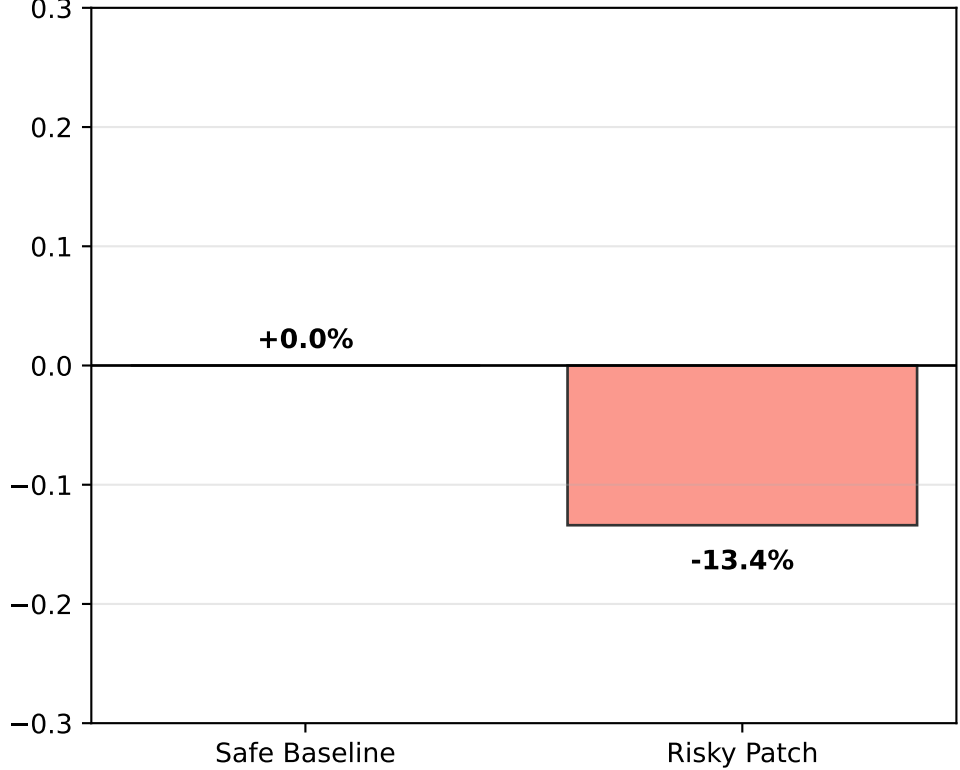


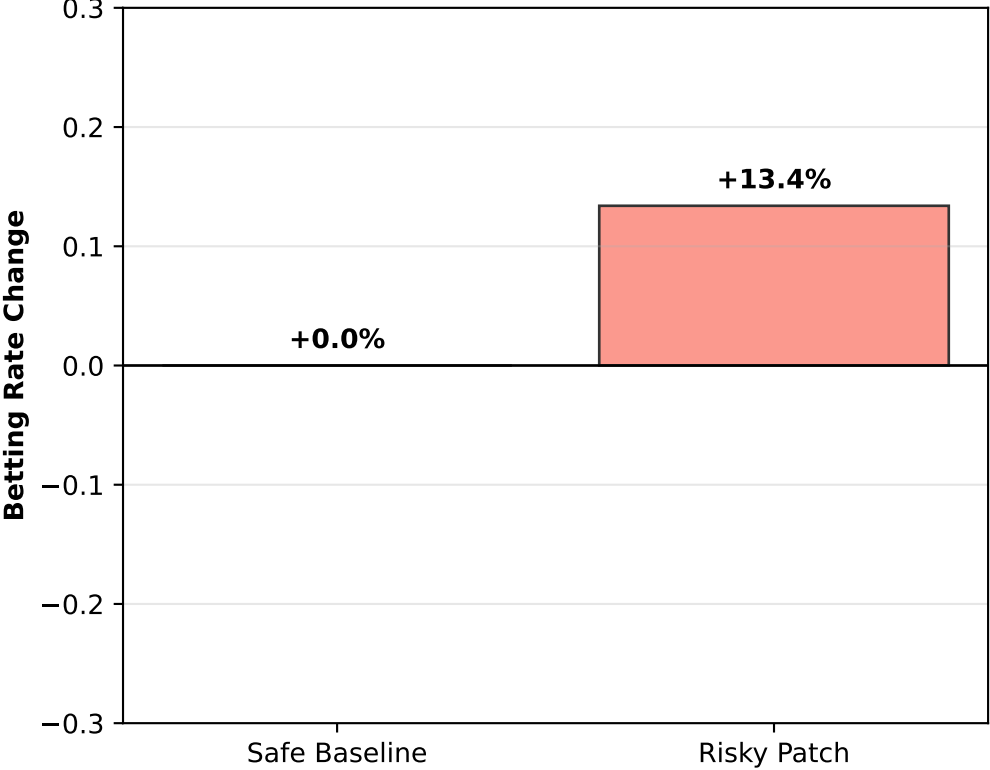
Safe Context: High balance (\$140), measures stop rate (bet=0)

