contrast, varying yields can cause shifting bottlenecks and reduce overall fab performance by more than the average yield loss. For example, if a fab is making multiple chip types that are used as a set and sold in fixed proportion, variability in the yields leads to a decrease in the number of good sets produced [9]. Sometimes the variability can be turned to advantage. If yield is serially or spatially correlated and if yields are especially low on part of a wafer, it may not even be worth the time to test neighboring wafers or parts of wafers [10], [11].

Spanos [12] analyzes a different source of experimental errors in semiconductor fabrication—measurement error. He shows that ignoring measurement errors can lead to incorrect inferences about process performance. By re-analyzing a data set with and without allowance for measurement error, he shows that apparently significant process changes may be due to measurement error. The present paper, in contrast, analyzes the corresponding effect caused by process yield variability.

There is also a great deal of literature on the causes and effects of process variation. The key point in this literature is that process variation is inherently bad because it leads to out-of-specification conditions, hurting product quality. Thus quality improvement is in large part a struggle to reduce variability [13, chap. 11]. The additional role of variation in creating noise in the learning process is recognized but not emphasized in this literature.

The statistics literature considers extensively the issue of noise in experiments, but pays little attention to the role of process variation in causing that noise. The underlying process variation is taken as given; the role of statistics is to quantify the resulting noise level and to use statistical tools to reduce it [3], [14].

This paper differs from previous work on semiconductor yield variability in two principle respects. First, it is primarily empirical, attempting to establish the magnitude of this problem in a sample of actual fabs. Perhaps because of the highly confidential nature of yield data throughout the industry, previous work has been primarily theoretical and has not included empirical measures. Second, it emphasizes the impact of yield variability on learning, rather than on short-run operating, cost, or quality issues. It attempts to estimate the amount by which yield variability makes it more difficult to learn about and improve causes of yield loss.

TABLE I
SUMMARY OF DATA SOURCES

Fab code	Product maturity (approx.)	Name for data set in tables	Number of lots	Comments
A	1.5 years	A1.5	5	Same prod as fab G
	3 years	A3	5	
C	1 year	C1	11	
	1.5	C1.5	9	
	2	C2	10	
	2.5	C2.5	12	
	3	C3	13	
B	unknown	B	13	
F	1 year	FF	8	
G	pre-qualify	G	6	Same product as A

II. THE MAGNITUDE OF PROCESS NOISE (EXPLORATORY RESULTS)

The magnitude of process noise in actual semiconductor fabrication was investigated empirically. Five fabs provided production yield data on one product apiece. Two of the fabs provided data for multiple time periods, allowing us to look for trends over time. (Table I). Each was a high volume, multi-product MOS fabrication facility. All except *G* were U.S. fabs in a single company; *G* was a foreign subcontractor. Fabs *A* and *G* made the same product using the same process. All the products were medium to high volume, where high volume is thousands of lots and millions of completed chips per year (see Table I). This section presents basic descriptions of the noise as revealed by the data. The following section estimates the effects of the noise. All absolute yield data are disguised to avoid revealing proprietary information; only data on noise can be fully presented.

Data were provided by individual engineers in each fab.¹ The data consist of wafer-by-wafer probe yield counts (good dice per wafer) for every wafer in each lot. Thus, the data give a precise measure of probe yield and line yield.

Fig. 1 shows several months of standard tracking data used in fab *A* to track yields over time. Each point shows probe yield for one wafer, on a linear scale. Each column shows production during a single week; two randomly selected wafers are shown from each lot. Fig. 1 shows high levels of yield variation. Because of the way the data are displayed, Fig. 1 mixes between-lot and within-lot variability in a way impossible to disentangle from this data.²

Fig. 2 shows complete dot plots of probe yields of individual wafers in fab *C1* (fab *C*, one year after the beginning of production for that product). All lots completed the production process sequentially during the same week, and are for the same product in the same fab. Each column represents one lot, while each dot represents the yield of one wafer in that lot. The yields are arbitrarily scaled to protect confidentiality. Within-

¹Fab engineers were asked to provide data from a medium to high volume product. In at least one case (fab *A*), this biased the data since the engineer who selected the data chose from lots with high line yields. Thus, the results for this fab will understate the noise levels of experimental lots. In fab *FF*, a number of anomalous lots, which appeared to be experimental lots, were excluded. One lot was excluded for the same reason as fab *C*. One entire period of production (4 lots) was excluded for fab *A* because of a number of anomalies.

²It would have been better to average the two wafers in each lot; this would at least give an indication of lot-to-lot yield variability. A standard control chart would be even more useful and require no additional data.

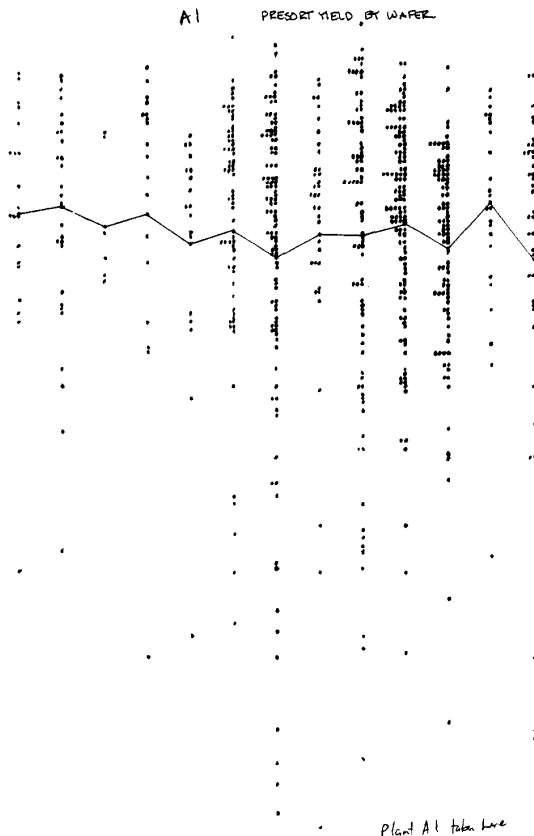


Fig. 1. Historical yield data for Fab A.

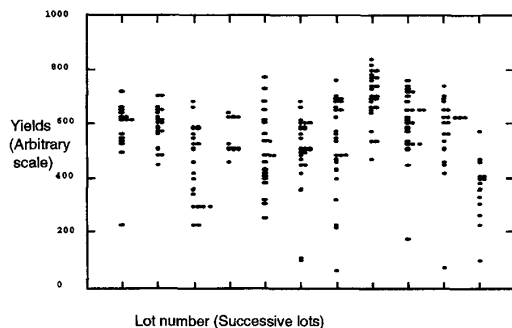


Fig. 2. Probe yields of individual wafers from fab C1.

lot variability in probe yield is the spread of each column. Between-lot variability is the difference among the columns.

