
Inspection for Circuit Board Assembly

Author(s): Phillipe B. Chevalier and Lawrence M. Wein

Source: *Management Science*, Vol. 43, No. 9 (Sep., 1997), pp. 1198-1213

Published by: [INFORMS](#)

Stable URL: <http://www.jstor.org/stable/2634633>

Accessed: 20/09/2011 05:23

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



INFORMS is collaborating with JSTOR to digitize, preserve and extend access to *Management Science*.

<http://www.jstor.org>

Inspection for Circuit Board Assembly

Phillipe B. Chevalier • Lawrence M. Wein

Institut d'Administration et de Gestion, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142

Several stages of tests are typically performed in circuit board assembly, and each test consists of one or more noisy measurements. We consider the problem of jointly optimizing the allocation of inspection and the testing policy in a system with a predefined inspection configuration; that is, at which stages should a board be inspected, and at these stages, whether to accept or reject a board based on noisy test measurements. The objective is to minimize the expected costs for testing, repair, and defective items shipped to customers. We analyze the problem and document an application of the model to an industrial facility. Since we were unable to gather all the necessary data, the model was applied in a limited and piecemeal fashion. Nevertheless, the proposed policy significantly improves upon the facility's historical policy. (*Quality Control; Inspection Decisions; Electronics Manufacturing*)

1. Introduction

This paper considers the problem of inspection in a circuit board assembly plant, which came to our attention while working with a Hewlett-Packard facility. Recent technological advances have given rise to circuit boards of increasing complexity, and consequently testing has become one of the most challenging aspects of circuit board manufacturing. Inspection costs can account for over half of the total manufacturing cost, and hence optimizing the utilization of inspection resources is a crucial task.

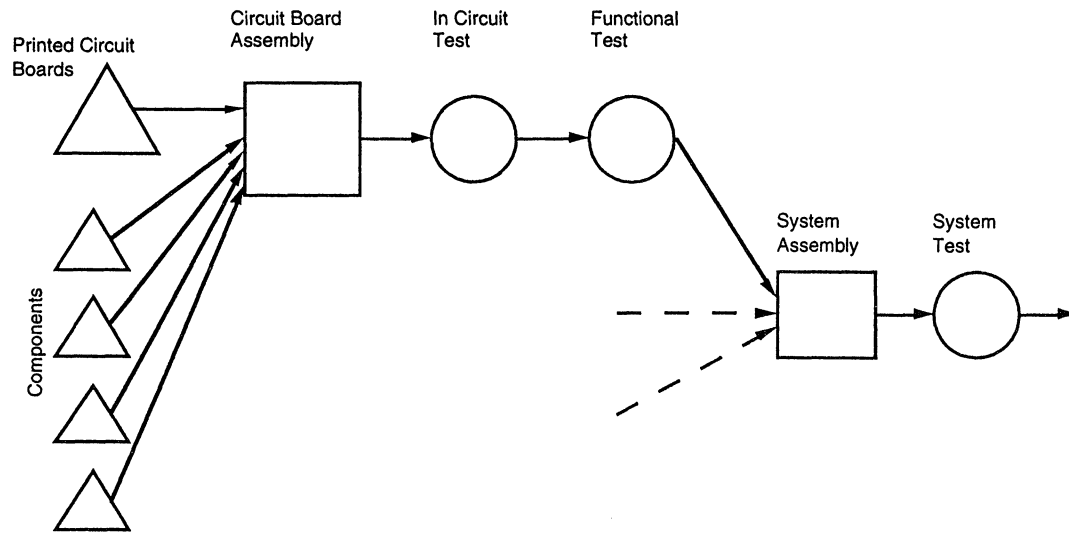
At the Hewlett-Packard facility, the assembly of circuit boards is performed in a single manufacturing stage, which is followed by several successive inspection stages; these stages will be described in detail in §2. A circuit board can fail in many different ways, and each inspection stage at the facility is designed to detect certain types of defects; typically, later inspection stages can detect more subtle problems. The inspection process at each stage consists of testing and repair. In testing, one or more measurements are taken from a board, and a decision is made whether to accept or reject the board; because the measurements are subject to error, Type I (false reject) and Type II (false accept) errors can occur during this classification. The defect on a rejected board is identified and corrected during repair. Testing costs

vary by stage, and repair costs depend upon the nature of the defect and perhaps the stage of detection.

The facility managers had two key concerns. First, they had the vague sense that too much testing was being performed. Second, they had no systematic procedure for accepting or rejecting boards at each stage as a function of the noisy test measurements. Consequently, this paper formulates and analyzes a mathematical optimization problem consisting of two interrelated decisions. Given an existing inspection configuration, which is typically dictated by the existing technology and board design, the first decision is to decide at which stage(s) to inspect a board; this problem is known in the literature as the *inspection allocation problem*. At the stages where inspection is performed, the *testing policy* decides whether to accept or reject each board based on the noisy measurements obtained from the test. The joint specification of inspection allocation and testing will be referred to as an *inspection policy*. The optimization problem is to find an inspection policy to minimize the total expected cost, which includes costs for testing, repair, and defective items leaving the facility. Since we assume that every defective board is repaired, no scrapping cost is included.

We analyze this problem in §3 and document an application of the model at the Hewlett-Packard facility in

Figure 1 Flow Chart of Hewlett-Packard's Circuit Board Assembly Process



§4. Despite gathering large amounts of data at this facility, we were unable to estimate the model inputs well enough to derive the optimal inspection policy. Consequently, a suboptimal policy, which is described in §4, was derived and applied in a fragmentary manner. Our limited numerical results, however, suggest that the suboptimal inspection policy easily outperforms the facility's historical inspection policy. Concluding remarks, including a description of the facility's use of the model, are contained in §5.

Many papers have been published on the optimal allocation of inspection in multistage serial systems. Early studies on this problem (see, for example, White 1966 and Lindsay and Bishop 1967) assumed perfect inspection (i.e., no Type I or Type II errors). Later papers allowed for imperfect inspection (Eppen and Hurst 1974, Yum and McDowell 1981, Garcia-Diaz et al. 1984), where one determines the number of times a test should be repeated. A survey of work published on inspection allocation can be found in Raz (1986). Recently, Villalobos et al. (1992) studied a dynamic version of the same problem, where inspection of an item at a particular stage can depend on the result of the inspection of that item at previous stages. Our paper appears to be the first in this research area to address the presence of distinct defect types, the joint optimization of inspection allo-

cation and testing, and the application of a model to an industrial facility.

2. Problem Description and Formulation

We begin by describing the circuit board assembly and testing process at the Hewlett-Packard facility, which is depicted in Figure 1. The primary manufacturing stage is circuit board assembly, where all components are soldered onto the printed circuit boards. Assembled boards then undergo *in circuit* testing, where every component on the circuit board is individually measured. The in circuit test is designed to ensure that the correct component is in the correct location, and that the soldering of components does not lead to any short (electrical contact where there should be none) or open (no contact where there should be) circuits. Although this test allows for easy problem identification, it is not very accurate because of the difficulty in isolating the influence of the various components. A *functional* test is performed next, where output measurements are taken from a board that is submitted to inputs that simulate its working conditions. This test can detect more subtle problems, such as out-of-specification components or a soldering that gives a poor contact, but requires much

more qualified personnel. Moreover, when an error arises, the source of the defect can be difficult to identify. Since it is not feasible to simulate all the possible working conditions on a complex board, the functional test is incomplete. After the boards have been assembled into a final product, a comprehensive *system* test is undertaken.

Although the Hewlett-Packard assembly process is quite typical, many variants of the configuration displayed in Figure 1 are possible; consequently, our mathematical model will allow more generality than depicted in this figure. Generally, downstream tests are more comprehensive, but require greater effort to diagnose and repair problems. The situation in traditional manufacturing settings, where additional work is performed between tests, is conceptually identical: Successive tests are increasingly comprehensive (if each test measures the cumulative functionality) and the cost of disassembling a unit to repair a defect increases as the unit progresses along the production line.

Approximately 50 different board types are manufactured at the Hewlett-Packard facility, and these boards go into various models of a single line of final products; the number of boards produced of each type is relatively low, ranging from 1,000 to 10,000 per year. For simplicity, our analysis considers only one type of circuit board; the multitype case is discussed briefly in §5.

