DEVELOPMENT OF A DUAL BURN-IN POLICY FOR SEMICONDUCTOR PRODUCTS BASED ON THE NUMBER OF DEFECTIVE NEIGHBORHOOD CHIPS

SEUNG HOON TONG

Memory Division, Semiconductor Business, SAMSUNG Electronics Co., Ltd., San#16, Banwol-dong, Hwasung-si, Gyeonggi-do, 445-701, Korea
shtong@samsung.com

BONG-JIN YUM

Department of Industrial Engineering, Korea Advanced Institute of Science and Technology, 373-1 Guseong-dong, Yuseong-gu, Daejeon, 305-701, Korea
bjyum@kaist.ac.kr

Most of the previous studies on developing an optimal burn-in policy for semiconductor products only deal with the burn-in process itself and little is concerned with utilizing the information on the quality levels of chips before being subjected to burn-in. Developed in this paper is a dual burn-in policy in which the number of chips \(d\) which do not pass the wafer probe (WP) test and lie in the neighborhood of a reference chip is utilized as an indicator on the quality level of that reference chip. The dual burn-in policy first classifies the chips which pass the WP test into two groups using a boundary value of \(d\), and then each group is subject to burn-in for its own duration. For a certain type of 256M DRAM product, the performance of the proposed dual burn-in policy is compared to that of the single burn-in policy in which all chips are subjected to the burn-in of the same duration without considering \(d\). The analysis results show that, for the cases considered, the proposed dual burn-in policy is more cost-effective than the single burn-in policy, implying that the additional information from the WP test is beneficial to establishing an efficient burn-in policy in semiconductor manufacturing.

Keywords: Burn-in; neighborhood chip; beta-binomial distribution; proportional hazard model.

1. Introduction

A semiconductor manufacturing process consists of two major stages: fabricating chips on a wafer and packaging individual chips cut from each wafer. At the end of the fabrication process, wafer probe (WP) is conducted to sort chips into electrically good and defective ones. A chip becomes defective if it contains the so called “killer defects”. Defective chips are usually discarded before putting into the packaging process. One exception is that, in the case of the defect tolerant memory product, it is first decided whether a chip with killer defects is repairable or not, and, if repairable, then it is laser repaired using redundant cells embedded in the chip and transferred to the packaging process. After the packaging process, every (in most cases) chip is subjected to a burn-in test in which an elevated temperature and voltage condition is used to weed out those chips that may cause infant mortality failures due to “latent defects”. The killer defect may include such defects as open, bridge, etc. which cause electrical function

---

1 Corresponding author. tel.: +82-42-869-3116; fax: +82-42-869-3110.
failures during the WP test. Although the latent defect is basically the same as the killer defect in nature, it passes the WP test without being detected since it is small in size and/or occurs in undetectable locations. A chip with latent defects may deteriorate with time and fail under stress (e.g., burn-in) or at the normal use condition. If it is not screened out during burn-in due to an insufficient burn-in duration or ineffective stress condition, then it may fail in the field resulting in a low early-period reliability.

While burn-in is beneficial to improving the reliability of products shipped to customers, it is costly, affects the productivity and could prevent timely shipment of product. Therefore, it is imperative for a semiconductor manufacturer to devise an efficient burn-in policy in terms of cost and product reliability. In this paper, we focus on the burn-in duration as a major factor for developing such an efficient burn-in policy.

If the burn-in duration is extended, then the direct burn-in cost and burn-in failure cost increase, while the field failure cost during the warranty period decreases. There exist numerous works\textsuperscript{1-6} which are concerned with determining the burn-in duration using these trade-offs. However, little is concerned with utilizing the information obtained at the WP test on the quality levels of chips before being subjected to burn-in.

In this paper, the number of chips ($d$) which do not pass the WP test and lie in the neighborhood of a reference chip which passes the WP test is considered as an indicator on the quality level of that reference chip. This is based on the well-known observation that defects on a wafer are clustered rather than randomly distributed,\textsuperscript{7,8} and therefore, a reference chip with a larger $d$ is more susceptible to failure during and after burn-in than the one with a smaller $d$, and vice versa.\textsuperscript{9} Developed in this paper is a dual burn-in policy in which the chips which pass the WP test are first classified into two groups using a boundary value of $d$, and then each group is subjected to burn-in for its own duration. To optimize the parameters of the proposed dual burn-in policy (i.e., boundary value of $d$ and burn-in duration for each group), a cost model is developed considering the direct burn-in cost, burn-in failure cost, field failure cost, and handling cost per chip. In order to evaluate the expected total cost per chip, the distribution of $d$ is modeled as a beta-binomial distribution, and the product reliability during and after burn-in is described using a proportional hazard model\textsuperscript{10} with $d$ as an explanatory variable. Then, in terms of the expected total cost, the proposed dual burn-in policy is compared to the single burn-in policy in which $d$ is not considered and all packaged chips are subjected to the burn-in of the same duration.

The rest of this paper is organized as follows. In Sec. 2, the related works are reviewed and their relationship with the present study is examined. Sec. 3 introduces a method for empirically estimating the distribution of $d$. In Sec. 4, a proportional hazard model is developed for predicting the product reliability, and an estimation method of its unknown parameters is described. In Sec. 5, the single and dual burn-in cost models are developed using the findings in Secs. 3 and 4. In addition, the two policies are optimized and compared for a certain type of defect tolerant 256M DRAM products in terms of the expected total cost per chip. Finally, Sec. 6 concludes the paper and discusses future research areas.
2. Review of Related Works

“Burn-in yield” is defined as the percentage of chips which survive burn-in and is also called “reliability yield” since it is closely related to the latent defect that may affect the early lifetime of the product. Similarly, “WP yield” denotes the percentage of chips that pass the WP test including those that are repaired when applicable. To develop an efficient burn-in policy, it is important not only to analyze the burn-in results but also to identify and utilize the relationship between the burn-in and WP test results. However, the existing works on yield modeling are mostly concerned with predicting WP yield, and correlating WP yield with burn-in yield (i.e., yield-reliability modeling) has not been sufficiently investigated. Early studies on yield-reliability modeling utilized only the basic information, WP yield, among the information that can be explicitly obtained from the WP test. Notable exceptions include Refs. 9, 15 and 16 which defined neighborhood chips in various ways for a reference chip that passes the WP test, and utilized the WP test results on these neighborhood chips in inferring the burn-in yield of each reference chip. This is based on the well-known observation that defects generated in the fabrication process are not randomly distributed but clustered on a wafer. That is, d for each reference chip provides additional information on the degree of defect contamination and can be used as an indicator for discriminating the reliability level of each reference chip. Barnett et al. also suggested an analytical model to predict burn-in yield using d. However, these existing yield-reliability models do not take the time variable into consideration, and therefore, cannot explain the fact that burn-in yield changes as the burn-in duration changes even in the case where the WP test results (e.g., WP yield, d, etc.) are the same.