The range of shapes shown in Fig. 2 is surprising, considering that all lots were produced under what should have been identical conditions. The mean yields, as well as the variance and skewness of yields, vary from lot to lot, suggesting that the underlying production process was not stable. A Bartlett test for homogeneity of group variances gave probability less than 0.005% that all ten lots from C1 had the same variances.

The within-lot standard deviation of production probe yields will prove to have the largest impact on experimental noise.

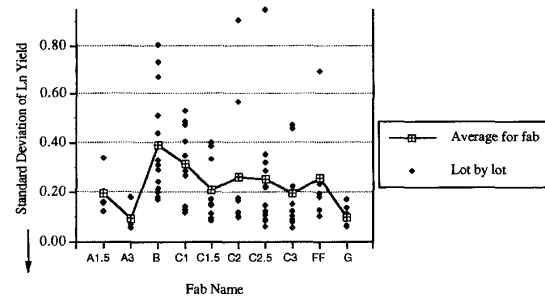


Fig. 3. Summary of noise by lot and fab.

Fig. 3 summarizes this measure for all five fabs. Each column corresponds to one of the fab/time combinations in Table I. For the rest of the paper, all yield data will be in natural logarithms to better indicate percent change in yields. Each point is the standard deviation of the log of probe yields of a single lot, which will be referred to as the “within-lot noise level”. Fig. 3 also shows the simple average of the within-lot noise levels in each fab.

Based on Fig. 3 we can make the following observations:

- Most lots within most fabs have high levels of within-lot variability.
- Noise levels vary considerably across fabs and time.
- The noise level varies greatly across lots in each fab. This is in addition to high lot-to-lot variation in mean yields.

These observations are consistent with manufacturing processes that are not under good process control. Whatever the cause, it is likely these factors will create high noise levels in experiments.

1) *Practical Implications:* These empirical results indicate that it is worthwhile for fab engineers to examine yield variability, both within and between lots. Comparisons with other plants making the same products may suggest targets to aim for. In addition, chronological tracking of data that is poorly selected, as in Fig. 1, throws away a lot of potentially useful information about what is happening in the fab. A deliberate sampling plan can easily show trends in mean and variability of yield, distinguish within-lot from between-lot variation, and even suggest hypotheses about causes of problems. Modern wafer tracking systems provide enough raw data for this purpose, but the fab engineering staff has to find ways of evaluating it without being overwhelmed by the details.

III. MAGNITUDE OF NOISE IN EXPERIMENTS

This section examines the effects of process noise on learning, by simulating experiments (engineering trials) using the data described in the previous section. Note that this section uses normal production data to simulate the conduct of experiments. This method gives a large data set which is comparable across fabs. In addition, the method used here, called bootstrapping, means that no assumptions are needed about the underlying distribution of yields. Even if different fabs or different part numbers have different types

of distributions (e.g., Neyman Type A versus Poisson), this methodology puts them all on a comparable basis.

We will assume that process changes multiply the yield by a constant which may be greater or less than 1.0. In other words, the underlying model of yields is an additive independent model in the log of yields.³ The model is therefore

$$Y_{\text{new}} = Y_{\text{old}} + \Delta Y + \varepsilon \quad (1)$$

where

- Y_{new} Probe yield after the process change.
- Y_{old} Original probe yield of the process.
- ΔY Change in average probe yield as a result of the experimental treatment (positive or negative). Larger is better.
- ε Noise in the probe yield. It can have any probability density distribution.

All quantities are measured in natural logarithms.

This section first gives a formal model of a split lot experiment, and shows how to simulate these experiments by *bootstrapping*. It then introduces the concept of a *power function*, which is a distribution-independent way of describing the results of any statistical procedure. Finally, it gives empirical results. These show that in some cases the results of experiments in these fabs are little better than flipping a coin.

A. Methodology

Learning is modeled as occurring through full-length split lot experiments. Each experiment consists of $2N$ wafers, N of which receive the experimental treatment at the critical process steps. The $2N$ wafers are processed as a single lot at all other process steps. The outcomes are measured at die probe. The standard test statistic for such an experiment is the difference in average yield between the two split groups. The larger the difference, the larger is the likely improvement from the new method. The test statistic is

$$\Delta Y_{\text{est}} = \left\{ \frac{1}{N_1} \sum_{i=1}^{N_1} Y_i \right\} - \left\{ \frac{1}{N_2} \sum_{i=1}^{N_2} Y_{N+i} \right\} \quad (2)$$

which deviates from the true effect of the treatment according to

$$\Delta Y_{\text{est}} = \Delta Y_{\text{true}} + \varepsilon_{\text{experiment}} \quad (3)$$

where

- ΔY_{est} Estimated yield improvement due to the new production method.
- ΔY_{true} Unknown true effect of the new production method.
- Y_i Log yield of the i th wafer; the first N wafers are the experimental group; the next N are the control group.

³This is a reasonable approximation for all but the largest changes, since total noise is the sum of noise in number of steps, while process changes affect only one or a few steps.

- N Initial sample size of each split group in the experiment. $N \leq 12$ since lot size in most fabs is 25.
- N_1 Number of wafers which survive in the experimental group; $N_1 \leq N$.
- N_2 Number of wafers which survive in the control group; $N_2 \leq N$.
- $\varepsilon_{\text{experiment}}$ Noise of the experiment, which depends on the process noise level and the number of wafers in the experiment.

If yields were distributed normally or according to another known distribution, and if line yields were 100% so that $N = N_1 = N_2$, we could use statistical theory to find the distribution of the experimental noise $\varepsilon_{\text{experiment}}$. However, using any single distribution to summarize the actual wafer by wafer data is risky. The lot-to-lot comparisons suggest that the manufacturing process parameters were not stable, and different distributions may apply in each fab. Also, the impact of missing wafers caused by line yield losses must be incorporated. This reduces the effective sample size below the nominal sample size N .

To evaluate the effectiveness of these experiments without assuming an underlying distribution function for probe yields, bootstrapping techniques were used to simulate what would have happened if experiments had been conducted on these lots in each fab [15]. Bootstrapping is a Monte Carlo method that uses limited amounts of empirical data to construct large samples which simulate experiments. The wafer by wafer probe yields from a single lot (discussed in Section II) were repeatedly sampled with replacement, to construct the two groups of $N/2$ wafers each, that would result from a single experiment. Wafers that did not survive the line yield were removed from each subsample. The test statistic, ΔY_{est} (difference of the average log yields), was then calculated for the case that $\Delta Y_{\text{true}} = 0$ (i.e., an experiment on a process change that has no effect). This gives the outcome of a single simulated experiment. This procedure was repeated 600 times for experiments with $N = 12$, and 2000 times for experiments with $N = 3$. Sampling was conducted equally from each lot of a particular fab/time period. Symmetry was then used to double these sample sizes to 1200 and 4000, respectively. These 5200 simulated experiments per fab/period form the basis for evaluating the error rates of real experiments in the fabs.