Since a board can fail in a multitude of ways, it is useful for modeling purposes to aggregate all possible errors into a distinct number of *defect types*. Our aggregation is such that defects within a given type have similar repair costs at each stage, and are nearly identical with respect to the ease of difficulty of detection at each inspection stage. Six defect types are used in the case study in §4: Opens and shorts, missing or wrong components, defective components, two different soldering defect types (one is much harder to detect than the other), and miscellaneous.

Let N denote the number of inspection stages and I be the number of defect types in our model. Recall that downstream tests are more comprehensive, in that more subtle defects tend to be detectable only in later stages. We model this state of affairs by defining δ_{in} to be the expected number of defects of type i per board that are detectable only from stage n onwards; consequently, a defect that is detectable at one stage is also detectable at

every subsequent stage. Hence, it is *as if*, on average, δ_{in} new defects of type i appear per board at stage n . Detectable defects are still subject to Type II errors however. We let $\delta_{i,N+1}$ be the expected number of type i defects per board that are undetectable by the inspection system. The quantities δ_{in} are estimated from historical data later in the paper.

Up to several thousand measurements are gathered at each testing stage, and we assume that each measurement can detect defects of only one defect type. Let K_{in} denote the number of measurements taken at stage n that are intended to detect defects of type i . This quantity is dictated by the existing testing configuration and is not a decision variable. At the Hewlett-Packard facility, K_{in} ranged from about 20 to several hundred over the various stages and defect types. For in circuit testing, at least one measurement is taken per component (e.g., one might measure the resistance between two different pins on a component), and each measurement can detect exactly one defect. In the later tests, however, as many as 20 defects within a defect type can potentially be detected from a single measurement. After analyzing data and speaking to engineers, we concluded that the probability of having two distinct defects arising from a single measurement was vanishingly small; consequently, our model allows for at most one defect to arise from a single measurement. Conversely, we also assume that a defect cannot influence more than one measurement at a *particular* stage of inspection. This assumption is partially justified by the existing testing procedure: The measurements at an inspection stage are taken in a specified order and, if a measurement detects a defect, then the defect is repaired before subsequent test measurements at this stage are taken. Moreover, different measurements at a given stage typically investigate different portions of a board or different types of functionality, and so a defect that is undetected by its direct measurement (either because of a Type II error or a lack of inspection) will not influence other measurements at this stage. Of course, if a Type II error occurs or if inspection is not undertaken at a particular stage, then measurements at *downstream* stages may be affected.

The test measurements are subject to error, and we let

$$v_{ink}^m = v_{ink}^t + \epsilon_{ink} \quad \text{for all } i = 1, \dots, I,$$

$$n = 1, \dots, N \text{ and } k = 1, \dots, K_{in},$$

where v_{ink}^m is the measured value, v_{ink}^t is the true value, and ϵ_{ink} is the measurement error. We assume that the random variable ϵ_{ink} is independent of v_{ink}^t and is independent of the other measurement errors. Our most difficult modeling task is the specification of v_{ink}^t ; the true value v_{ink}^t presumably depends upon the inspection policy at previous stages, including the acceptance intervals that are employed for each test measurement related to type i defects. The exact nature of this dependency can be very complex since the same quantity is not measured at different stages. Even if a model of this dependency was developed, data collection would be extremely difficult. Typically, the only available data are the noisy measurements under the existing inspection policy; experiments would probably need to be performed under each possible testing policy to gather the necessary data. We return to these empirical issues in §4.

We make the key assumption that the distribution of the random variable v_{ink}^t depends on the single quantity ρ_{in} , which is the expected number of detectable defects of type i present on a board when it enters inspection stage n . Although we do not have a strong justification for this assumption, for the purposes of tractability in both the analysis and the estimation, we need to assume that the distribution of v_{ink}^t depends on only a single quantity. Such a quantity needs to reflect the true quality of the board at stage n with respect to type i defects, and the quantity ρ_{in} is a natural surrogate. Notice that ρ_{in} equals δ_{in} , which was defined as the expected number of type i defects that become detectable at stage n , plus defects that were not repaired from earlier stages (as a result of bypassed inspections or Type II testing errors). The quantity ρ_{in} will be calculated later in this section in terms of the upstream inspection policy, and will represent the system state in the dynamic program in §3. For future use, we let $\rho_{i0} = 0$ and $\rho_{i,N+1}$ equal the expected number of type i defects per board received by a customer.

Let $\xi_{ink}(x; \rho_{in})$ be the probability density function of the true value v_{ink}^t given ρ_{in} , and $e_{ink}(x)$ denote the density function of the measurement error ϵ_{ink} . Let $f_{ink}(x; \rho_{in})$ be the density function of the measured values v_{ink}^m given

ρ_{in} , so that $f_{ink}(x; \rho_{in}) = \int_{-\infty}^{\infty} \xi_{ink}(x - y; \rho_{in}) e_{ink}(y) dy$. Let G_{ink} denote the interval in which the true value v_{ink}^t should reside to ensure that the board is functioning properly. A key role in our analysis will be played by $p_{ink}(v_{ink}^m; \rho_{in})$, which is the conditional probability that the true value v_{ink}^t is inside this interval, given the measurement value v_{ink}^m and the quantity ρ_{in} ; that is,

$$p_{ink}(v_{ink}^m; \rho_{in}) = \frac{\int_{G_{ink}} \xi_{ink}(y; \rho_{in}) e_{ink}(v_{ink}^m - y) dy}{f_{ink}(v_{ink}^m; \rho_{in})}. \quad (1)$$

Our subsequent analysis requires the functions $f_{ink}(x; \rho_{in})$ and $p_{ink}(x; \rho_{in})$ to be continuous and unimodal in x . If we assume that the true value density function $\xi_{ink}(x; \rho_{in})$ and the measurement error density $e_{ink}(x)$ are unimodal in x , then it follows that $f_{ink}(x; \rho_{in})$ is unimodal in x . Although we have been unable to prove that these unimodality assumptions also imply the unimodality of $p_{ink}(x; \rho_{in})$, it seems natural that $p_{ink}(x; \rho_{in})$ would increase as x gets closer to the interval G_{ink} ; moreover, our numerical results confirmed the unimodality of $p_{ink}(x; \rho_{in})$.

The testing policy T_n at stage n is defined by the intervals $[L_{ink}, U_{ink}]$, $i = 1, \dots, I$ and $k = 1, \dots, K_{in}$, where a board is accepted at stage n if $v_{ink}^m \in [L_{ink}, U_{ink}]$ for all $i = 1, \dots, I$ and $k = 1, \dots, K_{in}$, and is rejected otherwise. Let $\alpha_{in}(T_n, \rho_{in})$ be the expected number of false defects of type i per board at stage n under testing policy T_n , and let $\beta_{in}(T_n, \rho_{in})$ be the expected number of detectable defects of type i present on a board at stage n that are not detected at that stage. It follows that

$$\alpha_{in}(T_n, \rho_{in}) = \sum_{k=1}^{K_{in}} \left[\int_{-\infty}^{L_{ink}} p_{ink}(x, \rho_{in}) f_{ink}(x; \rho_{in}) dx + \int_{U_{ink}}^{\infty} p_{ink}(x, \rho_{in}) f_{ink}(x; \rho_{in}) dx \right], \quad (2)$$

and

$$\beta_{in}(T_n, \rho_{in}) = \sum_{k=1}^{K_{in}} \int_{L_{ink}}^{U_{ink}} (1 - p_{ink}(x, \rho_{in})) f_{ink}(x; \rho_{in}) dx. \quad (3)$$

In addition to deciding upon a testing policy, we also need to choose the inspection allocation policy at each stage, which specifies whether or not to test the boards at that stage. Let the binary decision variable z_n equal 1 if boards are inspected at stage n and 0 otherwise, where

we define $z_0 = 0$ for notational convenience. The objective is to minimize the total expected cost of testing, repair, and defects leaving the facility; this quantity will often be referred to as the *total inspection cost*. Although the inspection policy affects the work-in-process inventory levels at the various stages, inventory holding costs are not incorporated here because the congestion effects at the Hewlett-Packard facility were negligible; readers are referred to Tang (1991) for an inspection allocation problem with queueing costs. To express the total inspection cost, we define y_{in} to be the expected number of type i defects to be repaired at stage n . For $i = 1, \dots, I$, it follows that

$$y_{in} = z_n[\rho_{in} - \beta_{in}(T_n, \rho_{in}) + \alpha_{in}(T_n, \rho_{in})] \quad (4)$$

for $n = 1, \dots, N$,

and

$$\rho_{in} = z_{n-1}\beta_{i,n-1}(T_{n-1}, \rho_{i,n-1}) + (1 - z_{n-1})\rho_{i,n-1} + \delta_{in} \quad (5)$$

for $n = 1, \dots, N + 1$.