Recently, several authors developed yield-reliability models taking the time variable into consideration. Kim et al. and Kim et al. introduced the time variable into the yield-reliability model for describing the breakdown of gate oxide, a specific structure within a semiconductor device. However, these studies did not consider such an explanatory variable as d and dealt with only a specific type of failure (i.e., gate oxide breakdown). On the other hand, Barnett and Singh developed a general yield-reliability model that includes d as a parameter and considered the time variable without limiting the model to a specific type of failure. An advantage of their theoretical model is that it consists of parameters (e.g., average number of killer defects, degree of defect clustering, ratio of average number of latent defects to that of killer defects, etc.) which are meaningful for process evaluation and control. However, in order to use their model to determine the burn-in duration, two assumptions need to be satisfied. That is, the time limit within which every latent defect in each chip is precipitated to failure must be specified in advance, and the average number of latent defects per chip must be small. In general, however, it is not easy to explicitly specify such a time limit in practice. Besides, for the matured product, it can be expected that the average number of latent defects per chip is small, but for the product in its early stage of production, such an expectation is hard to be realized. Finally, as shown in Sec. 4, the actual burn-in yield of the product considered in this paper shows a piecewise linear pattern rather than a smooth one as
predicted by the theoretical model in Barnett and Singh. Therefore, in the present study, an empirical yield-reliability model based on actual data is developed to overcome the above-mentioned problems with the existing studies.

3. **Distribution of the Number of Defective Chips in the Neighborhood**

3.1. **Definition of neighborhood**

The quality level (i.e., the degree of contamination due to latent defects) of a reference chip that passes the WP test can be indirectly inferred from the WP test results on the neighborhood chips of that reference chip. This is due to the well-known fact that defects tend to be clustered rather than randomly distributed on a wafer. For instance, when the number of neighborhood chips of a reference chip is eight as shown in Fig. 1, the quality level of a reference chip whose \( d \) equals zero (i.e., all eight neighborhood chips pass the WP test) would be different from that of a reference chip whose \( d \) equals eight, implying that there would also be a difference between the burn-in yields. In the present study, the neighborhood of a reference chip is assumed to be as shown in Fig. 1, and therefore, \( d \) takes the values from zero to eight.

For a reference chip around the edge of a wafer, the number of neighborhood chips may be less than eight as shown in Fig. 2. In such cases, it is usually assumed that all eight neighborhood chips exist and all of the imaginary (i.e., missing) chips are defective. That is, \( d \) for an edge reference chip is determined by counting the number of real defective neighborhood chips and the number of imaginary chips. This is for taking into account the phenomenon that an edge chip generally has a lower yield than an inner chip.

3.2. **Beta-binomial distribution**

Let \( A_k \) be the event that the \( k \)-th neighborhood chip of a reference chip does not pass the WP test. If \( A_k \)'s are independent and \( \{Pr(A_k), k=1,2,\cdots,8\} \) are the same, then the distribution of \( d \) can be described as a binomial distribution. However, due to defect clustering, \( A_k \)'s may be mutually dependent and \( \{Pr(A_k), k=1,2,\cdots,8\} \) may not be the same. In such a case, the binomial distribution is not appropriate and a beta-binomial distribution is more suitable for describing the distribution of \( d \).

Let \( p \) be the parameter of a binomial distribution, and assume that \( p \) is a random variable and follows a beta distribution with probability density function \( f(p) \). Then, a beta-binomial distribution results from compounding a binomial distribution with respect to \( f(p) \). If \( d \) is assumed to follow a beta-binomial distribution, then its probability mass function \( u(i) \) can be expressed as follows:

\[
u(i) = \Pr(\text{WP passed} | d = i) = \int_0^1 \left\{ \binom{N}{i} p^i (1-p)^{N-i} \right\} f(p) dp
\]

\[
u(i) = \int_0^1 \left\{ \binom{N}{i} p^i (1-p)^{N-i} \right\} \left\{ \frac{p^{a-1} (1-p)^{b-1}}{B(a,b)} \right\} dp
\]
where \( N \) is the number of neighborhood chips for a reference chip \((N = 8)\), \( p \) is the failure probability of a neighborhood chip \((0 < p < 1)\), \( B(\cdot, \cdot) \) is a complete beta function, and \( a \) and \( b \) are the shape and scale parameters of a beta distribution, respectively \((a, b > 0)\).

The shape of the distribution of \( d \) changes depending on the maturity of the product. That is, for the matured product, the mode of the distribution appears at a small value of \( d \) while, for the immature product, the mode tends to appear at a large value of \( d \). Fig. 3 shows histograms of \( d \) for two types of DRAM products. The 256M DRAM product in Fig. 3(a) is of high maturity and the mode appears at \( d = 0 \), while the product in Fig. 3(b) is less matured than the product in Fig. 3(a) and the mode appears at \( d = 1 \). The beta-binomial distribution can describe distributions with different modes through its shape parameter \( a \).