B. Power Functions

We will start with the simplest possible test criterion. If $\Delta Y_{\text{est}} > 0$, treat the new production method as better; otherwise, stay with the old production method. This decision rule serves as a starting point for more complex decision rules, which will be discussed later.

From the bootstrap data, we construct the power function $G(\Delta Y)$ of the hypothetical experiment. $G(\Delta Y)$ is the probability of choosing the new production method, if the true value of the change is ΔY . The power function gives a complete measure of an experiment's information content, and can be used to evaluate the experiment according to any criterion,

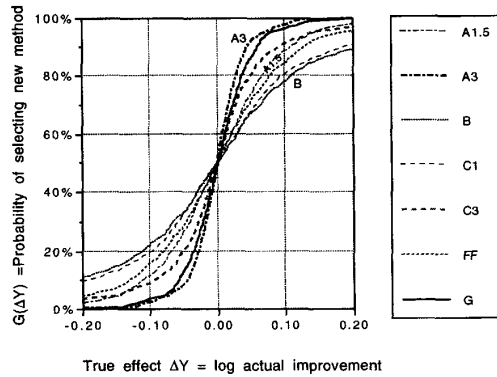


Fig. 4. Power functions for $N = 12$.

such as significance regions⁴. An ideal power function would rise vertically through $\Delta Y = 0$, with $G(-\beta) = 0$, $G(+\beta) = 1$, where β is arbitrarily small.

Fig. 4 shows the power functions for full lot experiments of $N = 12$ wafers per sample. Each line shows the probability (unknown to the experimenter) of accepting the new production method as a function of its true effect on yield ΔY_{true} . For any true value of log yield improvement, the height of the power function is the probability that the engineer will accept the hypothesis that $\Delta Y_{\text{true}} > 0$, i.e., that the new method is better.⁵ Probabilities of accepting inferior new methods ($\Delta Y_{\text{true}} < 0$) are given by symmetry.

For example, in fab *FF*, if the true value of a process change is $\Delta Y = 0.03$ (a 3% improvement in yield), the probability of accepting the new method is 63%, and the probability of rejecting it (type 2 error) is about 37%. Each power function is symmetric and passes through ($\Delta Y = 0$, prob. = 50%) because in this model, the engineer uses a symmetric test criterion. If the new method were in fact worse, with $\Delta Y_{\text{true}} = -0.03$, the probability of rejection would be 63% and the probability of acceptance (type 1 error) would be 37%. If the engineer sets a cutoff of $\Delta Y_{\text{est}} \geq \Delta Y_{\text{cutoff}} > 0$, in an effort to defeat the effect of noise, this would shift each curve to the right by ΔY_{cutoff} .⁶

C. Effect of Noise on Experimental Outcomes [Exploratory Results]

Learning in semiconductor manufacturing proceeds on the basis of multiple small improvements, in the neighborhood

⁴The procedure for estimating power functions from the bootstrap results is straightforward. Since in each simulated experiment the "true" value of the process change was the same, any variation in results of the experiments is due only to noise. Rank order each of the 1200 outcomes, from smallest to largest, and plot the results as in Fig. 4. Scale the vertical axis from 0 to 1, and the curve becomes a power function.

⁵Sampling error: Since Fig. 3 and associated tables are derived from bootstrap analysis with 1200 repetitions, they are subject to sampling errors. Since they are order statistics, these errors are largest in the tails. A visual assessment of the error is given by the waviness of the lines in Fig. 3; the true power functions are smooth. I also re-ran plant C1, 10 times. For $G(\Delta Y = 0.10)$, the runs had mean and s.d. of $0.793 \leq 0.10$.

⁶These probabilities are conditional on the experiment being successfully completed. There is also a small chance the experiment would have to be repeated due to too many wafers being lost in the production process (line yield). This is not shown in Fig. 4.

TABLE II
CONSEQUENCES OF WITHIN-LOT NOISE

Fab name	A1.5	A3	B	C1	C1.5	C2	C2.5	C3	FF	G
Avg. in-lot noise	0.196	0.094	0.391	0.315	0.209	0.261	0.251	0.196	0.256	0.100
Probability of missing process improvements:										
N=12, True $\Delta Y=0.01$	45.5%	37.5%	47.0%	47.0%	45.0%	43.5%	43.5%	43.5%	45.5%	40.5%
N=12, True $\Delta Y=0.05$	36.0%	18.5%	40.5%	40.0%	32.5%	32.5%	32.0%	28.5%	37.0%	24.5%
N=12, True $\Delta Y=0.10$	11.5%	2.5%	22.5%	19.5%	11.5%	12.0%	12.5%	9.0%	16.0%	4.0%
N= 3, True $\Delta Y=0.10$	22.5%	6.5%	34.0%	29.0%	20.0%	22.5%	21.0%	17.0%	25.0%	10.5%
Smallest effect that can be found with error probability $\leq 10\%$:										
for N=12	0.109	0.044	0.216	0.188	0.111	0.114	0.117	0.092	0.129	0.057
Number of lots required for 90% chance of detecting true process improvement of size .01 (approximate):										
Number of lots	148	41	630	444	221	392	394	184	365	50

of $0.01 \leq \Delta Y \leq 0.03$. This is the size of the signal being sought by the engineer; this is much smaller than the within-lot process noise of 0.10 and above, found in the empirical data. We now use the power functions to predict what will happen if experiments are run in each fab, on improvements of different sizes. Table II shows the predictions.

1) *Practical Implications:* We can make the following observations:

- The impacts of noise in most fabs were so large as to make the chance of overlooking process improvements (type 2 errors) quite high, except for very large improvements. To find a process change that has a 10% effect ($\Delta Y_{\text{true}} = 0.10$) is quite rare. But in fab *B* with a sample size of $N = 12$, even such a large effect would be missed in an experiment more than 20% of the time. Only fabs *A3*, *G*, and *C3* have probabilities of error below 10% for a change of 0.10. Experiments on process changes with $\Delta Y = 0.05$ have error rates ranging from 18% to 40%. As this model is formulated, 50% is the highest possible error rate, so fabs *B* and *C1* do little better than pure chance. None of the fabs does much better than pure chance for $\Delta Y = 0.01$.
- The consequences of noise differ considerably across fabs and time. Therefore, decision rules and experimental designs should be re-examined periodically. In particular, engineers should measure noise levels in new products and other situations with potentially high noise levels.
- All results are considerably worse for experiments conducted with samples of $N = 3$. In fact the noise levels are so high that all experiments should be run with $N \geq 12$.