We consider a per unit testing cost t_n at stage n , which typically includes operator time, test engineering, equipment depreciation and maintenance, and various overhead costs. A possible benefit, or negative cost, of testing is that it will lead to quicker learning and hence process improvements; however, in our case study, we did not attempt to incorporate this benefit into the per unit testing cost. Also, no fixed testing cost is included in the model. The repair cost r_{in} is the total cost incurred to diagnose and repair a type i defect on a board at stage n . We assume that all defective boards are repaired, and that the same repair cost is incurred, whether the defect is a real defect or a false defect. This assumption is reasonable because the diagnostic cost overwhelms the actual repair cost in this setting; in particular, before declaring that an out-of-specification measurement is due to a false defect, all possible problem sources must be investigated. The cost f of a defect on a board that leaves the plant includes the cost of a field repair, the cost of the analysis and repair of the defective board that comes back to the plant, and a cost measuring the customer's loss of goodwill. We assume that the cost is per defect and not per defective system, which simplifies the analysis. In any case, if there are two or more defects in a system, then it is unlikely that these defects would be

detected by the customer at the same time. Moreover, since the number of defects leaving the plant is very low, multiple defects on the same system are very unlikely; consequently, the results obtained would be very similar if a cost was incurred per defective system. We also assume that the cost f does not depend on the type of defect on a board. Although the more general case can be easily accommodated, this assumption seems practical, since the costs of lost goodwill and visiting the repair site dominate the other costs (by at least an order of magnitude), and are independent of the defect type.

Our optimization problem is to choose the inspection allocation policy z_n and the testing policy T_n for $n = 1, \dots, N$ to minimize

$$\sum_{n=1}^N \left(t_n z_n + \sum_{i=1}^I r_{in} y_{in} \right) + f \sum_{i=1}^I \rho_{i,N+1}, \quad (6)$$

subject to Equations (4)–(5). Notice that the probability distributions of the random variables ρ_{in} and δ_{in} do not matter because of the linear cost structure.

3. Analysis

In this section, we analyze problem (4)–(6). The testing problem is addressed in §3.1 and the inspection allocation policy is numerically derived in §3.2.

3.1. The Testing Policy

Problem (4)–(6) can be formulated as a dynamic program. If we let $J_n(\rho_{1n}, \dots, \rho_{In})$ denote the minimum expected total inspection cost from stage n onwards, where ρ_{in} is the expected number of detectable type i defects on a board arriving at stage n , then the dynamic programming optimality equations are

$$\begin{aligned} & J_n(\rho_{1n}, \dots, \rho_{In}) \\ &= \text{Min} \left[J_{n+1}(\rho_{1n} + \delta_{1,n+1}, \dots, \rho_{In} + \delta_{I,n+1}), \right. \\ & \quad t_n + \min_{T_n} \left[\sum_{i=1}^I (\rho_{in} - \beta_{in}(T_n, \rho_{in}) + \alpha_{in}(T_n, \rho_{in})) r_{in} \right. \\ & \quad + J_{n+1}(\beta_{1n}(T_n, \rho_{1n}) + \delta_{1,n+1}, \dots, \beta_{In}(T_n, \rho_{In}) \\ & \quad \left. \left. + \delta_{I,n+1}) \right] \right] \quad \text{for } n = 1, \dots, N, \end{aligned} \quad (7)$$

and

$$J_{N+1}(\rho_{1,N+1}, \dots, \rho_{I,N+1}) = f \sum_{i=1}^I \rho_{i,N+1}. \quad (8)$$

The first minimization on the right-hand side of Equation (7) represents the inspection allocation option; the first option corresponds to no inspection at stage n ($z_n = 0$) and the second option corresponds to inspection at stage n ($z_n = 1$); in the latter case, a second minimization is performed to determine the optimal testing policy. The second minimization can be rewritten more concisely as

$$\min_{T_n} \{h_n(\alpha_{1n}(T_n, \rho_{1n}), \dots, \alpha_{In}(T_n, \rho_{In})) + g_n(\beta_{1n}(T_n, \rho_{1n}), \dots, \beta_{In}(T_n, \rho_{In}))\}, \quad (9)$$

where

$$h_n(x_1, \dots, x_I) = \sum_{i=1}^I x_i r_{in} \quad \text{and} \quad (10)$$

$$g_n(x_1, \dots, x_I) = J_{n+1}(x_1 + \delta_{1,n+1}, \dots, x_I + \delta_{I,n+1}) + \sum_{i=1}^I (\rho_{in} - x_i) r_{in}. \quad (11)$$

If $\alpha_{in}(T_n, \rho_{in})$ and $\beta_{in}(T_n, \rho_{in})$ in (9) are replaced with the expressions in Equations (2)–(3), then the derivative of this function with respect to the upper acceptance limit U_{ink} is

$$-\frac{\partial h_n}{\partial x_i}(\alpha_{1n}, \dots, \alpha_{In}) p_{ink}(U_{ink}, \rho_{in}) f_{ink}(U_{ink}; \rho_{in}) + \frac{\partial g_n}{\partial x_i}(\beta_{1n}, \dots, \beta_{In})(1 - p_{ink}(U_{ink}, \rho_{in})) f_{ink}(U_{ink}; \rho_{in}),$$

where $\partial/\partial x_i$ is the partial derivative of the i th component and, to simplify notation, α_{in} stands for $\alpha_{in}(T_n, \rho_{in})$ and β_{in} stands for $\beta_{in}(T_n, \rho_{in})$. Setting this expression equal to zero and using (10)–(11), we obtain

$$p_{ink}(U_{ink}, \rho_{in}) = 1 - \frac{r_{in}}{\frac{\partial J_{n+1}}{\partial x_i}(\beta_{1n} + \delta_{1,n+1}, \dots, \beta_{In} + \delta_{I,n+1})}. \quad (12)$$

The second derivative of the objective function (9) will be positive if $(\partial f_{ink}(x; \rho_{in})/\partial x)|_{x=U_{ink}} < 0$ and $(\partial p_{ink}(x, \rho_{in})/\partial x)|_{x=U_{ink}} < 0$. One would expect the second-

order conditions to follow from our assumptions that $f_{ink}(x; \rho_{in})$ and $p_{ink}(x, \rho_{in})$ are continuous and unimodal in x , since the upper limit is at a point where both the frequency of measurement and the probability that a measurement corresponds to a valid board are decreasing. We will return to this convexity issue during the case study in §4.

The same argument can be used to show that the optimal lower acceptance limit satisfies

$$p_{ink}(L_{ink}, \rho_{in}) = p_{ink}(U_{ink}, \rho_{in}) = 1 - \frac{r_{in}}{\frac{\partial J_{n+1}}{\partial x_i}(\beta_{1n} + \delta_{1,n+1}, \dots, \beta_{In} + \delta_{I,n+1})}. \quad (13)$$

The second derivative of the objective function (9) will be positive if $(\partial f_{ink}(x; \rho_{in})/\partial x)|_{x=L_{ink}} > 0$ and $(\partial p_{ink}(x, \rho_{in})/\partial x)|_{x=L_{ink}} > 0$. Again, the second-order conditions are consistent with our intuition. Expression (13) has an intuitive meaning: the probability that a component is bad at the acceptance and rejection cutoff points should be equal to the marginal cost of an additional repair at the current stage divided by the marginal cost of an additional defect entering the next stage. At this cost ratio increases, it becomes less costly to accept defective boards, and the acceptance region is increased.

In our case analysis in §4.3, relation (13) will be refined further by making several simplifying assumptions that can be partially justified by data gathered from the Hewlett-Packard facility.

3.2. The Inspection Allocation Policy

Equation (13) expresses the optimal testing policy (L_{ink}, U_{ink}) in terms of the first derivative of the value function. Ideally, we would like to substitute this solution back into the inner minimization in (7) to express the optimality equations solely in terms of the value function, so that these equations could be solved using standard methods. Unfortunately, this approach cannot be used because $p_{ink}(v_{ink}^m, \rho_{in})$ in (1) is not readily invertible and β_{in} in the denominator of (13) is actually a function of the testing policy, as can be seen in (3), where L_{ink} and U_{ink} appear in the limits of integration.