### 3.3. Fitting beta-binomial distribution to data

To fit a beta-binomial distribution to the actual \( d \) data for a certain type of 256M DRAM product, the following is defined first.

\[
u(i) = u(d = i | \text{WP passed}) = \frac{N_{wi}}{N_w}, \quad i = 0, 1, \ldots, 8
\]

where \( N_{wi} \) is the number of chips, each of which passes the WP test and has \( i \) defective chips in its neighborhood, and \( N_w \) is the total number of chips that pass the WP test. Then, the following nonlinear regression model can be formulated.

\[
u(i) = u(i; a, b) + \varepsilon_i, \quad i = 0, 1, \ldots, 8
\]

where \( \varepsilon_i \) is an error term. In Eq. (1), the variance of \( u(i) \) (or equivalently, the variance of \( \varepsilon_i \)) is not constant across \( i \), and therefore, \( a \) and \( b \) should be estimated using the weighted least squares method. The weight \( w_i \) for \( u(i) \) is defined as the reciprocal of the variance of \( u(i) \) and can be approximated as follows.

\[
w_i \approx \frac{N_w}{u(i)\{1-u(i)\}}, \quad i = 0, 1, \ldots, 8.
\]

Then, unknown parameters \( a \) and \( b \) in Eq. (1) can be estimated using the NLIN procedure in SAS with the “weight” option, and the results are shown in Tables 1 and 2. The adequacy of the beta-binomial distribution can be assessed from the high \( R^2 \) or \( R^2 \) (adjusted) value in Table 1 and the fitted-value plot in Fig. 4(a). In addition, the standardized residual plots in Fig. 4(b) do not show any unusual pattern.
On the other hand, the binomial model does not adequately describe the same data as shown in Figs. 5(a) and 5(b). Based on the above analysis, the probability mass function of \( d \) is estimated as \( \hat{u}(i; \hat{a}, \hat{b}) \) with \( \hat{a} = 1.0187 \) and \( \hat{b} = 5.9421 \).

4. Proportional Hazard Model for Reliability

To formulate a cost model for developing a burn-in policy, the reliability of the product must be described as a function of time during and after burn-in. The proportional hazard model proposed by Cox10 has been used to describe the relationship between the lifetime of a product and covariates or explanatory variables.25, 26 The present study employs the proportional hazard model with \( d \) as an explanatory variable to describe the reliability of chips that pass the WP test and are subjected to burn-in of a certain duration.

4.1. Proportional hazard model

The proportional hazard model with \( d \) as an explanatory variable can be expressed as follows.

\[
h(t; d) = h_0(t) g(d) = \left[ \frac{\beta(t)}{\eta} \right]^{-1} \exp(\theta_d d) \tag{2}\]

where \( t \) represents the lifetime of a chip and equals zero at the start of burn-in, \( h_0(t) \) is the baseline hazard function assumed to be a Weibull hazard function with shape and scale parameters \( \beta \) and \( \eta \), respectively, and \( \theta_d \) represents unknown coefficient for \( d \).

The baseline hazard function in Eq. (2) represents the hazard function when \( d = 0 \). To describe the failure rate and reliability as a function of time, the hazard function of the Weibull distribution is taken as the baseline hazard function in the present study as shown in Eq. (2). That is, the proportional hazard model adopted in the present study is parametric in nature. Eq. (3) is a re-expression of Eq. (2) in terms of the reliability function.27

\[
R(t; d) = \left\{ R_0(t) \right\}^{\exp(\theta_d d)}
= \left[ \exp\left\{ -(t/\eta)^{\beta} \right\} \right]^{\exp(\theta_d d)}, \quad d = 0, 1, 2, \ldots, 8 \tag{3}\]

where \( R(t; d) \) represents the probability that a chip, which passes the WP test and has \( d \) defective chips in its neighborhood, survives \( t \) hours of burn-in. Taking logarithm of both sides of Eq. (3) twice yields the following.

\[
\ln \left\{ -\ln R(t; d) \right\} = \ln \left\{ -\ln R_0(t) \right\} + \theta_d d
\Rightarrow \ln \left\{ -\ln R(t; d) \right\} = - \beta \ln \eta + \beta (\ln t) + \theta_d d
\Rightarrow y = \theta_0 + \theta_1 (\ln t) + \theta_d d \tag{4}\]
where \( y = \ln \{ -\ln R(t; d) \} \), \( \theta_0 = -\beta \ln \eta \), and \( \theta_1 = \beta \).

When actual data are given, \( R(t; d) \) can be estimated as follows.

\[
R'_{ij} = R'(t = t_j; d = i) = n_{ij} / N_{wij}
\]

where \( N_{wij} \) represents the number of chips, each of which passes the WP test, has \( i \) defective chips in its neighborhood, and is subjected to \( t_j \) hours of burn-in; and \( n_{wij} \) represents the number of chips that survive \( t_j \) hours of burn-in among \( N_{wij} \). Then, the following regression model is obtained from Eq. (4).

\[
y'_{ij} = \theta_0 + \theta_t \left( \ln t_j \right) + \theta_d i + \epsilon_{ij}, \quad i = 0, 1, \ldots, 8, \quad j = 1, 2, \ldots, m
\]

where \( y'_{ij} = \ln \{ -\ln R'_{ij} \} \), \( \epsilon_{ij} \) represents an error term and \( m \) is the number of different burn-in durations.

Eq. (5) represents a linear regression model, and if burn-in yield data exist for at least two different burn-in durations, \( t = t_1 \) and \( t_2 \), then all the parameters in Eq. (5) can be estimated.

In Fig. 6, \( y' \) is plotted against \( d \) for the 256M DRAM product subjected to two burn-in durations. The cases where \((t, d) = (3, 7), (3, 8), (6, 7) \) and \((6, 8)\) are omitted since the corresponding numbers of reference chips are very small (i.e., 16, 2, 7 and 1, respectively) compared to the total number of chips (= 19,823 chips) that pass the WP test. Fig. 6 shows an interesting pattern. That is, \( y' \) and \( d \) do not have a first-order relationship as in Eq. (5), but have a piecewise linear relationship with a change point at \( d = 2 \).

In other words, there is little change in \( y' \) as \( d \) increases from 0 to the change point, but after passing the change point, \( y' \) increases (except the case where \((t, d) = (3, 6)\)).

To incorporate the piecewise relationship between \( y' \) and \( d \) into the model, Eqs. (2) and (3) are respectively modified as

\[
h(t; d) = h_0(t) g(d) = \left\{ \frac{\beta \left( \frac{t}{\eta} \right)^{\beta - 1}}{\eta} \right\} \exp \left\{ \theta_d d + \theta_d z (d - 2) z \right\}
\]

\[
R(t; d) = \left[ \exp \left\{ -(t/\eta)^{\beta} \right\} \right] \exp \left\{ \theta_d d + \theta_d (d - 2) z \right\}
\]

where \( z = 0 \) if \( d < 2 \) and 1 otherwise. Subsequently, Eq. (5) is modified as follows.

\[
y'_{ij} = \theta_0 + \theta_t \left( \ln t_j \right) + \theta_d i + \theta_d z (i - 2) z_i + \epsilon_{ij}, \quad i = 0, 1, \ldots, 8, \quad j = 1, 2, \ldots, m
\]

where \( z_i = 0 \) if \( i < 2 \) and 1 otherwise.