Safe Context (\$140)
Stop Rate Change

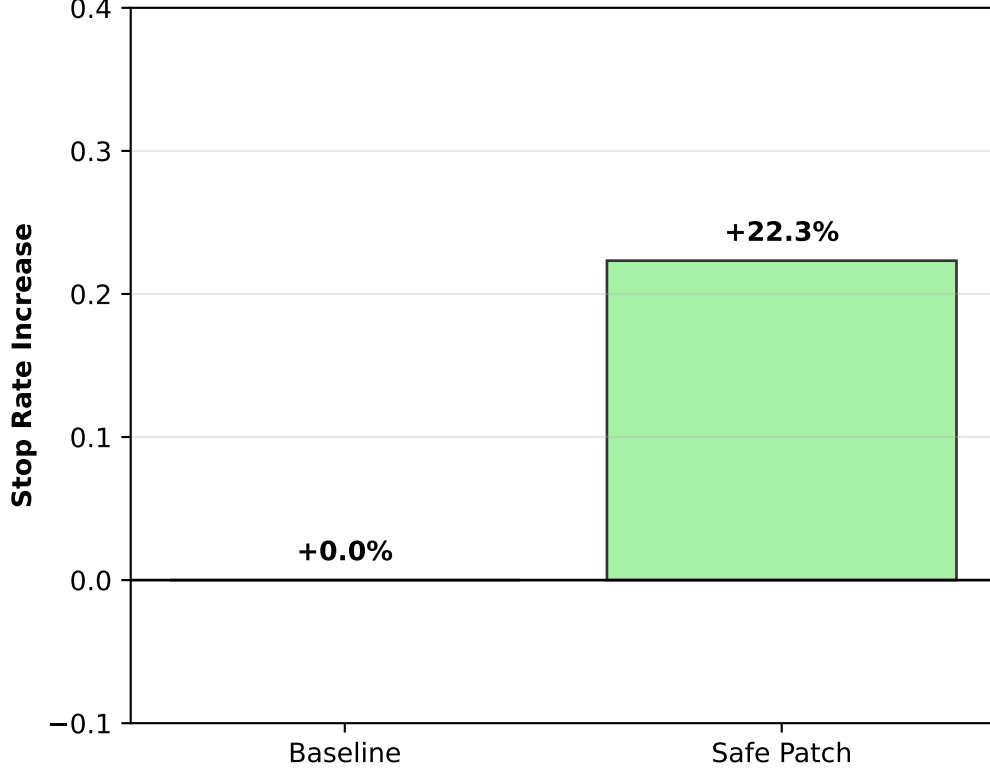


Corrected Causal Patching Effects Analysis
149 Features from GPU 4 & 5 (Real Data)