Instead, we start by discretizing the I -dimensional space of ρ_{in} for each stage. Notice that the right side of (13) is constant for each measurement $k = 1, \dots, K_{in}$. For each fixed value of ρ_{in} , we let p_{in} stand for $p_{ink}(L_{ink}, \rho_{in})$ or $p_{ink}(U_{ink}, \rho_{in})$, and rewrite (13) as

$$p_{in} = 1 - \frac{r_{in}}{\frac{\partial J_{n+1}}{\partial x_i} (\beta_{1n}(p_{1n}) + \delta_{1,n+1}, \dots, \beta_{In}(p_{In}) + \delta_{I,n+1})}$$

for $i = 1, \dots, I$, (14)

where the dependence of β_{in} on the testing policy is reintroduced into the notation. Starting from stage $n = N$, we numerically solve this set of I nonlinear equations for p_{in} ; since $f_{ink}(x; \rho_{in})$ and $p_{ink}(x, \rho_{in})$ are continuous and unimodal in x , $\beta_{in}(p_{in})$ is a decreasing function and J_{n+1} is concave, and hence the solution should not be too difficult to derive. With the solution p_{in} of (14) in hand, we derive the optimal testing policy (L_{ink}, U_{ink}) from (1), calculate $\alpha_{in}(T_n, \rho_{in})$ and $\beta_{in}(T_n, \rho_i)$ from (2)–(3), and solve (7) for both the optimal inspection allocation decision at stage n and the value function J_n . The last quantity allows us to solve (14) for stage $n - 1$ and carry on in an iterative fashion to $n = 1$.

4. Case Study

We now describe the application of this model to the Hewlett-Packard facility. The assembly and inspection process at this facility was detailed in §2. The nature of the final product, which cannot be revealed, requires very strict tolerances on the circuit boards and on their components, which partially explains why this facility tested all boards at all stages. Although boards of a particular type all undergo exactly the same tests, testing procedures across board types were not consistent and were highly dependent upon the particular engineers in charge. The managers at this facility estimate that the total cost of inspection represents about half of the total manufacturing cost.

As mentioned in the Introduction, our model's data requirements proved to be too demanding to successfully apply our results at the Hewlett-Packard facility; consequently, we were forced to derive suboptimal testing and inspection allocation policies and apply them in a fragmentary manner. Section 4.1 describes these barriers to implementation (which should also be

viewed as limitations of the model), and gives a brief description of our strategy to overcome these problems. Section 4.2 outlines the parameter estimation procedures. The suboptimal policies and numerical results for the testing policy and inspection allocation policy are given in §§4.3 and 4.4, respectively.

4.1. Barriers to Implementation

We encountered two key obstacles in implementing the results of §3. The first difficulty is due to a key characteristic of multistage inspection systems: The quality of an item at a particular stage depends upon the inspection policy employed at previous stages. In §2, we made the rather crude assumption that the probability distribution for the true measurement value v_{ink}^t was solely a function of ρ_{in} , which is the expected number of detectable type i defects on the board at stage n . Estimation of this probability distribution is challenging, because ρ_{in} is a function of the upstream testing and allocation policy, whereas all the historical data are from the existing system that employed a *single* inspection policy. Consequently, rather than estimating $\xi_{ink}(x; \rho_{in})$ for all ρ_{in} , we were only able to estimate $\xi_{ink}(x; \rho_{in})$ under the facility's historical ρ_{in} values. This obstacle prevented us from deriving the optimal testing and inspection allocation policies for this facility, and our efforts turned instead to deriving a suboptimal testing policy using (13) with the fixed historical values of ρ_{in} ; as a result, we will hereafter suppress the dependence on ρ_{in} in much of our notation.

The second barrier to implementation, which was not foreseen, did not allow us to achieve this lesser goal, however: After gathering data and estimating parameters, we found that our estimate for $p_{ink}(x)$ was not accurate enough to derive a reliable testing policy via (13). Unlike our first barrier, this problem is not inherent to the model and may not occur in settings with more data and/or less process noise. Hence, we propose a simpler testing policy in §4.3 that has several desirable properties; this testing policy is based on $\xi_{ink}(x; \rho_{in})$ under the historical ρ_{in} values.

The optimal inspection allocation policy satisfying (7) depends upon the Type I and Type II testing errors, which themselves are functions of the testing policy and the ρ_{in} values. Because of our first barrier to implementation described above, we could only estimate the test-

ing errors α_{in} and β_{in} under the historical testing policy and historical ρ_{in} values. By assuming a simple relationship between error rates at each stage and the expected number of defects at the stage, we used the historical error rate estimates to heuristically extrapolate α_{in} and β_{in} as functions of ρ_{in} under the fixed historical testing policy. These extrapolated functions allowed us to derive an inspection allocation policy given the facility's traditional testing policy; details can be found in §4.4.

Finally, it would be an enormous task to collect data and derive an inspection policy for all 50 board types. Instead, we applied our model in a very limited manner to exhibit its effectiveness, and left the appropriate software with the Hewlett-Packard engineers, who carried out the implementation.

4.2. Parameter Estimation

Because the parameter estimation procedures are rather involved and their description would disrupt the continuity of the paper, we relegate this material to the appendix; in this subsection, we merely state the parameters that were estimated: The cost parameters t_n , r_{in} , and f ; δ_{in} , which is the expected number of new type i defects per board appearing at stage n ; α_{in} and β_{in} , which are the Type I and Type II testing errors under the historical testing policy and historical ρ_{in} values; and the means μ_ξ and μ_e and standard deviations σ_ξ and σ_e of the true value v_{ink}^t and the measurement error ϵ_{ink} , respectively.

4.3. Testing Policy

The optimal testing policy (13) cannot be calculated because we have been unable to estimate $\xi_{ink}(x; \rho_{in})$ in terms of ρ_{in} . Consequently, we begin by calculating $p(v^m)$ (to improve readability, the subscript ink will be omitted throughout the remainder of the paper), which is defined in (1) and appears on the left side of (13), given the ρ_{in} values in (A3) that result from the facility's historical testing policy. During the course of this calculation, we will see that our inability to compute the right side of (13) becomes a moot point. As a first step toward this calculation, we assume that the true value v^t and the measurement error ϵ are independent and normally distributed with respective means μ_ξ and μ_e , and respective standard deviations σ_ξ and σ_e ; the estimated values for these four parameters for a sample of component measurements appear in Table 1. These nor-

Table 1 Sample of Component Measurements

Component	Nominal Value μ_ξ	Mean Measurement Error $\hat{\mu}_e$ (%)	Measurement Error Std. Dev. $\hat{\sigma}_e$ (%)	Component Value Std. Dev. $\hat{\sigma}_\xi$ (%)
R110	31600 Ω	0.0407	0.0162	0.1472
R111	1000 Ω	0.0530	0.0421	0.2449
R132	1780 Ω	1.9721	0.0700	0.2187
R158	10 Ω	1.6580	0.7778	0.3208
R162	1000 Ω	0.1059	0.0120	0.1192
R168	100 Ω	-0.0983	0.0370	0.2835
R170	17800 Ω	-2.2224	0.2614	0.4085
R210	1000 Ω	0.0582	0.0116	0.1718
R213	1000 Ω	0.0896	0.0125	0.2082
R316	14700 Ω	-0.1642	0.0313	0.3167
R317	464000 Ω	0.0122	0.1526	0.2318
R322	5110 Ω	-0.4706	0.0792	0.2541
C118	6.9 μF	-2.1074	0.1436	1.1643
C114	0.47 μF	-12.5308	1.4739	6.3439
C128	0.01 μF	-2.2309	0.9593	1.3864
C129	0.47 μF	-0.2691	0.2845	1.7317
C202	0.1 μF	1.0444	0.3189	4.6113
C205	0.1 μF	2.5666	0.3238	3.0688
C301	0.4 μF	-2.8866	1.2112	2.3799
C310	1 μF	-7.6399	0.1894	0.5920
C131	33 μF	-0.0072	0.2814	2.6723
L101	1 μH	37.5506	6.3128	1.1392
L103	1 μH	45.8132	3.6897	1.7435
Q105	59.753 μF	N/A	9.6449	4.3180
Q108	92.116 μF	N/A	0.4122	2.2315
Q110	28.299 μF	N/A	0.5867	5.4653
Q118	36.404 μF	N/A	0.3368	5.5723
Q122	48.807 μF	N/A	0.3554	1.9639
CR103	1.623 μF	N/A	0.0890	0.1148
CR106	0.726 μF	N/A	0.1210	0.1145
CR109	0.723 μF	N/A	0.1136	0.0695
CR111	2.169 μF	N/A	0.1133	0.2104
CR302	0.591 μF	N/A	0.1536	0.2206
CR305	0.597 μF	N/A	0.1566	0.1970