The proportional hazard model in Eq. (6) or (7) has the property that, for given \( d \), the change in \( y = \ln \{ -\ln R(t; d) \} \) due to a change in \( t \) depends on the amount of change in \( \ln t \), but not on \( d \). That is,
\[ y(t_2;d) - y(t_1;d) = \ln \left\{ -\ln R(t_2;d) \right\} - \ln \left\{ -\ln R(t_1;d) \right\} = \theta_i \left\{ \ln t_2 - \ln t_1 \right\}. \]

Note that, in Fig. 6, the changes in \( y' \) due to the change in the burn-in duration are approximately constant at all values of \( d \) (except the case where \( d = 6 \)), implying that the above property of the proportional hazard model holds for the 256M DRAM data.

4.2. Estimation

The unknown parameters \( \theta_0, \theta_1, \theta_d \) and \( \theta_d \) in Eq. (8) can be estimated using the linear regression technique. Since the variance of \( y'_{ij} \) (or equivalently, the variance of \( e_{ij} \)) is not constant across \( i \) and \( j \), the weighted least squares method should be used with weight \( w_{ij} \) for \( y'_{ij} \) being given by the reciprocal of the variance of \( y'_{ij} \) as follows (see Appendix A for derivation).

\[ w_{ij} \approx \frac{N_{wij}R'_{ij} \left( \ln R'_{ij} \right)^2}{1 - R'_{ij}} \] (9)

Table 3 shows the regression analysis results for actual 256M DRAM data in Fig. 6 under the model in Eq. (8).

The point, \((t, d) = (3, 6)\), in Fig. 6 deviates from the behavior of the other points, and therefore, its residual is examined together with some diagnostic measures. First, the corresponding standardized residual value is -2.03, which is marginal to be declared as an outlier.28 The corresponding Cook’s distance, leverage, and DFIT values28, 29 are 0.0435, 0.0405, and -0.516, respectively, indicating that the point does not have an undue influence on the regression coefficients and/or predicted values. Therefore, the point is decided to be retained in the subsequent analysis.

Since the P-value of the i term in Table 3 is large, another regression analysis is performed without the i term and the results are shown in Tables 4 and 5.

It is noticed from Table 5 that the \( R^2 \) or \( R^2 \) (adjusted) value is high and the standardized residual plots in Fig. 7 do not show any unusual pattern. Tables 4 and 5 show that the regression model without the i term is still statistically sound, implying that \( y \) from \( d = 0 \) to 2 is flat rather than having a positive or negative slope.

Based on the above analysis, the regression equation for the 256M DRAM product can be expressed as

\[ y' = -2.9971 + 0.75137 (\ln t) + 0.28562 (i - 2) z. \] (10)

In terms of the reliability function,

\[ \hat{R}(t; d = i) = \hat{R}(t; i) \]
\[ = \exp \left\{ -(t/\hat{\lambda}) \hat{\lambda} \right\} \exp \left\{ \hat{\theta}_d (i-2)z \right\}. \]
In Eq. (11), \( \hat{\eta} \) is determined as \( \exp(-\hat{\theta}_0/\hat{\theta}_1) \).

5. Burn-in Cost Models

5.1. Single burn-in policy

In the single burn-in cost model, three types of costs are considered per chip. They include direct burn-in cost \( c_1 \) (per hour of burn-in), burn-in failure cost \( c_2 \), and field failure cost \( c_3 \). The burn-in failure cost is incurred if a chip fails during burn-in while the field failure cost is incurred if a chip shipped to a customer fails within the warranty period. In general, the three types of costs are related as \( c_1 < c_2 < c_3 \).

When the burn-in duration is \( t_b \) and the warranty period is \( t_f \) as shown in Fig. 8, the expected total cost per chip can be expressed as follows.

\[
C_{\text{single}}(t_b) = (\text{direct burn-in cost}) + (\text{burn-in failure cost}) + (\text{field failure cost})
= c_1 t_b + c_2 \Pr\{T < t_b, \text{WP passed}\}
+ c_3 \Pr\{T < t_b + t_f, T > t_b, \text{WP passed}\}
\]

(12)

Eq. (12) can be rewritten as Eq. (13) using \( u(i) \) and \( R(t; d = i) \) derived in Secs. 3 and 4, respectively (See Appendices B and C).

\[
C_{\text{single}}(t_b) = c_1 t_b + c_2 \sum_{i=0}^{8} u(i)\left[1 - R(t_b; i)\right] + c_3 \sum_{i=0}^{8} u(i)\left[R(t_b; i) - R(t_b + t_f; i)\right]
\]

(13)

In addition, there usually exists a requirement on the reliability level of the product in the field. Such a requirement is often expressed as the average failure rate (AFR) during the warranty period \( t_f \), which is defined as

\[
AFR(t_f) = \frac{1}{t_f} \int_0^{t_f} h(\tau) d\tau = \frac{-\ln R(t_f)}{t_f}
\]

(14)

where \( R(\cdot) \) is the field reliability function. Since only the chips that survive the burn-in test are shipped to the customer, \( R(t_f) \) in Eq. (14) can be expressed as follows (See Appendix D).

\[
R(t_f) = \Pr\{T > t_b + t_f, T > t_b, \text{WP passed}\}
= \sum_{i=0}^{8} u(i) R(t_b + t_f; i) / \sum_{i=0}^{8} u(i) R(t_b; i)
\]

(15)

The product reliability at the field time \( t_f \) can be obtained using Eq. (15). In this paper, the field time is translated into the time in the burn-in time scale using the acceleration factor (AF). An AF is
defined as \( AF = t_{q, \text{normal}} / t_{q, \text{accelerated}} \) for an arbitrary \( q \) (0 < \( q < 1 \)) where \( t_{q, \text{normal}} \) and \( t_{q, \text{accelerated}} \) are the \( q \)-th quantiles of the lifetime distributions at the normal (i.e., field) and accelerated (i.e., burn-in) conditions, respectively. In addition, it is assumed in the present study that the warranty period \( t_f \) is one year (i.e., approximately 9,000 hrs in the field time and 9,000/AF hrs in the burn-in time scale) and is within the infant mortality period.

