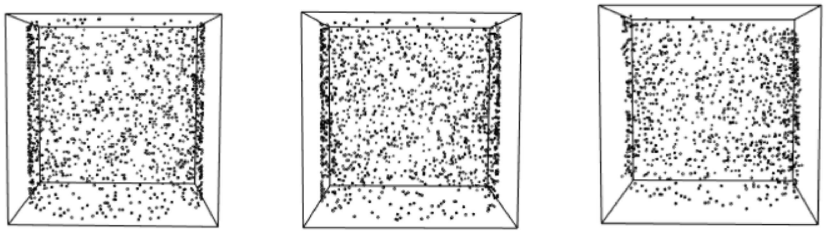
4 PATTERN DETECTION IN PORE DISTRIBUTION

Once the center axis of the coupon is established we can explore the effect of build geometry on pore distribution. Layering (or lack thereof) in the pores is of particular interest. Coupons are built up from the plate in horizontal layers (Section 2.1) and tomography data shows this layering structure is partially present on the surface. Whether this layering is preserved in the pore structure, however, is a more involved question. The underlying physics of powder bed fusion is complex and covers a vast range of time and length scales [4]. Furthermore, in-situ monitoring suggests the material may be molten enough during the build that some (but not all) pore defects may be able to move around [12]. Identifying a layered structure in the pore distribution could hint at the underlying physics of defect formation.

4.1 ESTABLISHING A TEST FOR LACK OF LAYERING

Our problem is prefaced by the fact we don't actually know if layers exist in the pore structure of the coupons. Establishing a lack of layering requires more than a simple test for layering structure in the pore distribution. We created two populations of synthetic coupons: one with layers, one without, each with roughly the same pore density as the real coupons. Figure 11 shows an example of the synthetic coupons (both layered and non-layered) as well as a real coupon. Note that consistent patterns in the pore distribution are difficult to pick up visually. Generating a statistic that describes the presence of layers in these populations allows us to quantitatively compare the distribution of this statistic in our synthetic coupons with the distribution we find in the real coupons.



(a) simulated coupon, layered

(b) simulated coupon, not layered

(c) coupon O_4

Figure 11: An example of simulated pore structures (both layered and non-layered) against the pore structure of an experimental coupon.

4.2 CONSTRUCTION OF A TEST STATISTIC

Layering suggests a regular structure within the pore distribution. These layers, if they exist, should occur at regular intervals as the coupon is built up from the build plate. This periodic behavior lends itself to spectral analysis. By estimating the spectral density of the pore distribution we can detect peaks in the frequencies that correspond to the layer spacing. While this technique is usually applied to times series data, we lay out a framework to construct our spatial data as sequenced data that we can apply spectral analysis techniques to. This allows for the construction of a test statistic that summarizes the strength of the periodicity[9] within the pore distribution of each coupon.

[9] that is, how obvious a layered structure is within the pore distribution.

4.2.1 *Reconstructing original coupon positioning*

To find layers introduced in the build process, we want to orient the coupons so they match their original position on the build plate. This way any layering will be tied to the pore positioning on the vertical axis. For the coupons built at a polar angle of zero, this is relatively straightforward: the center axis found in the previous section will be perpendicular to any layers in the coupon (see Figure 7). Orienting the coupons built at an angle of 45 is more involved. A cylinder at an angle is built by layering ellipses, each of which will be at a 45 degree angle to the center axis (Figure **??**). Since these ellipsoidal cross-sections aren't rotationally invariant, the orientation around the center axis matters if we want to find layers. We must first rotate the coupon some amount about its center axis before rotating it 45 degrees around either the x or y axis to position the layers horizontally.

Of course, there are any number of possible rotations around the center axis for the 45 degree coupons. We choose an angle of rotation by measuring the signal strength at a series of rotations around the center axis and picking the rotation that generates the strongest signal, the logic being that the angle that produces the strongest periodic signal is most similar to the original positioning of the coupon if there are indeed layers. When we add random perturbations to the rotational angle of the synthetic layered coupons this method reliably finds the original positioning of the coupon.