mality assumptions are difficult to validate because the two distributions are not at our disposal. We can assess the normality of the measurement distribution, however, which is the convolution of the two distributions in question. Notice that the results from the controlled experiment in the appendix are not appropriate for testing this assumption, since the experiment contains many repeated measurements. Hence, the production data set (see the last paragraph of the appendix for a

description of this data) was used, and for each of the 350 components, we applied the goodness-of-fit test associated with the kurtosis measurement in Equation (27.6a) of Duncan (1986) to the 80 measurement values. For more than 250 of the components, the normality assumption was not rejected at the 95% significance level. Under these normality assumptions, Equation (1) can be simplified to

$$p(x) = \frac{1}{2} \left(\operatorname{erf} \left(\frac{(G_u - \mu_\xi)\sigma_e^2 + (G_u + \mu_e - x)\sigma_\xi^2}{\sigma_e\sigma_\xi\sqrt{2(\sigma_e^2 + \sigma_\xi^2)}} \right) - \operatorname{erf} \left(\frac{(G_l - \mu_\xi)\sigma_e^2 + (G_l + \mu_e - x)\sigma_\xi^2}{\sigma_e\sigma_\xi\sqrt{2(\sigma_e^2 + \sigma_\xi^2)}} \right) \right), \quad (15)$$

where G_l and G_u are the lower and upper limit of the interval G such that the measured component is good if its true value lies inside it, and the error function $\operatorname{erf}(z) = 2/\sqrt{\pi} \int_0^z e^{-t^2} dt$.

We also assume that

$$\frac{|G_u - G_l|}{6\sigma_\xi} > 1, \quad (16)$$

and

$$\sigma_\xi > 3\sigma_e. \quad (17)$$

In quality management terminology, Inequality (16) states that the *machine capability index* is greater than 1; that is, the natural process range, $6\sigma_\xi$, is smaller than the product specification range, $|G_u - G_l|$. Since many companies are currently striving for "6 σ " capability (i.e., $|G_u - G_l| > 12\sigma_\xi$), (16) often holds in practice. Inequality (17) requires that the test measurements be reasonably reliable.

The experimental results that are partially contained in Table 1 were used to test assumptions (16)–(17). By substituting $\hat{\sigma}_\xi^2$ for σ_ξ^2 in (16), we found that the machine capability index was greater than 1 for all 79 components. To assess the validity of (17), readers are referred to Table 1; for all four inductors in our study, the estimated variance of the measurement noise, $\hat{\sigma}_e^2$, is higher than the estimated variance of the true values of these components, $\hat{\sigma}_\xi^2$. Hence, these measurements cannot distinguish between good and defective components. The estimated variance of the measurement noise is the same order of magnitude as the estimated variance of

the true values for the 13 diodes, and none of the inductors or diodes satisfies (17). In contrast, for the other three component types, 50 of the 62 components, and hence 50 of 79 in total, satisfy (17). For these other three component types, most of the components have an estimated noise variance much smaller than the estimated variance of their true values. These components can be tested with good accuracy. This analysis has important practical implications: Some components should not be tested (or, alternatively, a new test should be devised for them).

By conditions (16)–(17), one term of the right side of Equation (15) will always be equal to 1 or -1 , implying that $p(L) = p(U)$ for any L and U such that

$$L = \mu_\xi + \mu_e + (G_l - \mu_\xi)z \quad \text{and} \quad (18)$$

$$U = \mu_\xi + \mu_e + (G_u - \mu_\xi)z,$$

for any $z > 0$. Thus, by varying z , an entire class of policies can be easily generated; policies in this class are *Pareto optimal* with respect to Type I and Type II errors, in that it is impossible to reduce one type of error without increasing the other. In particular, the optimal solution of Equation (13) will be in this class of policies.

Figure 2 shows the form of the function $p(x)$ under assumptions (16)–(17). This function is equal to 1 between $G_l + \mu_e$ and $G_u + \mu_e$, and drops off very steeply to 0 somewhere outside this interval. The function is only symmetric if the tolerance interval is centered around the nominal value. Notice that Figure 2 and the normality of $f(x)$ imply that the second-order conditions leading to (13) are satisfied. Since the right side of (13) will take on values strictly between 0 and 1, the optimal cutoff points will most probably be located in the steep portions of the curve in Figure 2. Hence, the estimates for μ_ξ , μ_e , σ_ξ , and σ_e , which are used to obtain the estimate $\hat{p}(x)$ for $p(x)$, need to be very accurate. For example, suppose we seek to find the two cutoff points such that $p(x) = 0.9$ and decide to use the cutoff points $\bar{x}$ such that $\hat{p}(\bar{x}) = 0.9$. Furthermore, suppose that $\sigma_e = 0.9\hat{\sigma}_e$ (i.e., the estimate is 11.1% too high) and all other estimates are perfectly accurate. Under these assumptions, we used data from various components in our study and found that $p(\bar{x})$ takes on values between 0.2 and 0.4. Unfortunately, despite gathering a large amount of data, the 95% confidence intervals for the variance es-

estimates $\hat{\sigma}_\xi^2$ and $\hat{\sigma}_e^2$ are such that the estimates could be off by about 50% in either direction. To get out of this quandary, we attempted to group components in subsets that had very similar properties. The key characteristic used for the aggregation was that the standard deviation of the measurement noise and the standard deviation of the component values represented nearly the same percentage of the nominal value of the component. Although our aggregation led to tighter confidence intervals for σ_e^2 and σ_ξ^2 , the intervals were still too large to meaningfully optimize the testing policy.

In summary, even if we had a means of calculating the right side of (13), our variance estimates are too imprecise to derive a reliable testing policy. Consequently, we propose to simply set the lower and upper cutoff limits to the middle points of the ascent and descent of $p(x)$ in Figure 2; that is, we find the two points such that $p(x) = 0.5$. These points, denoted by x_l and x_u , are derived by setting the argument of either error function in (15) equal to 0, which yields

$$x_l = \mu_\xi + \mu_e + (G_l - \mu_\xi) \left(\frac{\sigma_\xi^2 + \sigma_e^2}{\sigma_\xi^2} \right), \quad (19)$$

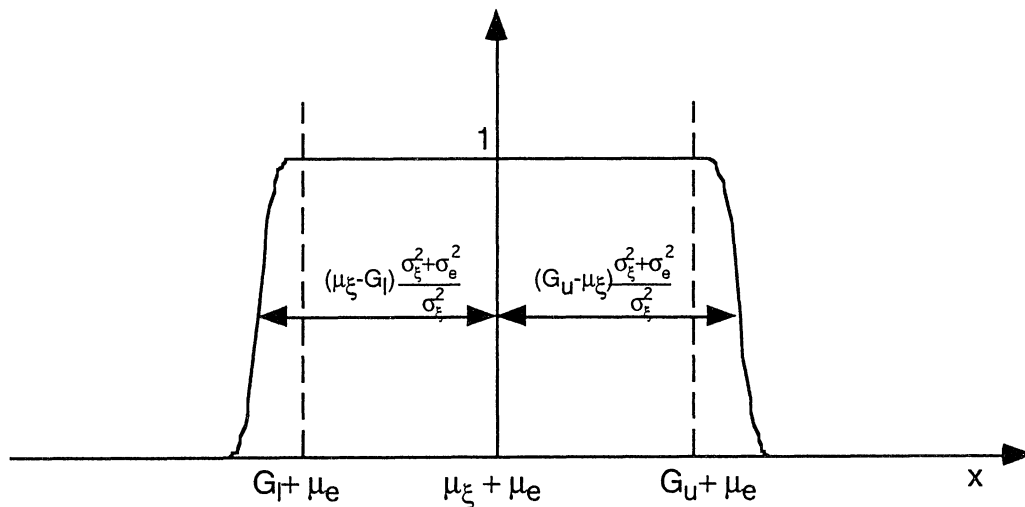
and

$$x_u = \mu_\xi + \mu_e + (G_u - \mu_\xi) \left(\frac{\sigma_\xi^2 + \sigma_e^2}{\sigma_\xi^2} \right). \quad (20)$$

Although the value in the right side of (13) will often be closer to 0.9 than 0.5 in practice, the corresponding difference in the cutoff points will typically be dwarfed by the inaccuracies in the measurement errors. Moreover, the cutoff points x_l and x_u have several desirable features. Even if the parameter estimates are inaccurate, the testing policy (x_l, x_u) is still Pareto optimal with respect to Type I and Type II errors. Also, this policy is easy to derive and Hewlett-Packard's engineers found the closed form expressions (19)–(20) to be intuitively appealing and easy to understand. More specifically, the proposed testing policy has the following intuitive meaning (consult Figure 2): The acceptance interval should be shifted from the tolerance interval by the expected value of the measurement noise; the tolerance on both sides should be multiplied by a factor that is the ratio of the sum of variance of the true values of the component and the measurement noise over the variance of the true values of the component. This ratio shows very clearly how the acceptance interval is affected by the measurement noise.