IV. DECISION RULES AND THE VALUE OF EXPERIMENTS (METHODOLOGY)

After process engineers run a trial that shows a small but positive impact ΔY_{est} from a new manufacturing method, they must decide whether to change the process permanently or leave it alone. The statistical approach to this decision is to use a significance test to evaluate the probability that $\Delta Y_{\text{true}} > 0$, and to make the change iff this probability is higher than some value such as 90%. This approach is not good in high noise environments. As shown in Table II, because of high experimental noise in most fabs, this criterion will lead to high levels of ΔY_{cutoff} and thereby will have a high chance of Type 2 errors (rejecting genuine improvements). How much

is lost because of such errors? What is the value maximizing decision rule for making the choice? What is the resulting aggregate rate of process improvement? By how much does it hurt to be running trials in a high noise fab instead of a low noise fab?

This section shows how to answer these questions, using a simple economic model for measuring the learning from an experiment. First it defines several rules for deciding whether to make a process change, called decision rules. Then it shows how to quantify the effectiveness of each experiment under different decision rules. It ends with a brief discussion of sequential experimentation. Section V evaluates these questions numerically for each fab in the data set.

A. Decision Rules: When to Change the Process

Once a trial has been run, the decision whether to make the process change should be based on whether the observed value of process improvement, ΔY_{est} , is above or below some cutoff value, ΔY_{cutoff} . (To simplify notation, Δ will be used in place of ΔY in Sections IV and V.) The statistical approach is to choose a cutoff $\Delta_{\text{cutoff}}(\alpha)$ which solves

$$1 - \alpha = \Pr[\Delta_{\text{true}} \geq 0 / \Delta_{\text{est}} = \Delta_{\text{cutoff}}(\alpha)] \quad (4)$$

where α is the significance level of the test, the allowable probability of type I errors (false positives) and where $\Pr[x/b]$ is read "the probability of x , given information set b ". The experimenter chooses the level of α based on subjective judgments of the relative costs of Type I and Type II errors. The statistical decision rule then follows.

Rule S: Adopt new method iff:

$$\Delta_{\text{est}} \geq \Delta_{\text{cutoff}}(\alpha). \quad (5)$$

Although it is commonly used, (4) for selecting a cutoff value is not the best possible because it does not consider the economic value of process improvement directly. In contrast, the decision theoretic rule chooses Δ_{cutoff} to maximize the expected value of the experiment. Let K be the net present value in dollars of future production of the product if the process change is not made. This depends on the chip's selling price, packaging cost, expected market life, and the discount rate. Let C be the fixed cost of implementing a process change. C should include the value of downtime while the change is being implemented, the opportunity cost of worker time implementing the change, and costs of any disruption caused by the change. From this we can calculate a "breakeven yield improvement" such that the cost of the process change exactly balances the value of the change.

$$\Delta_{\text{breakeven}} = C/K. \quad (6)$$

Rule B: Adopt new method iff:

$$\Delta_{\text{est}} \geq \Delta_{\text{breakeven}} \equiv \Delta_{\text{cutoff,B}}. \quad (7)$$

Typical values might be on the order of $C = \$0.5$ million, $K = \$100$ million, giving a cutoff of 0.5% yield improvement.

If the engineer has previous information about how large Δ is likely to be (such as from prior trials, or results in similar situations), they can do even better than Rule B. The prior knowledge can be included with the experimental results by a process known as *Bayesian updating* [16].

Rule W: Adopt new method iff:

$$\Delta_{\text{est}} \geq \Delta_{\text{breakeven}} + \frac{\sigma_{\text{est}}^2}{\sigma_{\text{prior}}^2} \times [\Delta_{\text{breakeven}} - \Delta_{\text{prior}}] \equiv \Delta_{\text{cutoff,W}}. \quad (8)$$

where Δ_{prior} is the estimated process improvement from information before the trial, σ_{est}^2 is the variance in the estimate of Δ , and σ_{prior}^2 is the variance in the information from prior sources.

If $\Delta_{\text{breakeven}} = \Delta_{\text{prior}}$ then Rule W reduces to Rule B. If the experimental noise level is high, as in some of the fabs in this paper, σ_{est}^2 will be large and the prior knowledge will receive a heavy weight. A rather counterintuitive result is that if $\Delta_{\text{breakeven}} < \Delta_{\text{prior}}$ then $\Delta_{\text{cutoff,W}} < 0$ and the process change may be accepted even if data from the experiment alone suggest that the process change is negative.

For typical levels of criterion α such as 5%, Δ_{cutoff} calculated according to Rules W or B will be smaller than $\Delta_{\text{cutoff}}(\alpha)$ calculated in Rule S, leading to more process changes under the optimization criterion than under the traditional statistical criterion. In other words, the traditional statistical rule is overly conservative. One interpretation of this difference is that $\Delta_{\text{breakeven}}$ serves to bias the engineer against small changes, which would more likely have been rejected by a significance test. The level of α in (5) is chosen based on the engineer's subjective assessment of the cost of false positives versus false negatives, which we have instead done using an economic model in (6). If the decision Rule W leads to "too many" small process changes, that argues that change cost C has been underestimated. A second interpretation is that the standard statistical approach is to do nothing unless a change is "almost sure to be right," while the optimization approach tries to get the best answer "in the aggregate," over many trials, even though some of those trials will give the wrong outcome.

Numerical evaluation of different decision rules (discussed later) indicates that Rule W is in fact not optimal given the actual distributions of $G(\Delta)$. A better decision rule is Halfway between Rule W and Rule B.

Rule H: Adopt new method iff:

$$\Delta_{\text{est}} \geq \Delta_{\text{cutoff,H}} = .5(\Delta_{\text{cutoff,W}} + \Delta_{\text{cutoff,B}}). \quad (9)$$

B. Lost Yield Improvement Due to Noise

Any decision rule will give errors some of the time because of experimental noise. We provide a criterion for comparing decision rules and different experiments. Let $G^*(\Delta)$ be the probability of deciding to make the change if $\Delta_{\text{true}} = \Delta$, where the * indicates that some decision rule with $\Delta_{\text{cutoff}} \neq 0$ is in use. Then the expected value of an experiment measured

in log yield improvement is

$$\begin{aligned} EV_{\text{real}} &\equiv \frac{1}{K} \int_{-\infty}^{\infty} v(\Delta) G^*(\Delta) f(\Delta) d\Delta \\ &\times [\text{Prob experiment completed}] \\ &= \left\{ \int_{-\infty}^{\infty} \Delta G^*(\Delta) f(\Delta) d\Delta - \Delta_{\text{breakeven}} \right. \\ &\quad \left. \times \int_{-\infty}^{\infty} G^*(\Delta) f(\Delta) d\Delta \right\} \\ &\times [\text{Prob experiment completed}] \end{aligned} \quad (10)$$

where $f(\Delta)$ is the prior probability density function for the size of process improvements Δ_{true} . [Prob. experiment completed] is $1 -$ (the chance that too many wafers will be lost due to line yields). EV_{real} is measured in logs; $K \times EV_{\text{real}}$ is the expected net present value improvement per experiment in dollars. To evaluate (10) it is easy to show that

$$G^*(\Delta) = G(\Delta - \Delta_{\text{cutoff}}) \quad (11)$$

where the $G(\cdot)$ function was estimated by bootstrapping in Section III.