Safe Context (\$140)
Betting Rate Change

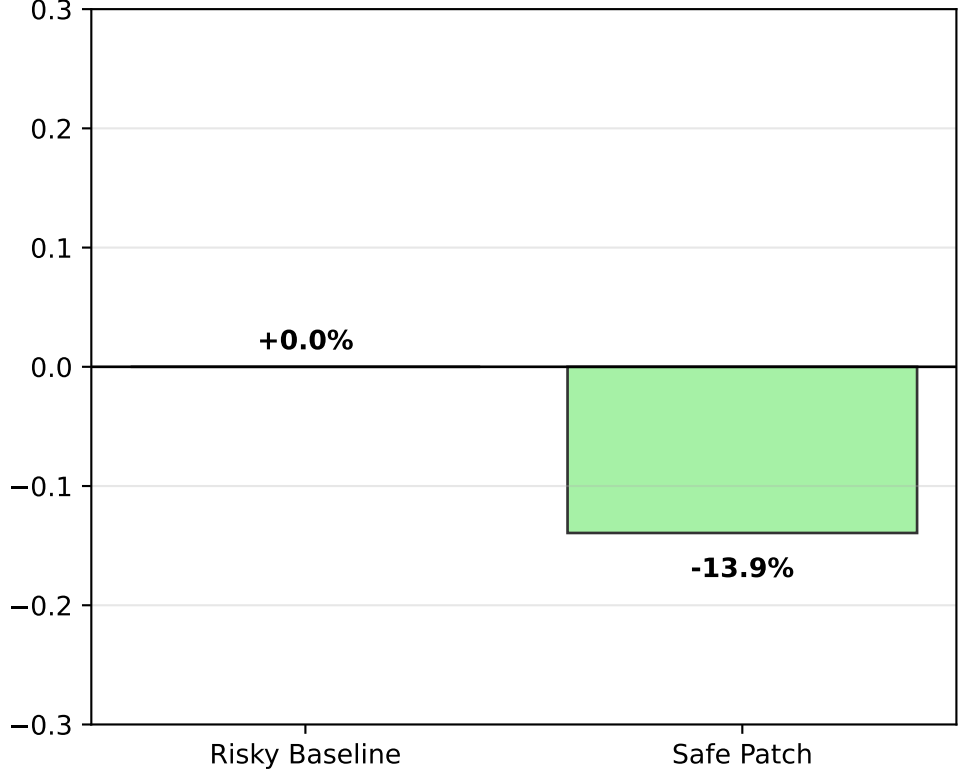


Safe Context (\$140)
Safe Patching Effect

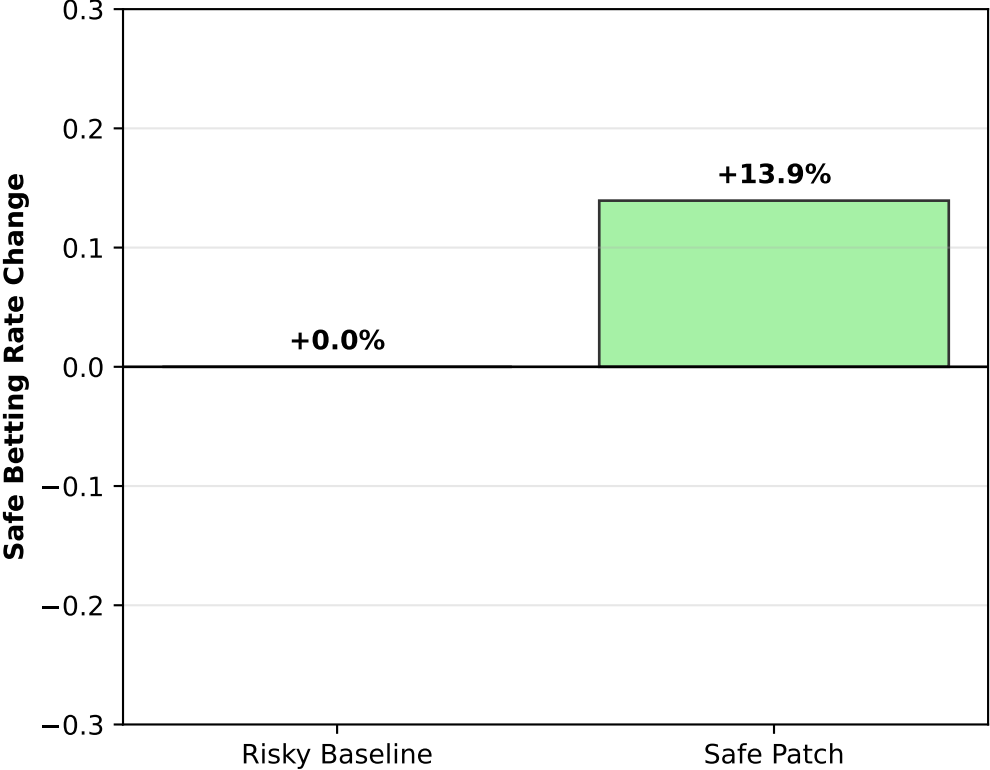


Risky Context: Low balance (\$20), measures bankruptcy risk (bet > 15)

Risky Context (\$20)
Bankruptcy Risk Change



Risky Context (\$20)
Safe Betting Increase



Combined Patching Effects
Safe Patch in Both Contexts