To apply the proposed testing policy (19)–(20), we focused on the in circuit test for the board type considered in Table 1. Although this test is used to identify all six defect types listed earlier, we only considered the test measurements for defective components, where a measurement is taken from each of $K_{in} = 79$

Figure 2 The Function $p(x)$



components. Hence, we calculated 79 acceptance intervals. The data partially displayed in Table 1 were used to derive the estimates in (A5) for each component. These quantities, along with μ_ξ and the interval $[G_l, G_u]$, allowed us to determine via (19)–(20) the proposed testing policy (x_l, x_u) for each component. For example, the interval $[G_l, G_u]$ for component R168 in Table 1 is [99, 101], and hence $x_l = 98.8847$ and $x_u = 100.9187$ for this component. Under normality assumptions, the expected number of Type I and Type II errors per board was calculated using (2)–(3). The resulting pair of values, $\alpha_{in}(T_{in})$ and $\beta_{in}(T_{in})$, correspond to the point *a* in Figure 3. The remainder of the Pareto optimal tradeoff curve in Figure 3 was derived by performing a similar aggregation using (18).

Hewlett-Packard engineers have observed that about 55% of the rejected components at the in circuit test are good components. Further analysis found that this percentage did not vary significantly by board type. To illustrate how this performance compares to that of our proposed testing policy, Figure 3 also shows a straight line that represents the set of policies such that 55% of the rejected components are good components. If we had obtained somewhat different estimates for σ_e^2 and σ_ξ^2 , then the proposed policy would have been on the

tradeoff curve in the neighborhood of *a*, and a significant improvement over the current policy would still be achieved; the improvement from the straight line to the point *a* in Figure 3 represents a cost reduction for this test of roughly 5%.

4.4. Inspection Allocation Policy

In this subsection, we find an inspection allocation policy given the facility's *historical* testing policy. For this purpose, three types of circuit boards were chosen that were representative of the variety of boards manufactured; these board types are different from the types analyzed in Table 1 or Chevalier's (1992) Appendix B. Since the analog or digital nature of a board is one of its distinguishing features, we chose one board type with mostly digital components, one with mostly analog components, and one with a mixture of components. These board types were produced in volumes that were typical for the facility, and their yield ranged from relatively low to relatively high. Also, these three board types did not share any test and thus can be considered independently.

Figure 4 displays the frequency of the different defect types detected at each stage for the three board types under consideration. Each defect type is represented by

Figure 3 The Current Testing Policy and the Pareto Optimal Tradeoff Curve

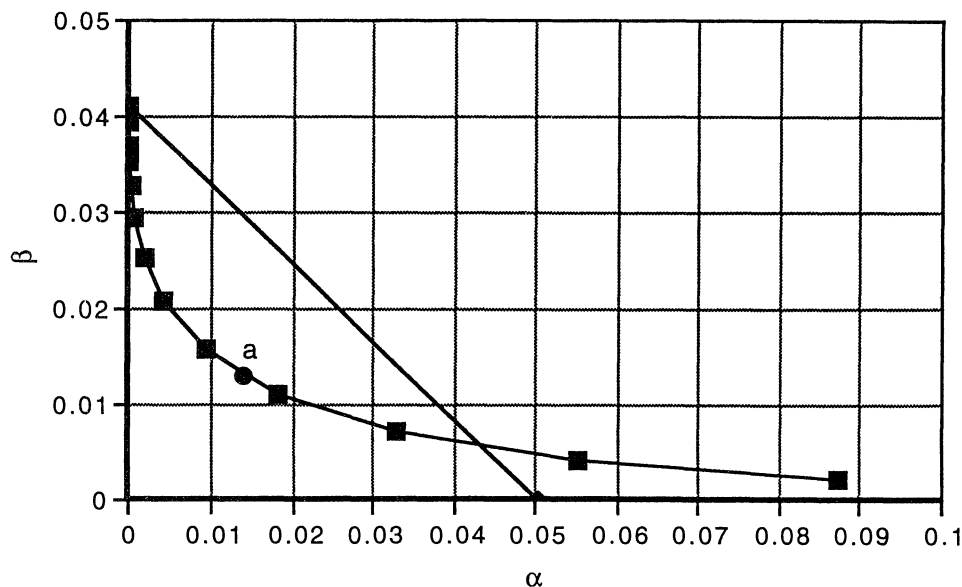
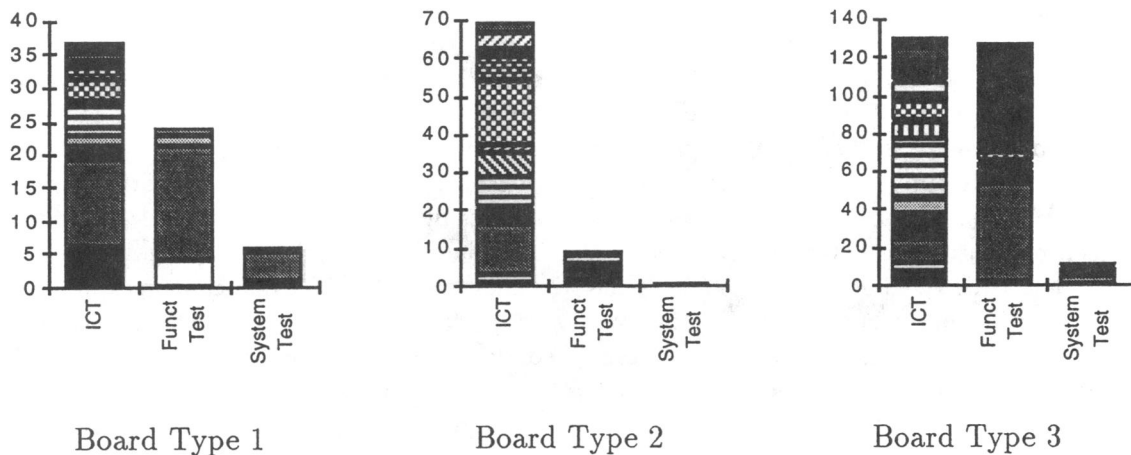


Figure 4 Frequency of Defects Detected at Each Stage for Each Board Type



the same pattern on all three charts. The values on each chart are arbitrary in order to disguise the data, but the relative values across the three charts are approximately correct; these values are proportional to the quantities ϕ_{in} defined in the appendix. The figure shows that the number of defects and the predominant types of defects vary significantly across the different board types. For example, type 3 boards have roughly three times as many defects as type 1 boards, and the medium gray defect type is predominant on type 1 boards but hardly present on type 2 boards.