Based on the above, the optimization problem for determining the burn-in duration \( t_b \) for the single burn-in policy can be formulated as follows.

Minimize \( C_{\text{single}}(t_b) \)
subject to \( AFR(t_f) \leq AFR_0 \) \hspace{1cm} (16)

where \( AFR_0 \) is the upper limit of \( AFR(t_f) \). In the above optimization, \( u(i) \) and \( R(t_b;i) \) are replaced with their estimates \( \hat{u}(i) \) and \( \hat{R}(t_b;i) \), respectively.

Since the burn-in duration is usually determined in an increment of one hour in practice, the optimal burn-in duration is also determined in one-hour unit using a grid search method over the search region [0, 96].

5.2. **Dual Burn-in Policy**

In the proposed dual burn-in policy, the chips which pass the WP test are first classified into two groups using the boundary value \( d_c \) of \( d \). Then, the chips whose \( d \) is less than or equal to \( d_c \) are subjected to burn-in of duration \( t_{bL} \), while the chips whose \( d \) is greater than \( d_c \) are subjected to burn-in of duration \( t_{bH} \) where \( t_{bL} < t_{bH} \). The justification of this policy is as follows. A chip with a larger value of \( d \) than others is more likely to fail early in the field, and therefore, the burn-in duration for such a chip is extended to increase the chance of screening, and thereby, to meet the reliability requirement in the field. On the other hand, when \( d \) is small, the burn-in duration is reduced to enhance productivity and to make economic benefits.

Under the proposed dual burn-in policy, the expected total cost per chip consists of the expected direct burn-in, burn-in failure, and field failure costs per chip for each group and the handling cost per chip (\( c_4 \)) for classifying the chips into two groups. That is,

\[
C_{\text{dual}}(t_{bL}, t_{bH}, d_c) = \left\{ \left( \text{direct burn-in cost} \right)_{\text{Low}} + \left( \text{burn-in failure cost} \right)_{\text{Low}} + \left( \text{field failure cost} \right)_{\text{Low}} \right\} \\
+ \left\{ \left( \text{direct burn-in cost} \right)_{\text{High}} + \left( \text{burn-in failure cost} \right)_{\text{High}} + \left( \text{field failure cost} \right)_{\text{High}} \right\} \\
+ \text{handling cost} \\
= c_1 t_{bL} \sum_{i=0}^{d_c} u(i) + c_2 \sum_{i=0}^{d_c} u(i) \{ 1 - R(t_{bL}; i) \} + c_3 \sum_{i=0}^{d_c} u(i) \{ R(t_{bL}; i) - R(t_{bL} + t_f; i) \}
\]


\[
+ c_1 t_{bH} \sum_{i=d_{c}+1}^{8} u(i) + c_2 \sum_{i=d_{c}+1}^{8} u(i) \left[ 1 - R(t_{bH}:i) \right] \\
+ c_3 \sum_{i=d_{c}+1}^{8} u(i) \left[ R(t_{bH}:i) - R(t_{bH} + t_f:i) \right] \\
+ c_4
\]

In general, \(c_4\) is smaller than \(c_1\) since sorting chips according to their \(d\) values at the WP test and handling the sorted chips in the packaging process and at the start of burn-in can be mostly automated.

It is shown in Appendix E that \(R(t_f)\) under the dual burn-in policy can be expressed as follows.

\[
R(t_f) = \frac{\sum_{i=0}^{d_c} u(i) R(t_{bL} + t_f:i) + \sum_{i=d_{c}+1}^{8} u(i) R(t_{bH} + t_f:i)}{\sum_{i=0}^{d_c} u(i) R(t_{bL}:i) + \sum_{i=d_{c}+1}^{8} u(i) R(t_{bH}:i)}. \tag{17}
\]

Then, \(AFR(t_f)\) can be calculated using Eq. (14).

The optimization problem for the dual burn-in policy can be formulated as for the single burn-in policy (see (16)). Again, \(u(i)\) and \(R(\cdot; i)\) are replaced with their estimates \(\hat{u}(i)\) and \(\hat{R}(\cdot; i)\), respectively, and a grid search method is used to determine optimal values of \(d_c\), \(t_{bL}\), and \(t_{bH}\). Note that \(d_c\) can take values from zero to eight. Optimal values of \(t_{bL}\) and \(t_{bH}\) are determined in one-hour unit over the search space [0, 96].

5.3. Comparisons of single and dual burn-in policies

For the 256M DRAM product considered in the present study, the relevant costs are estimated as: \(c_1 = 1\), \(c_2 = 20\), and \(c_3 = 40\). Note that \(c_1\) is standardized and the other costs are estimated relative to \(c_1\). It is assumed that \(AFR_0\) for one year (= 9,000 hrs) in the field is 1,000 FITs where one FIT corresponds to one failure in \(10^9\) hrs. In addition, the AF value for the 256M DRAM product under the current burn-in condition (specified by the applied temperature and voltage) is estimated as 20,000.

For a comparison of the single and dual burn-in policies, each policy is first optimized, and then the following \(\delta\) is calculated.

\[
\delta = (\text{optimal expected total cost per chip for the single burn-in policy}) \\
- (\text{optimal expected total cost per chip except } c_4 \text{ for the dual burn-in policy}).
\]

That is, if \(c_4\) is smaller than \(\delta\), then the proposed dual burn-in policy is preferred to the single burn-in policy, and vice versa.

Table 6 shows the computational results. The optimal dual burn-in policy is to classify the chips that pass the WP test into two groups using three as the boundary value \(d_c\), and to employ the burn-in durations of 15 and 29 hours for the first and second groups, respectively. For the single burn-in policy, the optimal burn-in duration turns out to be 17 hours. Note that \(\delta = 1.238\). Since \(c_4\) is usually smaller
than $c_1 (= 1)$, the dual burn-in policy is more cost-effective than the single burn-in policy for the product considered.

