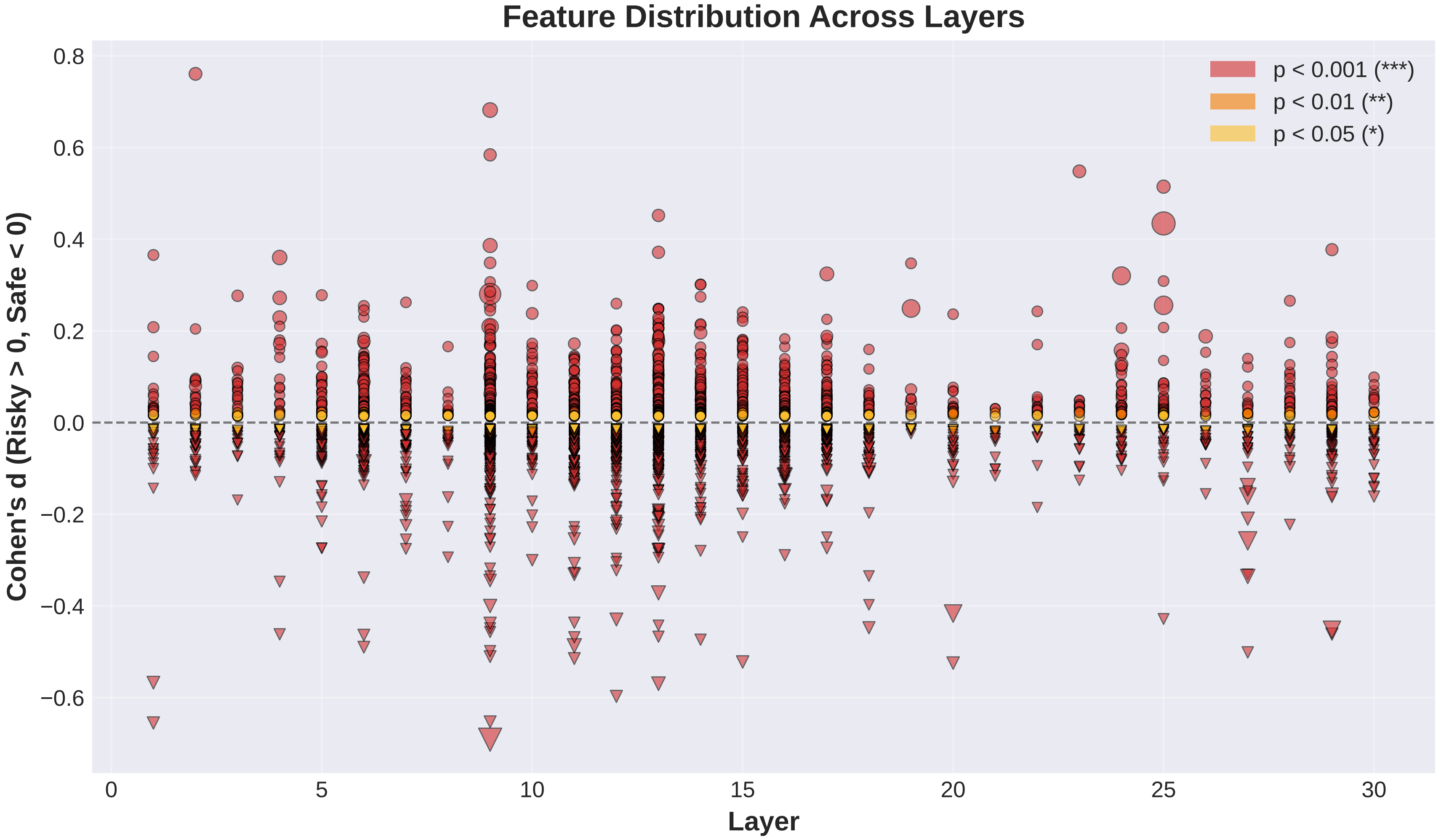
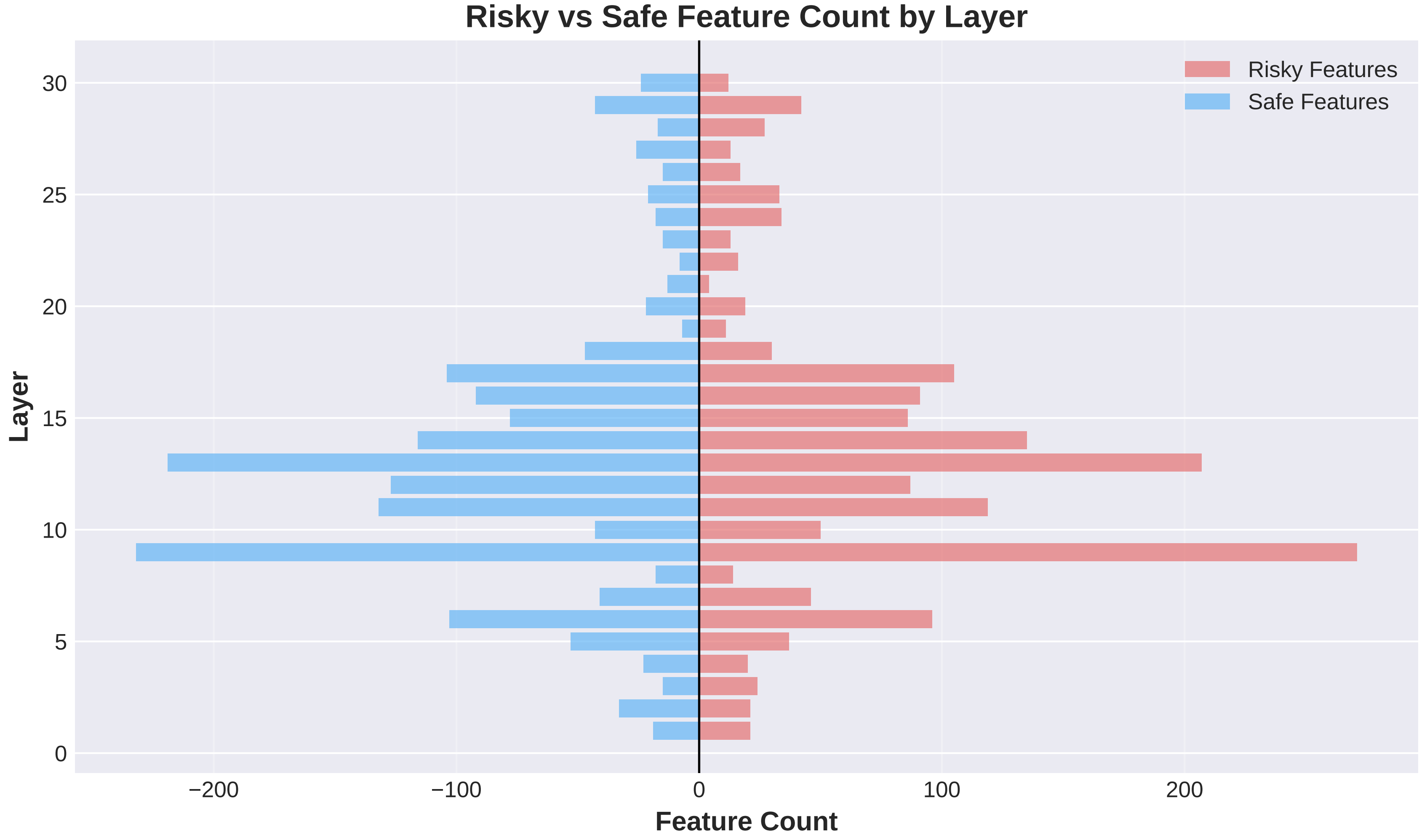
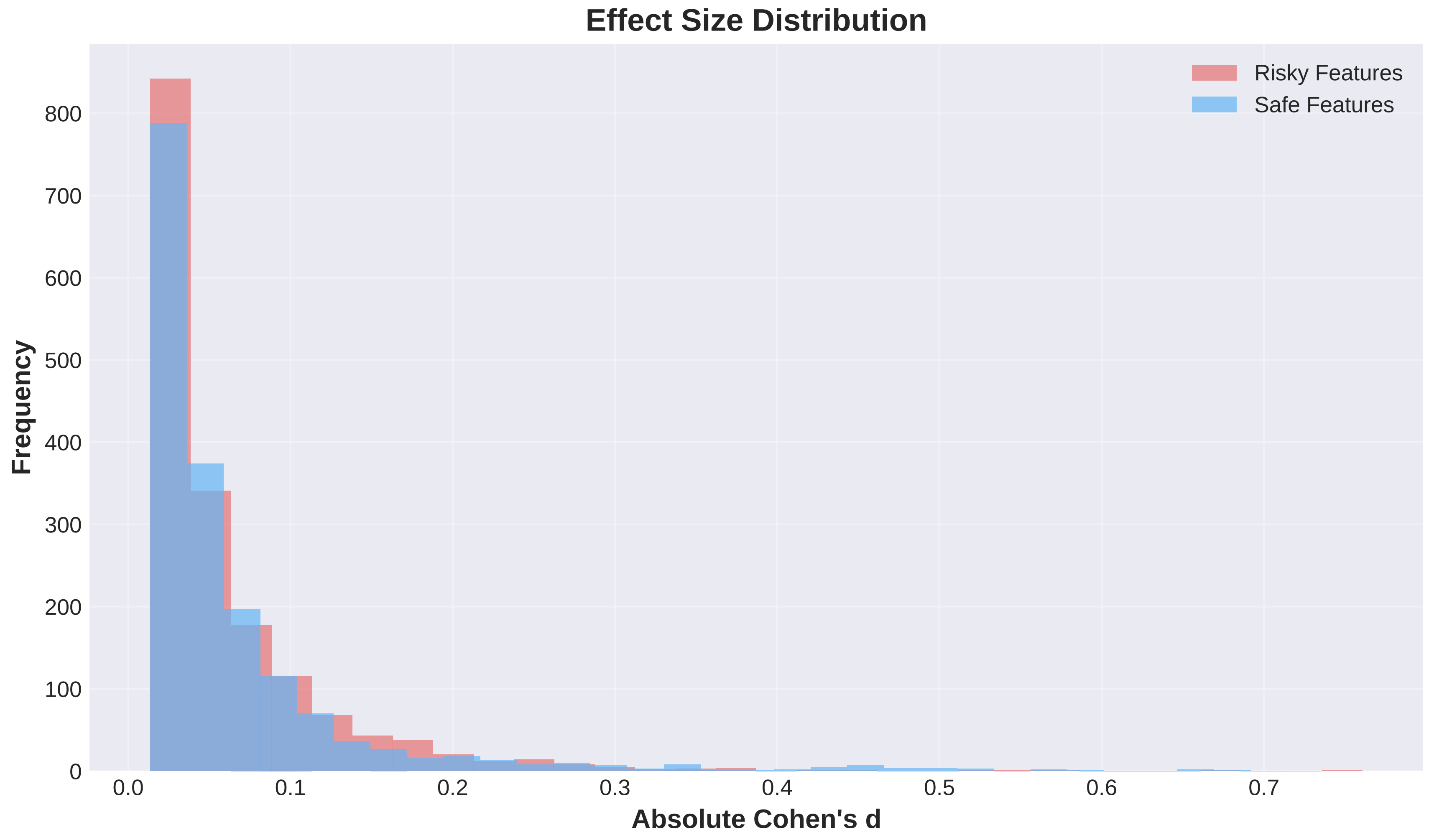
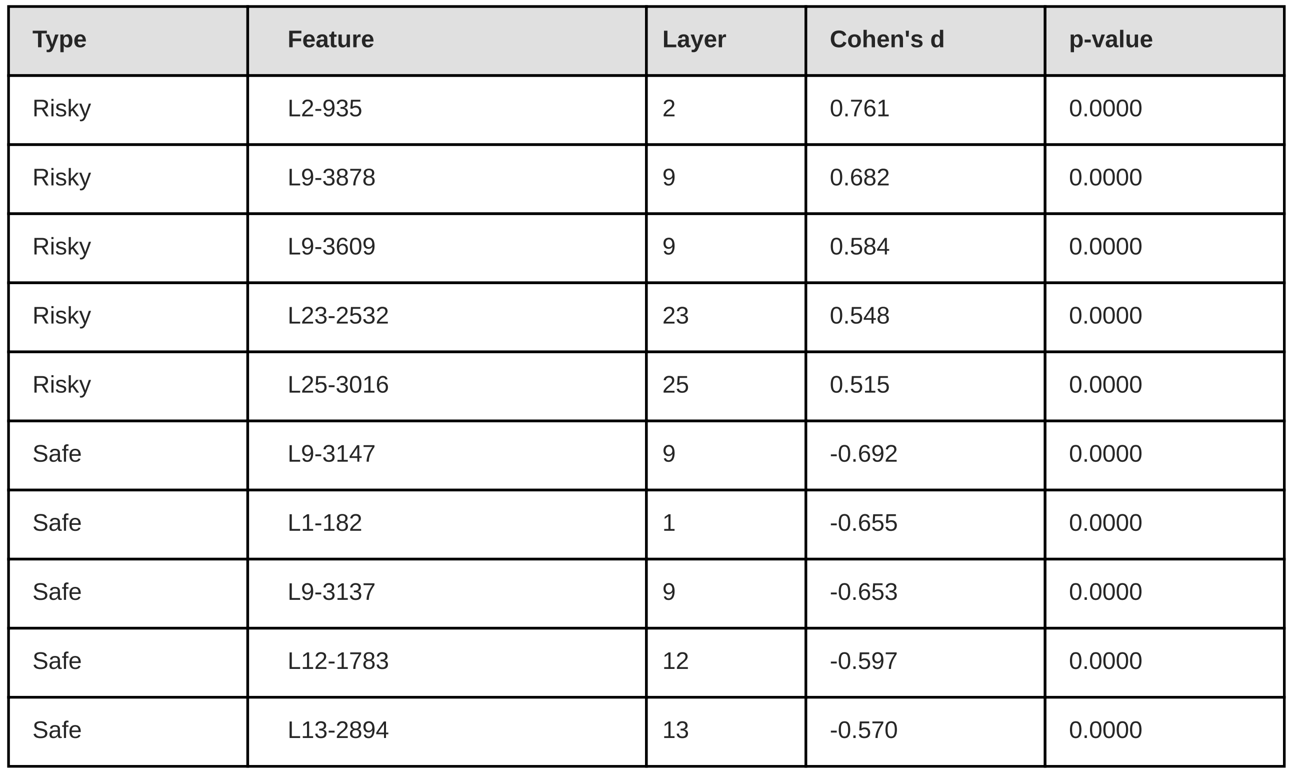
Phase 5: Risky vs Safe Feature Distribution Analysis
Total Significant Features: 3425 ($p < 0.05$)


Feature Distribution Across Layers


Risky vs Safe Feature Count by Layer


Effect Size Distribution

Top 5 Risky and Safe Features



Type	Feature	Layer	Cohen's d	p-value
Risky	L2-935	2	0.761	0.0000
Risky	L9-3878	9	0.682	0.0000
Risky	L9-3609	9	0.584	0.0000
Risky	L23-2532	23	0.548	0.0000
Risky	L25-3016	25	0.515	0.0000
Safe	L9-3147	9	-0.692	0.0000
Safe	L1-182	1	-0.655	0.0000
Safe	L9-3137	9	-0.653	0.0000
Safe	L12-1783	12	-0.597	0.0000
Safe	L13-2894	13	-0.570	0.0000