From this data and some additional data about the reliability of the different tests, we derived the estimates α_{in} and β_{in} in (A2), which represent the expected number of errors per board under the facility's historical inspection allocation and testing policy. In the dynamic programming optimality equation (7), however, these quantities are a function of the testing policy T_n at stage n and ρ_{in} , which reflects the inspection allocation and testing policy at previous stages. Since we are trying to

derive an inspection allocation policy given the facility's historical testing policy, the historical error rates α_{in} and β_{in} in (A2) need to be somehow extrapolated to incorporate the dependence on the upstream inspection allocation policy. To account for this dependence, we used the following heuristic procedure that was guided by the empirical experience of the engineers at Hewlett-Packard. We assume that the Type I error $\alpha_{in}(T_n, \rho_{in})$ is *unaffected* by the upstream allocation policy, and hence substitute $\alpha_{in}(T_n, \rho_{in})$ in (7) by α_{in} in (A2). We assume that the Type II error $\beta_{in}(T_n, \rho_{in})$ is *proportional* to ρ_{in} , which is the expected number of detectable type i defects per board at stage n ; that is, we replace $\beta_{in}(T_n, \rho_{in})$ in (7) by $\kappa_{in}\rho_{in}$, where the constant κ_{in} is computed as follows. Let $\hat{\rho}_{in}$ denote the left side of Equation (5) when, on the right side of (5), we set $z_{n-1} = 1$, use the value $\beta_{i,n-1}$ from (A2) for $\beta_{i,n-1}(T_n, \rho_{i,n-1})$ and use the value δ_{in} from (A1); thus, $\hat{\rho}_{in}$ is the expected number of detectable defects per board under the facility's historical inspection allocation and testing policy. Then the proportionality constant κ_{in} is simply given by $\beta_{in}/\hat{\rho}_{in}$. It is worth noting an interesting observation from our unreported numerical studies that supports our heuristic substitutions: If the true measurement value is normally distributed and its standard deviation is changed, then the Type I error rate remains nearly constant and the Type II error rate is nearly proportional to the number of detectable defects on the board. With these estimates, we calculate the proposed inspection allocation policy

Table 2 Proposed Inspection Allocation Policies

Board Type	Current Inspection Allocation	Proposed Inspection Allocation	Cost Reduction
1	1-2-3	1-3	20%
2	1-2-3	2	23%
3	1-2-3	2-3	6.5%

under the historical testing policy by solving (7)–(8) for all 2^3 possible allocation policies. An iterative numerical solution to (7)–(8) with these substitutions is straightforward; details can be found in Chevalier (1992).

The proposed inspection allocation for each board type is displayed in Table 2. As expected, the proposed allocation varies for the different board types. Referring to Figure 4 and Table 2, we see that the system test is only bypassed by board type 2, which has the lowest frequency of defects detected at system test. Since the total inspection cost represents about half of the total manufacturing cost, the savings realized by the proposed policy in the different cases are significant.

5. Concluding Remarks

Motivated by an inspection problem at a specific industrial facility, we developed and analyzed a mathematical model to address this problem. This paper contains a formulation and analysis of the problem and a description of an attempt to apply the model at a Hewlett-Packard facility. The representation of the true measurement values is undoubtedly the most difficult modeling issue we faced. Our model assumes that the probability distribution of the true measurement value is a function solely of the expected number of detectable defects on the board; the latter quantity is a surrogate for the testing policies employed at the stages that are upstream from where the measurement was taken. This assumption is the key limitation in the model. On the one hand, it represents a simplification of a very complex issue; on the other hand, this conditional probability density function was extremely difficult to estimate from the available data at Hewlett-Packard. Our inability to estimate this function was the first of two major obstacles we encountered during the model application phase. Consequently, we decided to forego the application of the jointly optimal policy, and to apply our results in a fragmentary fashion. Our second obstacle was not foreseen: Even after performing controlled experiments and aggregating similar component types, we discovered that our parameter estimates were not precise enough to reliably derive an optimal testing policy. We used this observation to our advantage, however, by deriving a suboptimal testing policy that was intuitively appealing to Hewlett-Packard's test engi-

neers. By assuming a simple relationship between the error rates and the number of detectable defects on a board, we used the historical data to construct an inspection allocation policy given their historical testing policy.

Although we were unable to apply the jointly optimal policy, the crude application appears to have been quite successful. For three representative board types in the case study, the optimal inspection allocation policy achieves a 10% to 20% reduction in expected inspection costs relative to the facility's historical inspection policy, under the facility's historical testing policy. For the in circuit test that detects defective components, the proposed testing policy outperforms the facility's historical testing policy on one board type, representing roughly a 5% cost reduction for this test. Since the cost of inspection represents about half of the total direct manufacturing cost, these cost savings are significant. Moreover, both policies are relatively easy to implement in practice.

Our model was used in a limited way for several years at the Hewlett-Packard facility. In particular, the proposed testing policy (19)–(20) was employed for some of their key tests. Ironically, if we had been able to derive the optimal testing policy, then they probably would not have used it (because of the complexity of its computation) and we would not have been motivated to pursue the suboptimal policy that was eventually implemented. For new board types that were being introduced, they also used the dynamic programming equations (7)–(8) to evaluate the total expected cost for various inspection allocation policies, using estimates of the Type I and Type II error rates. As a result, they omitted in circuit tests for a number of their boards; it is worth noting that, before our analysis, facility managers were leaning toward heavier use of in circuit tests and omitting functional tests. Although Hewlett-Packard engineers felt that new testing policies and the reduced amount of in circuit testing led to significant cost savings, they were unable to isolate and quantify the effects of our model on their cost savings or field defect rates. This facility used our model for about two years; by that time, technological changes in the products and boards made our model obsolete, and they did not have anyone to perform the necessary in-house tinkering of the model. Since a thorough investigation of

measurement errors had never been undertaken at this facility, perhaps the biggest contribution of our study was the statistical analysis of the data. We encountered many cases, previously unbeknownst to the facility's engineers, where the mean measurement error was an order of magnitude larger than the variance of the measurement error; see Table 1 for some examples. These tables are extremely useful for developing some quick insights into the optimal testing policy. For example, resistors such as R158, R170, and R317 should perhaps not be tested (or at least have extremely slack cutoff limits), since the measurement noise is so large relative to the true component noise. In contrast, fairly tight cutoff limits can be set for most of the resistors, where σ_e/σ_ξ is very small.

Two other uses of our model are described in Chevalier (1992). First, the marginal cost savings from reducing various types of defects are calculated; these quantities can be used to focus quality improvement efforts in an economic fashion. Hewlett-Packard manufactures their own test equipment, and the second use of the model is to evaluate the cost savings achieved by their new generation of testing equipment, which exhibits less measurement noise.

Several research issues remain unresolved. First, although our original model and analysis needed to be simplified in several essential ways before being applied at Hewlett-Packard, we believe that our results could be used in a more comprehensive way if a facility was in the fortunate state of having sufficiently rich data. To this end, an attempt could be made to gather the data necessary for implementing the jointly optimal policy. An outline for such an undertaking, which focuses on characterizing the true measurement v_{ink}^t by a family of probability density functions indexed by the expected number of defects present ρ_{in} , can be found on page 12 of Chevalier and Wein (1994). Alternatively, perhaps a different model could be developed that has less stringent data requirements, but still captures the impact of the upstream inspection policy on the downstream board quality. Our model considers one board type in isolation, and a second research issue is to analyze the case where a test measures the output of several boards as they are functioning together. In this case, the inspection allocation must coincide for the different board types that are jointly tested. Some preliminary

thoughts on how to address the multitype problem can be found in §3.3 of Chevalier (1992). Finally, we have taken the inspection configuration as given; additional improvements can no doubt be achieved by designing the board for improved testability.¹

¹ We are grateful to the people at the Hewlett-Packard plant in Andover, Massachusetts, for their advice and their tremendous effort in helping us obtain the necessary data for our model. We also thank Arnie Barnett and Steve Pollock for helpful discussions, and the area editor Hau Lee, the associate editor, and the referee for many helpful comments that greatly improved the paper. This research is supported by a grant from the Leaders for Manufacturing program at MIT and National Science Foundation Grant Award No. DDM-9057297.

Appendix: Parameter Estimation at the Hewlett-Packard Facility

Our model requires three different types of data. The first type is the cost parameters; these parameters were estimated from historical data and interviews with Hewlett-Packard engineers, and readers are referred to an earlier draft of our paper (Chevalier and Wein 1994) for the relevant details. The second type of data concerns the occurrence of defects and the last and perhaps most difficult type of data to gather pertains to the testing process.

Defect Data

The goal here is to estimate δ_{in} , which is the expected number of new type i defects per board appearing at stage n , and the Type I and Type II errors per board, which are α_{in} and β_{in} , respectively. A brute force approach to estimating these quantities would require probability density estimates for many different tests, and would be extremely cumbersome. Instead, we develop a simpler approach that employs historical data and engineers' estimates. More specifically, we used historical data to estimate two quantities that are required to estimate δ_{in} : the mean number of defects of type i per board detected at inspection stage n , which we denote by ϕ_{in} , and η_i , which is the mean number of type i failures per board that occurred during the warranty period of the equipment. A failure during the warranty period was considered to be caused by an undetected defect. In addition, the engineers in charge of each test provided estimates of the proportion a_{in} of incoming defects (i.e., defects that have not been detected thus far) of each type that they felt their test should detect, and the proportion b_{in} of defects of each type detected by the test that were actually good boards (false defects).