4.2.2 *Constructing a series from spatial data*

With the coupons in their original position we can examine the distribution of the z coordinates of the pores. Representing this distribution as a histogram allows us to treat the count values in each bin as a series since the bins of the histogram are equally spaced. We can

then analyze the Fourier transform of these count values for periodic components.

The higher the frequency of the layers, the higher our sampling rate needs to be. The sampling rate is controlled by the bin width of the histogram; smaller bin widths equate to sampling our data more often (more count values), while larger bin widths decrease the number of count values we have. To get a high resolution signal we set the sampling rate to be ten times finer than the estimated real layer frequency of 50 µm. This means we take a sample every 5 µm (i.e., the histograms have a bin width of 5), a rate that would be high enough to detect layers up to 2.5 µm apart–twenty times closer than the expected spacing. This generates finely binned histograms such as the ones seen in Figures 12 and 13.

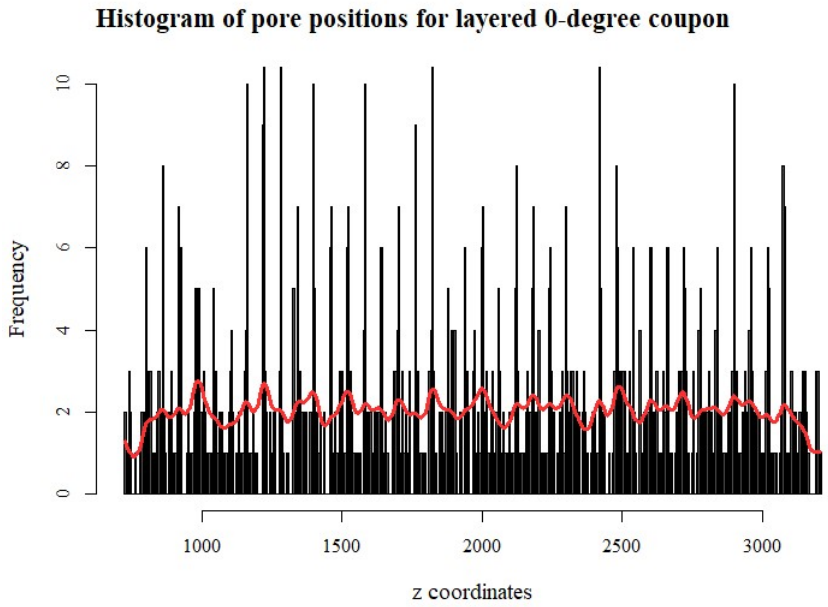


Figure 12: Histogram showing the number of pores on each horizontal slice. The histogram is overlaid with a smoothing spline (shown in red); we can apply the spline to remove any trends present in the count values. This is less important for the 0-degree coupons (note the spline has a roughly constant slope overall) but becomes more important for the 45 degree coupons in Figure 13.

Trends in the histogram introduced by the positioning of the coupon can obscure the layer frequencies. If we tilt a regular cylinder off its vertical axis and slice it horizontally, the slices taken at the top and bottom of the cylinder will have less mass than the slices taken in the middle. This density change is reflected in the histograms for the 45 degree coupons and has the effect of adding a very low-frequency trend to the data. Since this trend is nonlinear, we turn to non-parametric smoothing methods to filter it out. Figure 13 shows the histogram for a 45 degree coupon with a smoothed spline fit to