We compare this with what would happen in the absence of noise. $v(\Delta)$ and $f(\Delta)$ remain the same, but $G^*(\Delta)$ is replaced by a 0-1 step function at $\Delta_{\text{breakeven}}$. The resulting expected value of a perfect experiment (EVPE) is

$$\begin{aligned} EVPE &= \left\{ \int_{\Delta_{\text{breakeven}}}^{\infty} \Delta f(\Delta) d\Delta - \Delta_{\text{breakeven}} \right. \\ &\quad \left. \times \int_{\Delta_{\text{breakeven}}}^{\infty} f(\Delta) d\Delta \right\}. \end{aligned} \quad (12)$$

For each fab we can calculate by how much process improvement is slowed down due to noise as a ratio of lost learning (LL)

$$LL = 1 - \{EV_{\text{real}}/EVPE\}. \quad (13)$$

LL is independent of the economic scale factor K ; it depends mainly on the noise level in the fab, and on $f(\Delta)$. $f(\Delta)$ is the relative density of large and small yield changes resulting from different process changes. It depends on the maturity of the technology and the degree of insight of the engineer choosing what trial to run. $f(\Delta)$ cannot be observed directly; large numbers of real experiments would be needed to estimate its true shape. An article about a prototype production facility at Hewlett Packard gave the following comments about $f(\Delta)$:

It is estimated that an experiment designed to improve a process will most likely result in a yield improvement from -2.5% to 12.5% [percent, not percentage points].

It can be further estimated that some small, but positive yield improvement (typically 5%) for each completed engineering experiment is the most probable result. [17]

This suggests that $f(\Delta)$ is symmetric with a mean of $+0.05$ and a standard deviation of about 0.03 . However, as a process matures, $f(\Delta)$ will move to the left since the most promising hypotheses are investigated first. All but two of the data sets in this paper are for products one year old or more. Anecdotal, process improvements of more than a few percent are rare for such processes.

TABLE III
LEARNING DETERMINANTS EVALUATED

Method	Variable	Low Value	Mid Value	High Value
Lower noise fab	Fab ID	All 10 fab/periods in Table I		
Idea quality, ¹⁷ implement. cost	Δ_{prior} $\Delta_{\text{breakeven}}$	0	-----	.02
Idea extremism	σ_{prior}	.03	-----	.06
Sample size	N	3	-----	12
Decision rule	Δ_{cutoff}	H	B	S

V. BETTER NOISE CONTROL AND ITS BENEFITS [EXPLORATORY RESULTS]

This section compares several approaches to dealing with noise. Learning from trials can be improved by a number of methods, most of which require engineering effort to learn and set up. Therefore, it is useful to get an idea of their impact before pursuing them. Specific improvements modeled here include:

- being in a lower noise fab,
- decreasing implementation costs C : decreases $\Delta_{\text{breakeven}}$,
- proposing better initial hypotheses for improvements: raises Δ_{prior} ,
- proposing more radical ideas for improvement (e.g., bigger shifts away from standard set-points): raises σ_{prior} ,
- using a better decision rule: Rule H or W instead of Rule S,
- increasing the sample size N .

All of these are under a degree of management and engineer control, although some (decision rules) are easier to change than others (better hypotheses).

Based on the methodology in the previous sections, this section estimates the expected values of learning and lost learning in each fab under different conditions. This is done by assuming values for the parameters of $\Delta_{\text{breakeven}}$ and $f(\Delta)$, then applying the bootstrap-derived power function in Section III to the valuation equations of Section IV.⁷ This section, using an economic model, analyzes the following cases, as shown in Table III.⁸

The outcome measures are EVPE, EV_{real} , and lost learning as a percent of EVPE. A full factorial design was used to analyze the variables in Table III except for sample size; $N = 3$ was examined only for a few lower noise fabs. In most cases, decision Rule B was almost as good as Rule H, so it will not be reported extensively. Table IV reports detailed results for Fab G ($N = 3, 12$) and Fab C1 ($N = 12$). LL is lost learning as a percentage of EVPE (smaller is better), as shown in Table IV.

1) *Practical Implications:* According to Table IV

- All learning determinants have a substantial impact on learning in some situations. Doing everything "well" can reduce lost learning to as low as 13% of potential

⁷ Aside from numerical accuracy issues, using the power functions for this is equivalent to bootstrap simulation of the experiments, then measuring each experiment directly and averaging to get expected values.

⁸ We will assume that the distribution of ideas $f(\Delta)$ is normal (mean is Δ_{prior} , standard deviation is σ_{prior}) so that EVPE is fully determined by specifying $(\Delta_{\text{prior}} - \Delta_{\text{breakeven}})$ and σ_{prior} . With other assumptions about $f(\Delta)$ somewhat different numerical estimates of EVPE would be reached, but the lost learning levels would be approximately the same. Note that this is the only place in the paper where a specific distribution is assumed.

TABLE IV
LOST LEARNING RESULTS FOR 2 FABs

Scenario				Results (Larger EV _{real} and smaller LL are better)							
Case ID	Δ prior- Abkveven	sigma prior	Decision rule	EVPE	Fab G, N=12		Fab G, N=3		Fab C1, N=12		
					EV _{real}	LL	EV _{real}	LL	EV _{real}	LL	
1	0	0.03	H	.0120	.0068	43.5%	.0046	61.8%	.0031	74.1%	
2	0	0.03	B	.0120	.0068	43.5%	.0046	61.8%	.0031	74.1%	
3	0	0.03	S	.0120	.0043	63.8%	.0019	83.7%	.0007	94.5%	
4	0.02	0.03	H	.0245	.0206	16.0%	.0195	20.4%	.0180	26.7%	
5	0.02	0.03	B	.0245	.0192	21.8%	.0164	33.1%	.0145	41.1%	
6	0.02	0.03	S	.0245	.0119	51.4%	.0064	73.9%	.0034	86.1%	
7	0.02	0.06	H	.0353	.0304	13.6%	.0265	24.8%	.0226	35.9%	
8	0.02	0.06	B	.0353	.0303	14.1%	.0260	26.1%	.0221	37.4%	
9	0.02	0.06	S	.0353	.0248	29.6%	.0153	56.7%	.0070	80.1%	
10	0	0.06	H	.0239	.0189	21.1%	.0148	38.3%	.0110	53.9%	
11	0	0.06	B	.0239	.0189	21.1%	.0148	38.3%	.0110	53.9%	
12	0	0.06	S	.0239	.0154	35.8%	.0087	63.5%	.0035	85.6%	

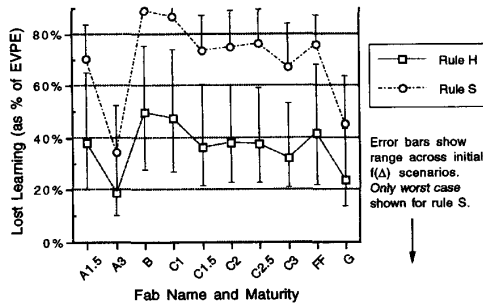


Fig. 5. Lost learning by Fab ($N = 12$).

learning (run 7, decision Rule H, Fab G , $N = 12$). Doing it “poorly” can almost wipe out the value of experimentation (run 7, decision Rule S, Fab $C1$ has lost learning of 80%).