From these data, we calculated the number of defects per board for defect types $i = 1, \dots, I$ that become detectable at each stage $n = 1, \dots, N$ ($I = 6$ and $N = 3$ at the Hewlett-Packard facility), which is (recall that inspection was undertaken at every stage)

$$\delta_{in} = a_{in} \left(\sum_{m=1}^3 \phi_{im}(1 - b_{im}) + \eta_i - \sum_{m=1}^{n-1} \delta_{im} \right) \quad \text{for } n = 1, \dots, 3,$$

$$\delta_{i4} = \sum_{m=1}^3 \phi_{im}(1 - b_{im}) + \eta_i - \sum_{m=1}^3 \delta_{im}. \quad (\text{A1})$$

Notice that δ_{i4} , which is the expected number of type i defects per board that are undetectable, will be zero when a_{i3} is one. For defect type $i = 1, \dots, 6$ and stage $n = 1, \dots, 3$, the historical expected number of Type I and Type II errors per board are, respectively,

$$\alpha_{in} = \phi_{in}b_{in} \quad \text{and} \quad \beta_{in} = \sum_{m=1}^n (\delta_{im} - \phi_{im}(1 - b_{im})). \quad (A2)$$

These historical estimates will be used later. Notice that the procedure culminating in (A2) is much simpler and probably more reliable than attempting to estimate these historical quantities via (2)–(3). Finally, by (5), note that the historical ρ_{in} values are given by

$$\rho_{in} = \beta_{i,n-1} + \delta_{in}. \quad (A3)$$

Testing Data

To derive the testing policy, we need the intervals G_{ink} and the density functions $e_{ink}(x)$ and $\xi_{ink}(x)$ for each quantity measured at each stage. For a system to function properly, the true value of each quantity measured should be in the corresponding interval G_{ink} . We used the specifications to which the component was bought to determine this interval.

To estimate the distribution of the measurement noise $e_{ink}(x)$ and the distribution of the true value of the quantity measured $\xi_{ink}(x)$, we only have the empirical distribution of the measured values at our disposal. This measured value is the sum of the true value of the component and the measurement noise; unfortunately, neither of these two quantities can be estimated independently. The true values are almost impossible to measure, since most components used at this facility are surface mount components, which are extremely small and fragile. Estimating the measurement noise is also a delicate task since the distribution of this noise will depend on many things, such as the type of component that is being measured, how the measurement is *guarded* (guarding is the technique used to try to isolate a component from the rest of the circuit board) and the topology of the board. As a result, the measurement noise can only be determined via experiments for each different board type.

At our request, a controlled experiment was performed at the Hewlett-Packard facility to study the measurement noise at the in circuit test level. At this facility, several in circuit testers, also called *testheads*, are used in parallel, and boards are typically tested on the first available testhead. Consequently, we wanted to find out what portion of the measurement noise is attributable to variations across different testheads. The experiment repeated each of 79 measurements $K = 10$ times consecutively on $H = 3$ different testheads for $B = 10$ different boards of the same type, and generated nearly 24,000 data points. One measurement was taken from each of the 79 components on each board, which consisted of 28 resistors, 22 capacitors, 13 diodes, 12 transistors, and 4 inductors. For each of the 79 measurements, the noise was modeled by expressing the measurements y_{bhk} as

$$y_{bhk} = \mu + \tau_b + \theta_h + \psi_{bh} + \epsilon_{bhk} \quad \text{for } b = 1, \dots, B, \quad h = 1, \dots, H, \quad k = 1, \dots, K, \quad (A4)$$

where:

μ is the reference value for the component being measured;

τ_b is the average deviation of the measurements taken on the b th board, and has zero mean and standard deviation σ_τ ;

θ_h is the average deviation of measurements taken on the h th head, and has zero mean and standard deviation σ_θ ;

ψ_{bh} is the average deviation of measurements taken on the h th head and the b th board, and has zero mean and standard deviation σ_ψ ;

ϵ_{bhk} is the residual variation of a measurement that cannot be explained by the testhead, the board, or the interaction between the testhead and the board, and has zero mean and standard deviation σ .

Standard statistical techniques give us estimators for all these quantities; we refer the interested reader to Duncan (1986) for their derivation.

This model incorporates the implicit assumption that the residual noise ϵ has the same variance on all testheads. The data indicated that this assumption did not hold for all components, which suggests that the three testheads were not equally calibrated when the data were gathered. Without this assumption, the estimation of the model parameters would have been greatly complicated. Moreover, by gathering data from only three testheads, a good estimation of the distribution of the variance of the residual noise across different testheads was not possible. Also, one might think that a fixed effects model would be more appropriate for the testheads, since the facility uses only a fixed number of different testheads; however, the variation between testheads evolves over time as a result of usage, maintenance, calibration, etc., and thus the random effects model seems more appropriate.

The estimated parameters from the statistical model in Equation (A4) are used to estimate the parameters of the distributions of the measurement noise and the component values. We assume that μ_ϵ , which will also be referred to as the *nominal* value of the component, is known (e.g., a 100-ohm resistor has a nominal value of 100 ohms), and let

$$\hat{\mu}_\epsilon = \bar{y} - \mu_\epsilon, \quad \hat{\sigma}_\epsilon^2 = \hat{\sigma}^2 + \hat{\sigma}_\tau^2 + \hat{\sigma}_\theta^2, \quad \text{and} \quad \hat{\sigma}_\epsilon^2 = \hat{\sigma}^2. \quad (A5)$$

For a representative sample of each of the five component types, Table 1 contains estimates that were derived from this experiment. Readers are referred to the Appendix of Chevalier and Wein (1994) for tables containing estimates for all 79 components considered in our study.

An important question for the design of future experiments is whether the variance of the noise associated with the different testheads can be predicted. To address this question, a regression was run to predict σ_ϵ from σ and the absolute value of μ_ϵ . The coefficients of correlation were consistently high; for example, $r^2 = 0.92$ for the resistors and $r^2 = 0.98$ for the capacitors.

Finally, the facility's emphasis on product output makes it impractical to run a similar controlled experiment for each of the 50 board types. Consequently, we also gathered data, which consisted of a measured value from each of 350 components on 80 different boards, from another board type during an actual production run, and used these data to assess the qualitative similarity between the two board types. We were encouraged to find that these results, which can be found in Appendix B of Chevalier (1992), and the corresponding results in Table 1 are strikingly similar.

References

- Chevalier, P. B., "Two Topics in Multistage Manufacturing Systems," Doctoral Dissertation, Operations Research Center, MIT, Cambridge, MA, 1992.
- Chevalier, P. B. and L. M. Wein, "Inspection for Circuit Board Assembly," Working Paper, Operations Research Center, MIT, Cambridge, MA, 1994.
- Duncan, A. J., *Quality Control and Industrial Statistics*, Irwin, Homewood, IL, 1986.
- Eppen, G. D. and E. G. Hurst, "Optimal Allocation of Inspection Stations in a Multistage Production Process," *Management Sci.*, 20 (1974), 1194–1200.
- Garcia-Diaz, A., J. W. Foster, and M. Bonyuet, "Dynamic Programming Analysis of Special Multistage Inspection Systems," *IIE Trans.*, 16 (1984), 115–125.
- Lindsay, G. F. and A. B. Bishop, "Allocation of Screening Inspection Effort—A Dynamic-Programming Approach," *Management Sci.*, 10 (1967), 342–352.
- Raz, T., "A Survey of Models for Allocating Inspection Effort in Multistage Production Systems," *J. Quality Technology*, 18 (1986), 239–247.
- Tang, C. S., "Designing an Optimal Production System with Inspection," *European J. Oper. Res.*, 52 (1991), 45–54.
- Villalobos, J. R., J. W. Foster, and R. L. Disney, "Flexible Inspection Systems for Serial Multi-Stage Production Systems," Working Paper, Department of Industrial Engineering, Texas A&M University, College Station, TX, 1992.
- White, L. S., "The Analysis of a Simple Class of Multistage Inspection Plans," *Management Sci.*, 9 (1966), 685–693.
- Yum, B. J. and E. D. McDowell, "The Optimal Allocation of Inspection Effort in a Class of Nonserial Production Systems," *IIE Trans.*, 13 (1981), 285–293.

Accepted by Hau L. Lee; received September 2, 1992. This paper has been with the authors 12 months for 2 revisions.