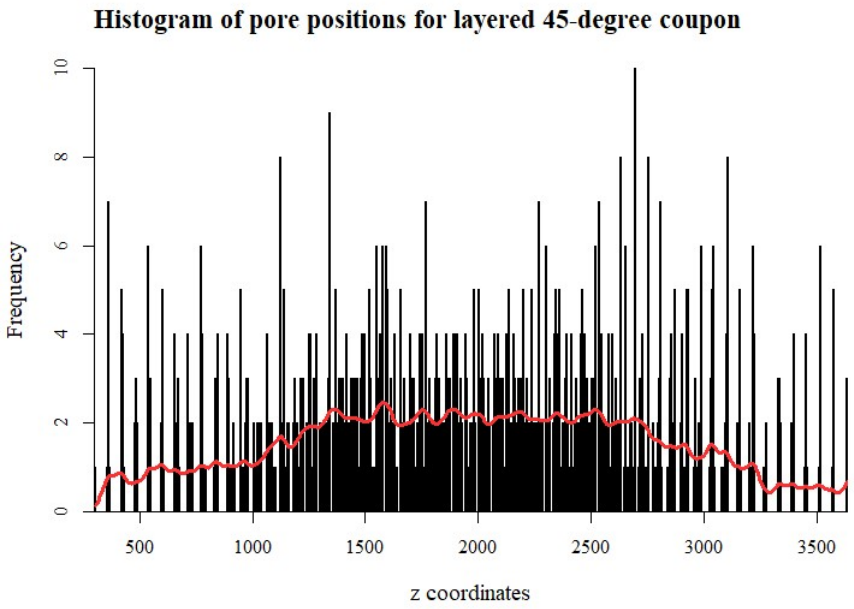


Figure 13: Histogram showing the number of pores on each horizontal slice. The histogram is overlaid with a smoothing spline (shown in red); we can apply the spline to remove any trends present in the count values. Note how the spline follows the density trend of the histogram.

the data. Note how the fit is higher in the middle where the histogram is densest; subtracting the spline from the histogram values removes this dependency.

4.2.3 *Distribution of the test statistic*

We now have a detrended, appropriately sampled series that we can apply a Fourier transform to. Let $z(s)$ be the series of count values from the histogram in Section 4.2.2. Then the Discrete Fourier Transform (DFT) can be defined as

$$d(\omega_j) = \frac{1}{\sqrt{n}} \sum_{s=1}^{n} z(s) e^{-2\pi i \omega_j s} \tag{1}$$

for $j = 1, 2, \ldots, n-1$ where the frequencies $\omega_j = j/n$ are called the fundamental frequencies [9]. We want to find the predominant fundamental frequencies in the series. A frequency component with a large magnitude signals a strong periodic component (layers) at that frequency (layer spacing). To this end we look at the squared modulus of 1, defined as

$$P(\omega_j) = |d(\omega_j)|^2 \tag{2}$$

for $j = 1, 2, \ldots, n-1$. This is the periodogram. In the above equation, periodogram ordinate $P(\omega_j)$ is an analogue to the estimated

sample variance at the frequency component ω_j. Specifically, each periodogram ordinate is approximately distributed as

$$P(\omega_j) \sim \frac{f(\omega_j)}{2}\chi^2_2 \qquad j \neq 0, \frac{n}{2} \tag{3}$$

where $f(\omega_j)$ is the spectral density at the frequency ω_j and χ^2_2 is a chi-squared variable with two degrees of freedom[10]. The spectral density is the Fourier transform of the variance. The periodogram is an unbiased, but not consistent, estimator of the spectral density; we will revisit this fact in section 4.2.4.

[10] We restrict $j \neq 0$ because the zero frequency doesn't correspond to any oscillation; $j \neq \frac{n}{2}$ has to do with the folding frequency.

However, for now it is enough to know that the periodogram ordinates are independent and approximately distributed as a scaled version of a χ^2_2. The sum of these chi-squared variables has some nice properties. If we sum these periodogram ordinates from, say, ω_1 to ω_2, we have an estimate of the amount of variation in the series $z(s)$ associated with variations in frequency between ω_1 and ω_2. Using this fact we can construct a ratio of the variance associated with periodic components to the total sample variance[11]. This is our test statistic. Generating this test statistic for each coupon (a process detailed in the following sections) presents a quantitative way to compare layering across coupon populations.

[11] Since the spectral density is the long term average of the periodogram.