- Moving from a high noise fab ($C1$) to a low noise fab (G , $N = 12$) gives a roughly 30% point reduction in lost learning.
- Decreasing the sample size is like moving to a higher noise fab and increases the lost learning.
- Using the S (Statistical) decision rule instead of H or B is costly, in many cases doubling the lost learning. Using a better rule is an easy change to make (selecting Δ_{cutoff} using economic criteria rather than purely statistical criteria).
- The $f(\Delta)$ initial hypothesis parameters have a substantial effect on EVPE and EV_{real}, and on lost learning. Notice that more radical ideas raise the value of possible learning. (Compare EVPE in run 7 and run 4, for example.)

Fig. 5 looks in more detail at the effect of fab noise levels on lost learning. Each entry is the mean across the four combinations of $f(\Delta) : \{\Delta_{\text{prior}} - \Delta_{\text{breakeven}} = \text{Low, High}; \sigma_{\text{prior}} = \text{Low, High}\}$. There is a 2.5:1 ratio between best and worst fabs, and their ranking is consistent with the ordering of their power functions in Fig. 4 and noise levels in Fig. 3. Fig. 5 shows that statistical decision rule S with $\alpha = 10\%$ is completely dominated by Rule H , and it sacrifices more than 60% of the potential learning in 8 of the 10 fabs/time periods. Even Rule H is poor in an absolute sense, with average lost learning close to 40% in the same 8 fabs.

The only way to get lost learning below 30% according to this analysis is to use a good decision rule (H or B), use full lots (24 wafers), and either be in a low noise fab ($A3$ or G)

or have a good initial hypothesis $\Delta_{\text{prior}} - \Delta_{\text{breakeven}} \geq 0.02$. The first two conditions can be fulfilled with sufficient effort, but the last condition requires either extended work on noise reduction or good luck.

VI. GENERAL CONCLUSIONS: FAST LEARNING DESPITE NOISE

Given that process noise levels in most of the fabs were large enough to cause high experimental noise and consequent lost learning, what can be done? Standard statistical methods alone cannot overcome the high noise. An approximate calculation indicates that from 40 to 400 lots per experiment would be needed to meet standard statistical criteria in each fab.⁹ In the process of visiting a number of fabs I observed four classes of countermeasures, used with varying degrees of intensity and expertise among different companies, fabs, and engineers:

- *statistical methods*, such as larger sample size, experimental design, and better data analysis,
- *transforming the problem by changing the outcome variable*; examples of this approach are short-loop experiments, defect analysis, and creation of special purpose test sites,
- *transforming the problem by changing the site or nature of experimentation*; examples of this approach include wafer tracking [19] and laboratory investigations in place of engineering trials [20],
- *reduction of process noise levels* in the fab, by methods such as SPC, to detect changes [5].

These methods can be very effective, especially if used in concert. This section looks briefly at each in turn.

1) *Statistical Methods*: Statistical methods of dealing with noise take the core definition of the trial as fixed (what is measured, how, and where), but use mathematical methods to improve the design or analysis. Routine statistical methods, such as significance tests, were discussed earlier. More advanced statistical methods are often recommended by statisticians.

Multi-lot trials: This brute force approach requires enough lots to have enough wafers in the sample to overcome the noise. The drawbacks are very high costs, and long delays since all the lots have to finish processing before the results can be analyzed. A variant on this approach is *sequential experimentation*. In this approach one or a few lots are run and analyzed. If the results are clearly good (high ΔY_{est}) the method is adopted. If the results are clearly bad, it is rejected. If ΔY_{est} is intermediate, another set of lots are run under the same conditions, and the results are averaged. This approach reduces costs but greatly increases the duration of the trial and therefore reduces the rate of learning, compared with multi-lot trials.

Fractional factorial design: Fractional factorial designs (such as Taguchi’s orthogonal arrays) allow multiple variables to be investigated simultaneously, with only a slight increase in effective noise level for any of the variables. However they do not solve the basic problem of high noise. They also require very careful operational control of the experiment, since each

⁹Last line of Table II. The underlying calculations are presented in [18].

lot now requires a number of setups, and any processing error can cause misleading results.

Bayesian decision rules: Bayesian methods compare the costs of gathering more information with the costs of making different kinds of errors. Rule H, analyzed in the previous section, is an example of a Bayesian method, and it gave greatly improved learning per experiment. However lost learning was still quite high in most of the scenarios.

2) *Transforming the Problem by Changing the Outcome Variable:* Each sequential step in wafer fabrication has its own random defect mechanisms, and adds its own noise to the final yield. This paper has looked at full length controlled experiments, in which the dependent (outcome) variable is the die yield at wafer probe. Since die yield is the key economic driver, such experiments are directly relevant to process economic performance. Yet it is often useful to look at intermediate yield drivers rather than final yield, in order to reduce noise levels.¹⁰

Causal analysis of defects: Engineers test process changes because of an underlying causal model of how defects occur. They hypothesize that a specific change will reduce one or a few defect mechanisms. If these defect types can be measured directly on wafers at the end of the trial, then this removes all of the process variability caused by other defect mechanisms. This can be as much as a 100× improvement in noise levels. However, it leaves the engineer vulnerable to unanticipated and unmeasured side effects from the trial, which could reduce the overall yield. Causal analysis of defects is also a slower and more labor-intensive process than automated probe yield testing, limiting the number of trials that can be done in this way.

Short-loop experiments: Because short-loop trials remove wafers before completion [22], they cut out the noise from downstream steps. In addition, they usually look only at one or a few defect mechanisms, thereby picking up the noise reduction benefits of causal analysis.

Test structures: Test structures represent an enhanced approach to short-loop trials [23].

Within-wafer effects: A referee points out that many process changes have differential effects on different parts of the wafer. In this situation, weighted sampling of different parts of the wafer will help.

Drawbacks and obstacles to changing variables: The big (potential) drawback of changing variables is loss of *fidelity*. Measuring only certain yield drivers may overlook other problems which are created or exacerbated by the process change being tested. This is especially likely when working with new and novel processes. In addition, it is often not clear what the causal relationships are between intermediate and final variables. The process engineer may not know of a reliable way of measuring a particular problem, earlier in the process.

In addition, measuring variables other than probe yield is slower. It may be quite labor intensive, reducing the feasible sample size, or increasing costs, substantially.