Since some uncertainties may exist in the estimates of $c_2$, $c_3$, and AF, a sensitivity analysis is conducted with respect to these parameters. It is first assumed that the true values of $c_2$, $c_3$, and AF lie within $\pm 30\%$ of the estimated values. Then, for each parameter, three true values are selected as

\[(70\% \text{ of the estimate, } 100\% \text{ of the estimate, } 130\% \text{ of the estimate}).\]

This results in 27 combinations of true parameter values, and for each combination, the true expected total cost per chip is calculated for each policy which was determined using the estimated parameter values. Table 7 shows the sensitivity analysis results. It is first noticed that $\delta$ values change from 1.153 to 1.321, implying that the advantage of the proposed dual burn-in policy is still maintained under the assumed uncertainties in the parameter estimates. On the other hand, the actual AFR values become greater than the required value ($= 1,000 \text{ FIT’s}$) for both policies when the true AF is smaller than the estimated one. This implies that an overestimation of AF should be avoided as much as possible.

6. Conclusion

Burn-in performs an important function to screen those products that are likely to fail early in the field, and thereby, to meet the reliability requirement. However, it is costly and may degrade productivity, and therefore, developing an efficient burn-in policy is imperative to manufacturers.

In this paper, a dual burn-in policy is developed and its effectiveness relative to the single burn-in policy is illustrated using the actual 256M DRAM data from a semiconductor manufacturing process. The proposed policy first classifies the chips that pass the WP test into two groups according to the number of defective neighborhood chips, and different burn-in durations are applied to the two groups. Then, a cost-based optimization problem with a reliability requirement is formulated and solved for the decision variables of the proposed policy. To the best of the authors’ knowledge, the present study is the first attempt to utilize the information on the WP test results in developing a burn-in policy.

The present study may be extended in several directions. In the proposed dual burn-in policy, the number of groups for the chips that pass the WP test is restricted to two. Policies with more than two groups need to be investigated to compare their effectiveness to the present one. In addition, for the defect tolerant memory product, the number of repairs for a chip at the WP test is another important indicator on the quality level of that chip. It may be a fruitful area of future research to develop a burn-in policy based on this indicator and compare its performance with the present one.

Acknowledgment
The authors thank the editor and a referee for their encouragement and suggestions that improve this article.

Appendices

A. Derivation of \( w_{ij} \) in Eq. (9)

\( R'_{ij} \) is defined as \( n_{wij}/N_{wij} \) where \( n_{wij} \) is distributed as Binomial \((N_{wij}, p_{ij})\). Therefore, \( E(R'_{ij}) = p_{ij} \) and \( Var(R'_{ij}) = p_{ij}(1-p_{ij})/N_{wij} \). Since \( y'_{ij} = \ln(-\ln R'_{ij}) \), a Taylor series expansion of \( y'_{ij} \) around \( p_{ij} \) yields

\[
y'_{ij} = \ln\left(-\ln R'_{ij}\right) = \ln(-\ln p_{ij}) + \frac{1}{p_{ij} \ln p_{ij}} (R'_{ij} - p_{ij})
\]

Therefore,

\[
E(y'_{ij}) \approx \ln(-\ln p_{ij}), \quad \text{and} \quad Var(y'_{ij}) \approx Var(R'_{ij})/(p_{ij} \ln p_{ij})^2 \approx (1-R'_{ij})/(R'_{ij}N_{wij}(-\ln R'_{ij})^2).
\]

Finally, \( w_{ij} \) is defined as the reciprocal of \( Var(y'_{ij}) \).

B. Derivation of \( \Pr\{T < t_b \mid \text{WP passed}\} \) in Eq. (12)

\[
Pr(T < t_b \mid \text{WP passed}) = \sum_{i=0}^{8} \Pr(T < t_b, d = i \mid \text{WP passed})
\]

\[
= \sum_{i=0}^{8} \frac{\Pr(T < t_b, d = i, \text{WP passed})}{Pr(\text{WP passed})} \cdot \frac{Pr(T < t_b, d = i, \text{WP passed})}{Pr(d = i, \text{WP passed})}
\]

\[
= \sum_{i=0}^{8} \Pr(d = i \mid \text{WP passed}) \Pr(T < t_b \mid d = i, \text{WP passed})
\]

\[
= \sum_{i=0}^{8} u(i) \left[ 1 - \Pr(T > t_b \mid S(i)) \right]
\]

\[
= \sum_{i=0}^{8} u(i) \left[ 1 - R(t_b; i) \right]
\]

where \( S(i) \) as an event that an arbitrary chip passes the WP test and has \( i \) defective chips in its neighborhood.

C. Derivation of \( \Pr\{T < t_b + t_f , T > t_b \mid \text{WP passed}\} \) in Eq. (12)

\[
Pr(T < t_b + t_f, T > t_b \mid \text{WP passed}) = \sum_{i=0}^{8} \Pr(T < t_b + t_f, T > t_b, d = i \mid \text{WP passed})
\]
\[
\begin{align*}
&= \sum_{i=0}^{8} \Pr\{T < t_b + t_f \mid T > t_b, d = i, \text{WP passed}\} \\
&= \sum_{i=0}^{8} \frac{\Pr(d = i, \text{WP passed}) \Pr(T > t_b, d = i, \text{WP passed})}{\Pr(d = i, \text{WP passed})} \\
&= \sum_{i=0}^{8} \frac{\Pr(T < t_b + t_f, T > t_b, d = i, \text{WP passed})}{\Pr(T > t_b, d = i, \text{WP passed})} \\
&= \sum_{i=0}^{8} \Pr(d = i, \text{WP passed}) \cdot \Pr(T > t_b \mid d = i, \text{WP passed}) \\
&= \sum_{i=0}^{8} u(i) \Pr(T > t_b \mid S(i)) \left[1 - \Pr\{T > t_b + t_f \mid T > t_b, S(i)\}\right] \\
&\quad \cdot \Pr\{T > t_b + t_f \mid T > t_b, d = i, \text{WP passed}\} \text{ (C.1)} \\
&= \sum_{i=0}^{8} u(i) \Pr(T > t_b \mid S(i)) \left[1 - \frac{\Pr\{T > t_b + t_f, T > t_b, S(i)\}}{\Pr(T > t_b, S(i))}\right] \text{ (C.2)} \\
&= \sum_{i=0}^{8} u(i) \Pr(T > t_b \mid S(i)) \left[1 - \frac{\Pr\{T > t_b + t_f, S(i)\}}{\Pr(T > t_b, S(i))}\right] \text{ (C.3)} \\
&= \sum_{i=0}^{8} u(i) \Pr(T > t_b \mid S(i)) \left[1 - \frac{\Pr\{T > t_b + t_f \mid S(i)\}}{\Pr(T > t_b, S(i))}\right] \text{ (C.4)} \\
&= \sum_{i=0}^{8} u(i) \left[R(t_b;i) - R(t_b + t_f;i)\right] \\
&= \sum_{i=0}^{8} u(i) [T > t_b + t_f | T > t_b, \text{WP passed}] \text{ in Eq. (15)}
\end{align*}
\]