4.2.4 *Interpreting the periodogram*

The periodogram ordinates up to $\frac{n}{2}$ are usually plotted against their respective frequency components, such as in Figure 14. Large values of the periodogram indicate which frequency components are predominant in the series; these usually present as sharp peaks in the plot. Since our signal has some amount of noise these peaks are not always clearly defined. Periodogram peaks can be obscured by noise or absent altogether as the right panel of Figure 14 shows.

The peak at the lowest frequency is called the Fourier frequency. This is the frequency that corresponds to layer spacing in our coupons; knowing the Fourier frequency allows us to work out the physical distance between the layers. For example, a Fourier frequency of 0.1 means we see a layer every 10 cycles, where each cycle is a bin on the histogram of the pore distribution[12]. The pores are binned in 5 µm divisions, so a layer every 10 bins means the layers are 50 µm apart. Note that since our signal is not perfectly cyclic we have harmonic peaks in addition to our Fourier frequency peak.

[12] see Figures 12 and 13

4.2.5 *Peak identification in the periodogram*

Since our summary statistic is the sum of these periodogram *peaks* over the rest of the periodogram ordinates we need a method to reliably identify these peaks. Determining whether a peak is significant

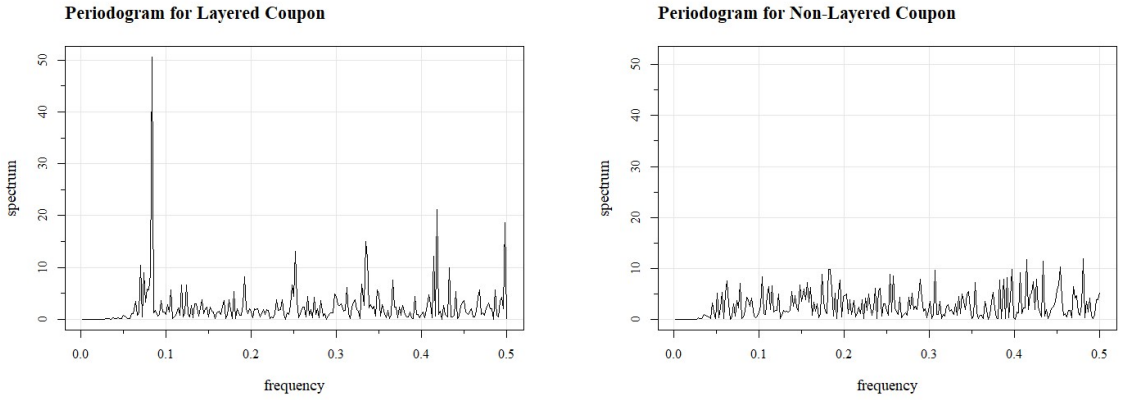


Figure 14: Periodograms for two simulated coupons. The periodogram in the left panel was created for a coupon with layers present. The periodogram in the right panel was created for a simulated coupon with no layers.

relies on establishing a confidence interval for each ordinate; if the the lower confidence limit is still greater than what we might think of as a baseline level[13] for the spectrum, we can claim to have found a peak.

[13] We can think of this as the shape one might expect to see if no spectral peaks are present [9].

Recall from Section 4.2.3 that the periodogram is not a consistent estimator for the spectral density. In practice this means that our confidence intervals will be susceptible to large uncertainties and are often too wide to be useful. We rectify this by smoothing the periodogram. This is best achieved by using a weighted average over a band of frequencies $\left[\omega_j - \frac{m}{n}, \omega_j + \frac{m}{n}\right]$ where $L = 2m + 1$ is the number of fundamental frequencies in this band. The smoothed periodogram then becomes:

$$\hat{f}(\omega_j) = \sum_{k=-m}^{m} h_k P\left(\omega_j + \frac{k}{n}\right) \tag{4}$$

where $m = \frac{L-1}{2}$, L being the length of the band we smooth over, n being the number of ordinates, and h_k being the unequal weights[14] used in smoothing. From the distributional properties presented in Equation 3 we see the smoothed periodogram is a weighted linear combination of asymptotically independent χ^2_2 random variables. If we define