¹⁰ Several published case studies of directed process improvement illustrate various of these approaches to noise mitigation [19], [21]; [21] is particularly interesting in its use of multiple methods to reduce the effects of noise.

3) *Transforming the Problem by Changing the Site or Nature of Experiments:* Normal production fabs have a number of demands on them, and they are not optimized for experimentation. Therefore some companies use *pilot lines* for much of their learning. However, pilot lines are subject to continual change, use equipment which is not fully debugged, and run a high rate of controlled experiments which are themselves disruptive. Hence noise control cannot be taken for granted in pilot lines or development fabs.

Another approach which is becoming cheaper and more popular is the use of *natural experiments* based on detailed analysis of data from ongoing production. Wafer tracking systems [19] increasingly incorporate the hooks for collecting such data [19], [24]. Because of the lack of control groups in the data (no split lots), the noise level per lot is higher in natural experiments than in controlled experiments. However, extremely large sample sizes are possible, canceling some of the noise.

Natural experiments are theoretically suspect for learning because of a number of statistical problems. For example, if x and yield are both caused by an unobserved process condition, then no matter how high the correlation between x and yield in the data, increasing x will not necessarily improve yield, and in fact could hurt it.¹¹ Despite this inability to prove causality, natural experiments are still an excellent way to develop ideas for further testing.

4) *Reduction of Process Noise in the Fab:* The large difference in noise levels between fabs A and G , which made the same product, confirms that direct reduction of process noise is feasible. Many of the sources of manufacturing variability can and should be removed, through methods such as total quality maintenance and SPC to spot process excursions [5]. Some fab practices, intended to improve performance, actually work to increase noise. An important example is the practice of "tweaking" equipment and recipes frequently. Deming, among others, has argued strongly against frequent process adjustments on the grounds that they hurt average performance.

The concept of robust design (of both products and promises) offers considerable promise for reducing yield variability as well as increasing mean yield [25]. In a similar way, more conservative chip design rules, while they may reduce maximum die yield, can reduce process noise and therefore increase the potential rate of learning.¹²

VII. CONCLUSION

Using data from a sample of fabs and products, this paper has estimated the magnitude of improvements possible by using different statistical and nonstatistical methods of

¹¹ Box *et al.* [3] refer to natural experiments as the use of "happstance data," and show numerous potential problems with their value.

¹² Among the issues not dealt with in this paper are the root causes of the high noise levels, and the reasons why noise is not addressed more directly and solved by fab management. Noise in final probe yields is due at a micro level to specific machine/process variability at specific process steps. But the large differences in noise among fabs using similar equipment and processes suggests deeper levels of causality. Phenomena of this complexity cannot be captured by studying experimentation in isolation from other aspects of fab operation.

noise mitigation. Standard statistical methods, especially by themselves, are ineffective, or costly and slow. Non-statistical, Bayesian and combined methods are much more effective. Taken together, they can reduce lost learning from as high as 80% to below 20% of potential learning.

The payoff to dealing better with noise in experimentation is faster performance improvement, or "accelerated learning curve progress." It is one class of methods for increasing the amount learned from each learning activity or cycle [22]. Some of the changes recommended here can be implemented easily and individually, while others require fab-wide efforts.

Although the specific results in different fabs will vary, the same principles should work, and the same methods should be usable. The empirical data in this paper covers only four products and five fabs. Within this small data set, there is considerable variation in noise levels across products, fabs, and product maturity. This variation is itself one of the results of the paper, and suggests that noise is a serious problem. A larger and more comprehensive sample of companies, fabs, products, and time series will be needed to see whether the high and variable noise levels are a general feature of VLSI fabrication¹³.

APPENDIX

DERIVATION OF DECISION RULES

This appendix derives several rules for deciding whether to implement a process change on a permanent basis, based on both economic and statistical issues. The Statistical decision rule derived in Section IV is

Rule S Adopt new method iff $\Delta_{\text{est}} \geq \Delta_{\text{cutoff}}(\alpha)$. (5)

The power functions in Fig. 4 allow construction of such tests by locating the value of Δ which solves

$$G(\Delta_{\text{cutoff}}) = 1 - \alpha. \quad (A1)$$

In contrast, the decision theoretic rule chooses Δ_{cutoff} to maximize the expected value of the experiment. To derive this rule, let $v(\Delta)$ be the net present value in dollars of a yield change of size Δ . Let $\pi(t)$ be the marginal value to the fab of an additional die at time t . If the market is competitive,

$$\pi(t) = (\text{Wholesale selling price}) \\ - (\text{Marginal packaging cost}). \quad (15)$$

Let $Q(t)$ be the base case quantity of chips produced at t , with no process change. Then an adequate model for $v(\Delta)$ is

¹³Interested readers are encouraged to contact the author, who is willing to analyze data they collect, or provide his software so they can conduct their own analyses.

¹⁴(A1) is a more sophisticated version of the standard t -test for equality of means, which assumes that ε is distributed according to a normal distribution [14].

¹⁵In (15), if the market is not fully competitive, marginal revenue should be substituted for selling price. The distinction between price and marginal revenue is discussed in any microeconomics textbook.

the linear model

$$v(\Delta) = \Delta \times K - C \\ K = \int_0^T e^{-rt} \pi(t) Q(t) dt \quad (A2)$$

where T is the time horizon, r is the relevant discount rate, C is the fixed cost of implementing a process change, and K is the value of future production. Solving (A1) for $\Delta_{\text{breakeven}}$ such that $v(\Delta_{\text{breakeven}}) = 0$ gives

$$\Delta_{\text{breakeven}} = C/K. \quad (A3)$$

The objective function is to *maximize the expected net present value of all process changes*. This is done by making the change iff:

$$E[v(\Delta)/\text{available information}] \geq 0. \quad (A4)$$

Then, decision rule (A2) becomes *make the process change if the expected value of Δ is greater than $\Delta_{\text{breakeven}}$* .

Adopt new method iff:

$$E(\Delta/\text{experimental results}) \geq \Delta_{\text{breakeven}}. \quad (A5)$$

An obvious value for $E(\Delta/\text{experimental results})$ is ΔY_{est} defined in (2). This leads to decision Rule B (for breakeven): Rule B: Adopt new method iff

$$\Delta_{\text{est}} \geq \Delta_{\text{breakeven}} = \Delta_{\text{cutoff,B}}. \quad (7)$$

If the experimenter has previous information about how large Δ is likely to be, they can do better. Let $f(\Delta)$ be the prior probability density function for the size of the process improvement Δ_{true} . Then $E(\Delta/\text{experimental results})$ is a weighted average of Δ_{est} and the pre-experiment value of Δ , $\Delta_{\text{prior}} \equiv \int f(\Delta) d\Delta$. The weight depends on the relative uncertainties of $f(\Delta)$ and the experimental results. If both $f(\Delta)$ and $G(\Delta)$ are normal, this has a closed form solution.¹⁶ Let $f(\Delta)$ be normal (mean is Δ_{prior} , standard deviation is σ_{prior}), and assume for now that the distribution $G(\Delta)$ of experimental results is normal (mean is Δ_{est} , standard deviation σ_{est}). Then by a process known as *Bayesian updating* [16]

$$E(\Delta/\text{all experimental data}) = \frac{\frac{\Delta_{\text{prior}}}{\sigma_{\text{prior}}^2} + \frac{\Delta_{\text{est}}}{\sigma_{\text{est}}^2}}{1/\sigma_{\text{prior}}^2 + 1/\sigma_{\text{est}}^2}. \quad (A6)$$

Substituting (A4) into (A3) and solving for the level of Δ_{est} , which will produce equality gives the decision Rule W (for Weighted).