\[
\text{Pr}(T > t_b + t_f | T > t_b, \text{WP passed}) = \sum_{i=0}^{8} \text{Pr}(T > t_b + t_f, d = i | T > t_b, \text{WP passed}) \\
= \sum_{i=0}^{8} \frac{\text{Pr}(T > t_b + t_f, d = i, T > t_b, \text{WP passed})}{\text{Pr}(T > t_b, \text{WP passed}) \cdot \text{Pr}(d = i, \text{WP passed})} \\
= \frac{1}{\text{Pr}(T > t_b, \text{WP passed})} \sum_{i=0}^{8} \text{Pr}(d = i, \text{WP passed}) \\
\]

\[
= \sum_{i=0}^{8} \frac{u(i) \Pr(T > t_b \mid S(i)) \Pr(T > t_b + t_f \mid T > t_b, S(i))}{\text{Pr}(T > t_b, \text{WP passed})} \\
= \frac{\sum_{i=0}^{8} u(i) \Pr(T > t_b \mid S(i)) \Pr(T > t_b + t_f \mid T > t_b, S(i))}{\text{Pr}(T > t_b, \text{WP passed})}
\]
\[
\sum_{i=0}^{8} u(i) \Pr \{ T > t_b + t_f \mid S(i) \} \frac{\Pr(T > t_b \mid WP\ passed)}{\sum_{i=0}^{8} u(i) \Pr \{ T > t_b + t_f \mid S(i) \}}
\]
(from Eqs.(C.1) – (C.4), Appendix C)

\[
\sum_{i=0}^{8} \frac{u(i) \Pr \{ T > t_b + t_f \mid S(i) \}}{\Pr(T > t_b \mid WP\ passed)}
\]

\[
\sum_{i=0}^{8} \frac{u(i) \Pr \{ T > t_b + t_f \mid S(i) \}}{\Pr(T > t_b \mid WP\ passed)}
\]

E. Derivation of field reliability under dual burn-in policy in Eq. (17)

\[
R(t_f) = \sum_{i=0}^{8} \Pr(T > t_{bi} + t_f, d = i \mid BI\ passed, WP\ passed)
\]

where \( t_{bi} = \) tbd for \( i \leq d_c \) and \( t_{bi} = t_{bi}^{HF} \) for \( i > d_c \).

\[
R(t_f) = \sum_{i=0}^{8} \frac{\Pr(T > t_{bi} + t_f, d = i, BI\ passed, WP\ passed)}{\Pr(BI\ passed, WP\ passed)}
\]

\[
= \sum_{i=0}^{8} \frac{\Pr(T > t_{bi} + t_f, d = i, BI\ passed \mid WP\ passed) \Pr(WP\ passed)}{\Pr(BI\ passed \mid WP\ passed) \Pr(WP\ passed)}
\]

\[
= \sum_{i=0}^{8} \frac{\Pr(T > t_{bi} + t_f, d = i \mid WP\ passed)}{\Pr(BI\ passed \mid WP\ passed)}
\]

(E.1)

Numerator of (E.1)

\[
= \sum_{i=0}^{8} \frac{\Pr(T > t_{bi} + t_f, d = i, WP\ passed)}{\Pr(WP\ passed)}
\]

\[
= \sum_{i=0}^{8} \frac{\Pr(d = i, WP\ passed) \Pr(T > t_{bi} + t_f, d = i, WP\ passed)}{\Pr(d = i, WP\ passed)}
\]
\begin{align*}
  &= \sum_{i=0}^{8} \Pr(d = i \mid \text{WP passed}) \Pr(T > t_{bi} + t_f \mid d = i, \text{WP passed}) \\
  &= \sum_{i=0}^{8} u(i) \Pr(T > t_{bi} + t_f \mid S(i)) \\
  &= \sum_{i=0}^{d_1} u(i) \Pr(T > t_{bl} + t_f \mid S(i)) + \sum_{i=d_1+1}^{8} u(i) \Pr(T > t_{bh} + t_f \mid S(i)) \\
  &= \sum_{i=0}^{d_1} u(i) R(t_{bl} + t_f ; i) + \sum_{i=d_1+1}^{8} u(i) R(t_{bh} + t_f ; i).
\end{align*}

Denominator of (E.1)
\begin{align*}
  &= \sum_{i=0}^{8} \Pr(\text{BI passed}, d = i \mid \text{WP passed}) \\
  &= \sum_{i=0}^{d_1} \Pr(T > t_{bl}, d = i \mid \text{WP passed}) + \sum_{i=d_1+1}^{8} \Pr(T > t_{bh}, d = i \mid \text{WP passed})
\end{align*}

using similar procedures as in Appendix B, we have
\begin{align*}
  &= \sum_{i=0}^{d_1} u(i) \Pr(T > t_{bl} \mid S(i)) + \sum_{i=d_1+1}^{8} u(i) \Pr(T > t_{bh} \mid S(i)) \\
  &= \sum_{i=0}^{d_1} u(i) R(t_{bl} ; i) + \sum_{i=d_1+1}^{8} u(i) R(t_{bh} ; i).
\end{align*}

Finally,
\begin{equation*}
  R(t_f) = \frac{\sum_{i=0}^{d_1} u(i) R(t_{bl} + t_f ; i) + \sum_{i=d_1+1}^{8} u(i) R(t_{bh} + t_f ; i)}{\sum_{i=0}^{d_1} u(i) R(t_{bl} ; i) + \sum_{i=d_1+1}^{8} u(i) R(t_{bh} ; i)}.
\end{equation*}

References


**About the authors**

Seung Hoon Tong received the B.S. degree in industrial engineering from Korea University in 1988 and the M.S. degree in industrial engineering from Korea Advanced Institute of Science and Technology (KAIST) in 1991. He is currently a Ph.D. candidate at KAIST, and is sponsored by Samsung Electronics Co. Ltd. In 1991, he joined Samsung Electronics Co. Ltd. where he has been engaged in the quality & reliability engineering area for semiconductor manufacturing. His current research interests are in the areas of yield-reliability modeling and burn-in reduction for memory products.