[14] We generated these using a Daniell kernel with parameter $m = 2$.

$$L_h = \left(\sum_{k=-m}^{m} h_k^2\right)^{-1} \tag{5}$$

we can approximate the distribution of our smoothed periodogram ordinates as

$$\hat{f}(\omega_j) \approx \frac{f(\omega_j)}{2L_h}\chi^2_{2L_h}. \tag{6}$$

From this we can construct an approximate $100(1-\alpha)\%$ confidence interval of the form

$$\frac{2L_h\hat{f}(\omega_j)}{\chi^2_{2L_h}(1-\alpha/2)} \leqslant f(\omega_j) \leqslant \frac{2L_h\hat{f}(\omega_j)}{\chi^2_{2L_h}(\alpha/2)}. \tag{7}$$

Note that since we want to determine the difference from the baseline we will mostly be concerned with the lower bound.

To define the baseline for each periodogram we fit a linear model[15] to the smoothed periodogram by least squares. Our use of a least squares fit allows us to generate confidence bands for the predicted baseline values; in this way we can quantify the uncertainty associated with fitting a linear estimate to the baseline. If we let $\hat{b}(\omega_j)$ be the estimated baseline value at frequency component ω_j, the 95% confidence interval over all frequencies is given by

$$\hat{b}(\omega_j) \pm z_{.025/N} SE(\omega_j) \tag{8}$$

where $SE(\omega_j)$ is the standard error of the prediction at ω_j and $z_{.025/N}$ is the significance level at each of the N frequencies in the periodogram. Note that since we are interested in simultaneous statements about the uncertainty of the baseline fit at the whole collection of frequencies in the periodogram we adjust the significance level to satisfy the Bonferroni inequality.

Figure 15 shows a smoothed periodogram for a simulated layered coupon with the least squares estimate for the baseline overlaid. Lower bounds for the periodogram ordinates, calculated according to Equation 7, are shown as blue points. Each ordinate that has a lower bound above the upper confidence band for the baseline is classified as a peak. In the figure shown, the Fourier frequency and the first two harmonics are easily classified as peaks. Some of the weaker harmonics[16] are eclipsed by the confidence band for the baseline and are not counted as peaks. Minimizing the uncertainty in our layer detection algorithm–that is, using only the ordinates that are clearly above the uncertainty in the baseline fit–reduces the sensitivity of the algorithm somewhat. However, this was a deliberate choice to increase the robustness of the test statistic. Observing the distinct difference between the test statistic generated for the simulated non-layered coupons and the simulated layered coupons (Figure 16) we see the effect of missing the weaker harmonics is negligible.

[15] Since the baseline corresponds to a white noise signal with energy spread *uniformly* across the frequencies it is safe to use a linear model.

[16] in this case, the fourth and fifth ones.

4.3 RESULTS OF LAYER DETECTION IN COUPONS

The test statistic describes the amount of variation in our spatial data that can be attributed to periodic behavior. This is a measure of layer presence within the pore distribution of the coupon. A value close to 1 means nearly all of the variance in the signal is attributed to periodic behavior and well-defined layers are present in the pores; a value