Rule W: Adopt new method iff:

$$\Delta_{\text{est}} \geq \Delta_{\text{breakeven}} + \frac{\sigma_{\text{est}}^2}{\sigma_{\text{prior}}^2} \times \\ [\Delta_{\text{breakeven}} - \Delta_{\text{prior}}] \equiv \Delta_{\text{cutoff,W}} \quad (8)$$

¹⁶The shape of $f(\Delta)$ is determined by expert judgement (engineering estimate). If no prior information is available, this is equivalent to assuming that $f(\Delta)$ has infinite variance in the following equations.

ACKNOWLEDGMENT

The author is grateful for valuable comments on previous versions of this paper as provided by Dr. Richard Dehmel, Dr. Andy Urquhart, Prof. Larry Wein, Prof. Roy Welsch, and an anonymous referee. Prof. Michael Watkins assisted with data analysis. Data collection was funded by the Harvard Business School and an anonymous company. The author remains solely responsible for the paper's omissions and errors.

REFERENCES

- [1] R. Jaikumar and R. E. Bohn, "A dynamic approach to operations management: An alternative to static optimization," *Int. J. Product. Econ.*, vol. 27, no. 3, pp. 265-282, 1992.
- [2] W. I. Zangwill and P. B. Kantor, "Toward a theory of continuous improvement," Graduate School of Business, University of Chicago, 1993.
- [3] G. E. P. Box, W. G. Hunter, and J. S. Hunter, *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*. New York: Wiley, 1978.
- [4] D. J. Friedman and S. Albin, "Clustered defects in IC fabrication: Impact on process control charts," *IEEE Trans. Semiconduct. Manufact.*, vol. 4, no. 1, pp. 36-42, 1991.
- [5] C. J. Spanos, "Statistical process control in semiconductor manufacturing," *Proc. IEEE*, vol. 80, no. 6, pp. 818-830, 1992.
- [6] C. H. Stapper, "The defect-sensitivity effect of memory chips," *IEEE J. Solid-State Circuits*, vol. SC-21, no. 1, pp. 193-198, 1986.
- [7] ———, "Fact and fiction in yield modeling," *Microelectron. J.*, vol. 20, no. 1-2, pp. 129-151, 1989.
- [8] S. Albin and D. J. Friedman, "The impact of clustered defect distributions in IC fabrication," *Management Sci.*, vol. 35, no. 9, pp. 1066-1078, 1989.
- [9] F. Avram and L. M. Wein, "A product design problem in semiconductor manufacturing," *Oper. Res.*, vol. 40, no. 5, pp. 986-998, 1992.
- [10] J. Ou and L. M. Wein, *Sequential Screening in Semiconductor Manufacturing I: Exploiting Lot-to-Lot Variability* (Working paper). Cambridge, MA: Sloan School, MIT, 1992.
- [11] M. D. Longtin, L. M. Wein, and R. E. Welsch, *Sequential Screening in Semiconductor Manufacturing II: Exploiting Spatial Dependence* (Working paper). Cambridge, MA: Sloan School, MIT, 1992.
- [12] C. J. Spanos, "Statistical significance of error-corrupted IC measurements," *IEEE Trans. Semiconduct. Manufact.*, vol. 2, no. 1, pp. 23-28, 1989.
- [13] W. E. Deming, *Out of the Crisis*. Cambridge, MA: MIT, Center for Advanced Engineering Study, 1986.
- [14] R. V. Hogg and J. Ledolter, *Engineering Statistics*. New York: Macmillan, 1987.
- [15] B. Efron and G. Gong, "A leisurely look at the bootstrap, the jackknife and cross-validation," *Amer. Statistician*, vol. 37, no. 1, pp. 36-48, 1983.
- [16] H. Raiffa and R. Schlaifer, *Applied Statistical Decision Theory*. Cambridge, MA: MIT Press.
- [17] M. W. Brooksby, P. L. Castro, and F. L. Hanson, "Benefits of quick-turnaround integrated circuit processing," *The Hewlett Packard J.*, vol. 32, no. 6, 1981.
- [18] R. E. Bohn, *Noise and Learning in Semiconductor Manufacturing (Working paper)*, University of California San Diego, 1993.
- [19] G. M. Scher, "Wafer tracking comes of age," *Semiconduct. Int.*, pp. 126-131, 1991.
- [20] R. E. Jones and T. C. Mele, "Use of screening and response surface experimental designs for development of a 0.5-um CMOS self-aligned titanium silicide process," *IEEE Trans. Semiconduct. Manufact.*, vol. 4, no. 4, pp. 281-287, 1991.
- [21] M. Hiatt and A. Urquhart, "Experimental technique for resist process evaluation," *Semiconduct. Int.*, pp. 146-151, 1987.
- [22] D. Dance and R. Jarvis, "Using yield models to accelerate learning curve progress," *IEEE Trans. Semiconduct. Manufact.*, vol. 5, no. 1, pp. 41-46, 1992.
- [23] S. Magdo and M. Gupta, "Using a test site for the rapid introduction of 32-kb bipolar RAM," *IEEE Trans. Semiconduct. Manufact.*, vol. 5, no. 1, pp. 62-67, 1992.
- [24] D. A. Hodges and W. C. Holton, *CIM at Japan's Best VLSI Manufacturers*, (Working paper), Dept. EECS, UC Berkeley, 1988.
- [25] M. S. Phadke, *Quality Engineering Using Robust Design*. Englewood Cliffs: Prentice Hall, 1989.



Roger E. Bohn (M'82) received the B.A. degree in applied mathematics from Harvard University, *summa cum laude*, in 1976, and the Ph.D. degree in management from M.I.T. in 1982.

He is now on the faculty of the University of California, San Diego. From 1982 to 1989, he taught at the Harvard Business School. His research interests are in the general areas of technology and operations management, and his present research concerns technological knowledge and learning in manufacturing. He has co-authored various papers

in the IEEE TRANSACTIONS ON POWER SYSTEMS, on the topic of real-time pricing of electricity, and is the co-author of *Spot Pricing of Electricity* (Kluwer Academic Press: Norwood, MA).