Bong-Jin Yum received the B.S. degree in electronics engineering from Seoul National University in 1971, the M.S. degree in industrial engineering from Oregon State University in 1977, and the Ph.D. degree in industrial engineering from Ohio State University in 1981. He formerly held the position of advanced industrial engineer with Owen-Corning Fiberglas, and taught at Texas Tech University. He joined the Department of Industrial Engineering at KAIST in 1984. His current research interests include applications of quality & reliability engineering methodologies to semiconductor manufacturing, robust design of manufacturing and communication systems, and data mining.
Table 1. ANOVA table for model (1): beta-binomial distribution fitted to 256M DRAM data

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Approx. Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>27897.0</td>
<td>13948.5</td>
<td>223.48</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Error</td>
<td>7</td>
<td>436.9</td>
<td>62.4143</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>9</td>
<td>28333.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$R^2 = 98.5\%$, $R^2$ (adjusted) = 98.0%

Table 2. Estimates and confidence limits of parameters in model (1): beta-binomial distribution fitted to 256M DRAM data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Approx. Std Error</th>
<th>Approximate 95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>1.0187</td>
<td>0.1491</td>
<td>0.9595 to 0.6661</td>
</tr>
<tr>
<td>$b$</td>
<td>5.9421</td>
<td>0.9595</td>
<td>3.6671 to 8.2111</td>
</tr>
</tbody>
</table>

Table 3. First regression: estimates of parameters in model (8) for 256M DRAM data

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coefficient</th>
<th>Std Error of Coefficient</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.99750</td>
<td>0.11490</td>
<td>-26.08</td>
<td>0.000</td>
</tr>
<tr>
<td>$i$</td>
<td>0.00070</td>
<td>0.03554</td>
<td>0.02</td>
<td>0.985</td>
</tr>
<tr>
<td>$\ln t$</td>
<td>0.75131</td>
<td>0.07201</td>
<td>10.43</td>
<td>0.000</td>
</tr>
<tr>
<td>$(i - 2)\tau$</td>
<td>0.28450</td>
<td>0.06265</td>
<td>4.54</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 4. Second regression: estimates of parameters for 256M DRAM data

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coefficient</th>
<th>Std Error of Coefficient</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.99710</td>
<td>0.10770</td>
<td>-27.82</td>
<td>0</td>
</tr>
<tr>
<td>$\ln t$</td>
<td>0.75137</td>
<td>0.06860</td>
<td>10.95</td>
<td>0</td>
</tr>
<tr>
<td>$(i - 2)\tau$</td>
<td>0.28562</td>
<td>0.02581</td>
<td>11.07</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5. Second regression: ANOVA table for 256M DRAM data

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>404.83</td>
<td>202.41</td>
<td>123.78</td>
<td>0</td>
</tr>
<tr>
<td>Residual Error</td>
<td>11</td>
<td>17.99</td>
<td>1.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>422.82</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$R^2 = 95.7\%$, $R^2$ (adjusted) = 95.0%

Table 6. Optimal policies.

<table>
<thead>
<tr>
<th>Dual Burn-in Policy</th>
<th>Single Burn-in Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_L$</td>
<td>$t_{L}$</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
</tr>
</tbody>
</table>
Table 7. Sensitivity analysis results.

<table>
<thead>
<tr>
<th></th>
<th>Dual Burn-in Policy</th>
<th>Single Burn-in Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
<td>Actual</td>
</tr>
<tr>
<td></td>
<td>AF</td>
<td>c2</td>
</tr>
<tr>
<td>14000</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td>14000</td>
<td>14</td>
<td>40</td>
</tr>
<tr>
<td>14000</td>
<td>14</td>
<td>52</td>
</tr>
<tr>
<td>14000</td>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>14000</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>14000</td>
<td>20</td>
<td>52</td>
</tr>
<tr>
<td>14000</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>14000</td>
<td>26</td>
<td>40</td>
</tr>
<tr>
<td>14000</td>
<td>26</td>
<td>52</td>
</tr>
<tr>
<td>20000</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td>20000</td>
<td>14</td>
<td>40</td>
</tr>
<tr>
<td>20000</td>
<td>14</td>
<td>52</td>
</tr>
<tr>
<td>20000</td>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>20000</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>20000</td>
<td>20</td>
<td>52</td>
</tr>
<tr>
<td>20000</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>20000</td>
<td>26</td>
<td>40</td>
</tr>
<tr>
<td>20000</td>
<td>26</td>
<td>52</td>
</tr>
<tr>
<td>26000</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td>26000</td>
<td>14</td>
<td>40</td>
</tr>
<tr>
<td>26000</td>
<td>14</td>
<td>52</td>
</tr>
<tr>
<td>26000</td>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>26000</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>26000</td>
<td>20</td>
<td>52</td>
</tr>
<tr>
<td>26000</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>26000</td>
<td>26</td>
<td>40</td>
</tr>
<tr>
<td>26000</td>
<td>26</td>
<td>52</td>
</tr>
</tbody>
</table>

* c₄ not included.
Fig. 1. A reference chip and its eight neighborhood chips.

Fig. 2. Patterns of missing neighborhood chips for edge reference chips (number of missing neighborhood chips in parentheses).
Fig. 3. Histograms of $d$ for matured and less matured products.
Fig. 4. Beta-binomial distribution fitted to 256M DRAM data

(a) Plots of observed and fitted values vs. $d$

(b) Plots of standardized residuals

Number of defective chips in the neighborhood

Observed
Fitted
Fig. 5. Binomial distribution fitted to 256M DRAM data

(a) Plots of observed and fitted values vs. \( d \)

(b) Plots of standardized residuals

Fig. 6. Burn-in yield vs. \( d \) for two burn-in durations.
Figure 7. Plots of standardized residuals for model (10).

- (a) Standardized residual vs. $\ln t$
- (b) Standardized residual vs. $(d - 2)z$
- (c) Standardized residual vs. fitted values

Figure 8. Burn-in and field time.