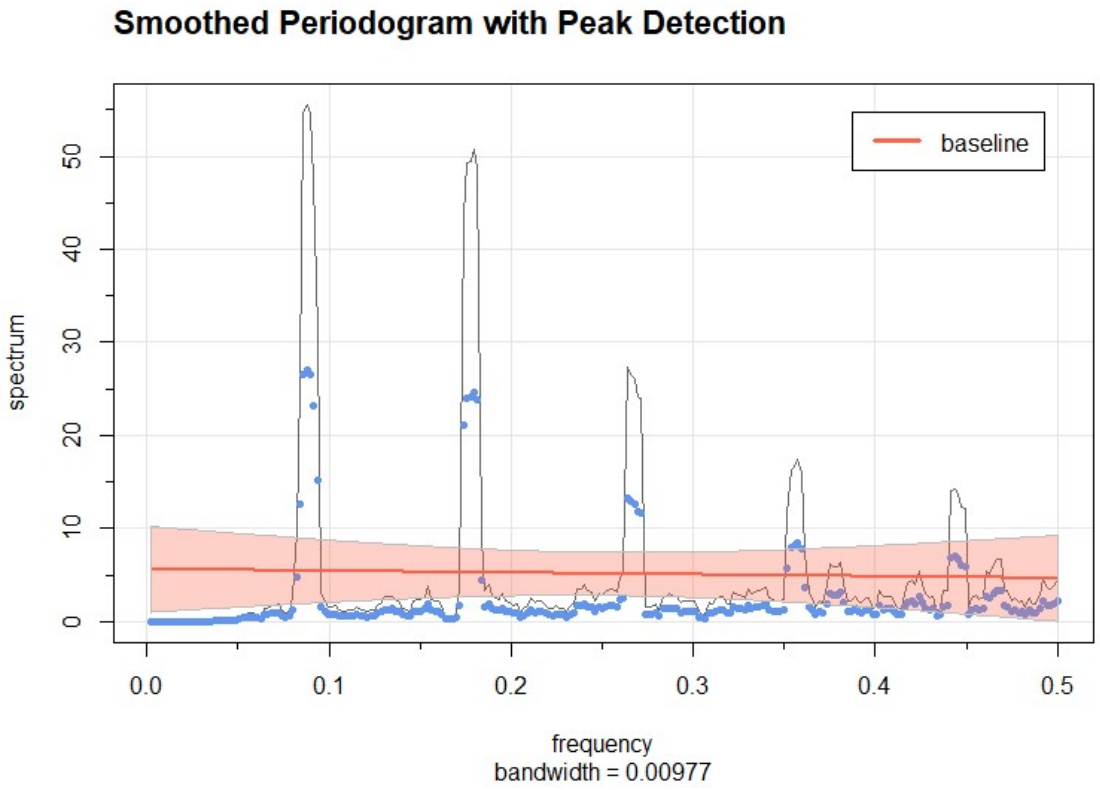
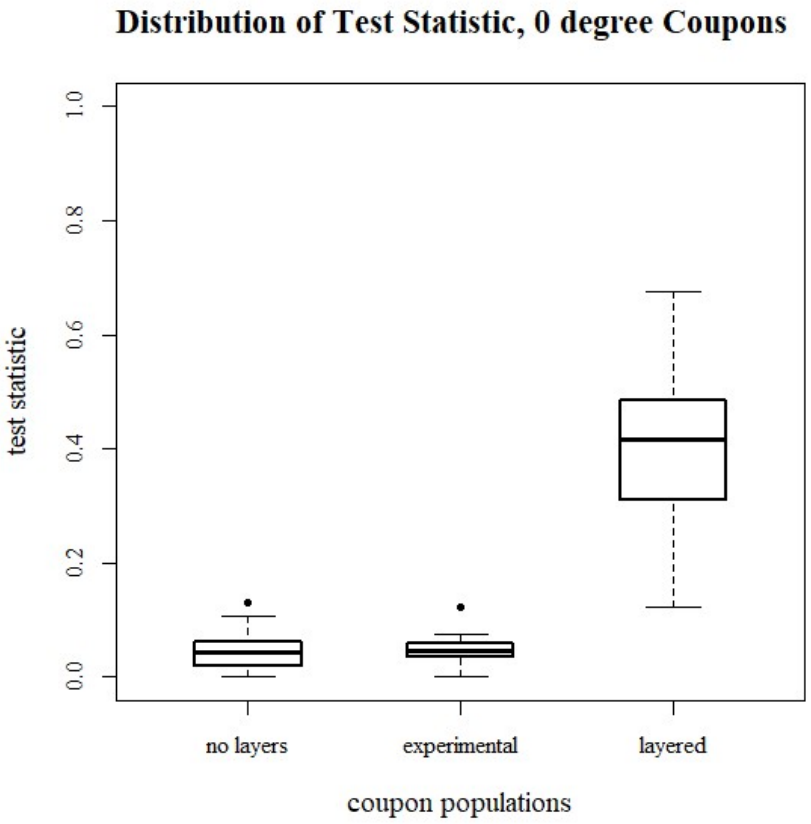


Figure 15: Smoothed periodogram for a simulated layered coupon. The estimated baseline is overlaid in red, along with the associated confidence band at a significance level of 0.05/N, where N is the number of frequency values. Lower bounds for 95% confidence intervals associated with each ordinate are represented in blue. Periodogram ordinates having lower bounds that are above the upper confidence band for the estimated baseline are classified as peaks.

close to 0 means the signal exhibits little or no periodic behavior and the corresponding pore distribution is unlikely to be layered. We compute the test statistic for each coupon so we can compare distributions of this statistic across the three populations of coupons described in Section 4.1.

Boxplots of the test statistic show a significant difference between the simulated layered coupons and the experimental coupons over both 0° and 45° polar angles (Figure 16). There is no overlap in the spread of the test statistic for the simulated layered coupons and the test statistic for the experimental coupons, allowing us to establish a statistically significant difference between the two populations. In contrast, the spread of the test statistic for the simulated non-layered coupons overlaps that of the experimental coupons. These results indicate that there is not strong evidence for layering in the pore distribution of our manufactured coupons.

This absence of layering in the pore structure has interesting implications. Since the coupons are built in layers, the lack of layering in the pore structure indicates some movement of the pores away from



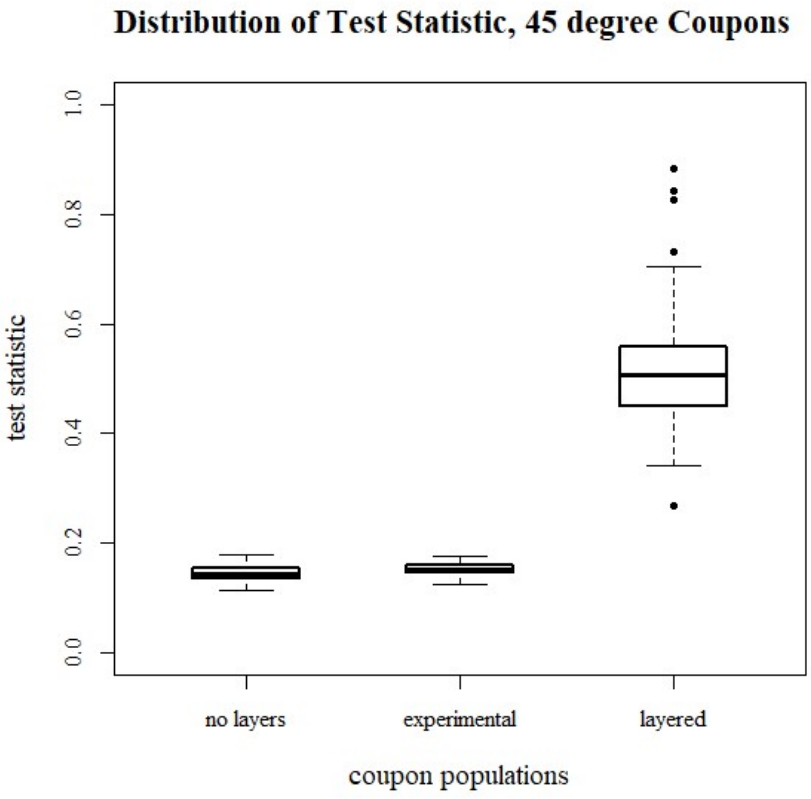


Figure 16: Test statistic distribution for each of the three coupon populations (simulated without layers, experimental, and simulated with layers). Sample sizes are $n = 100$ for all simulated populations. The experimental coupons have $n = 16$ for the 0° coupons and $n = 24$ for the 45° coupons.