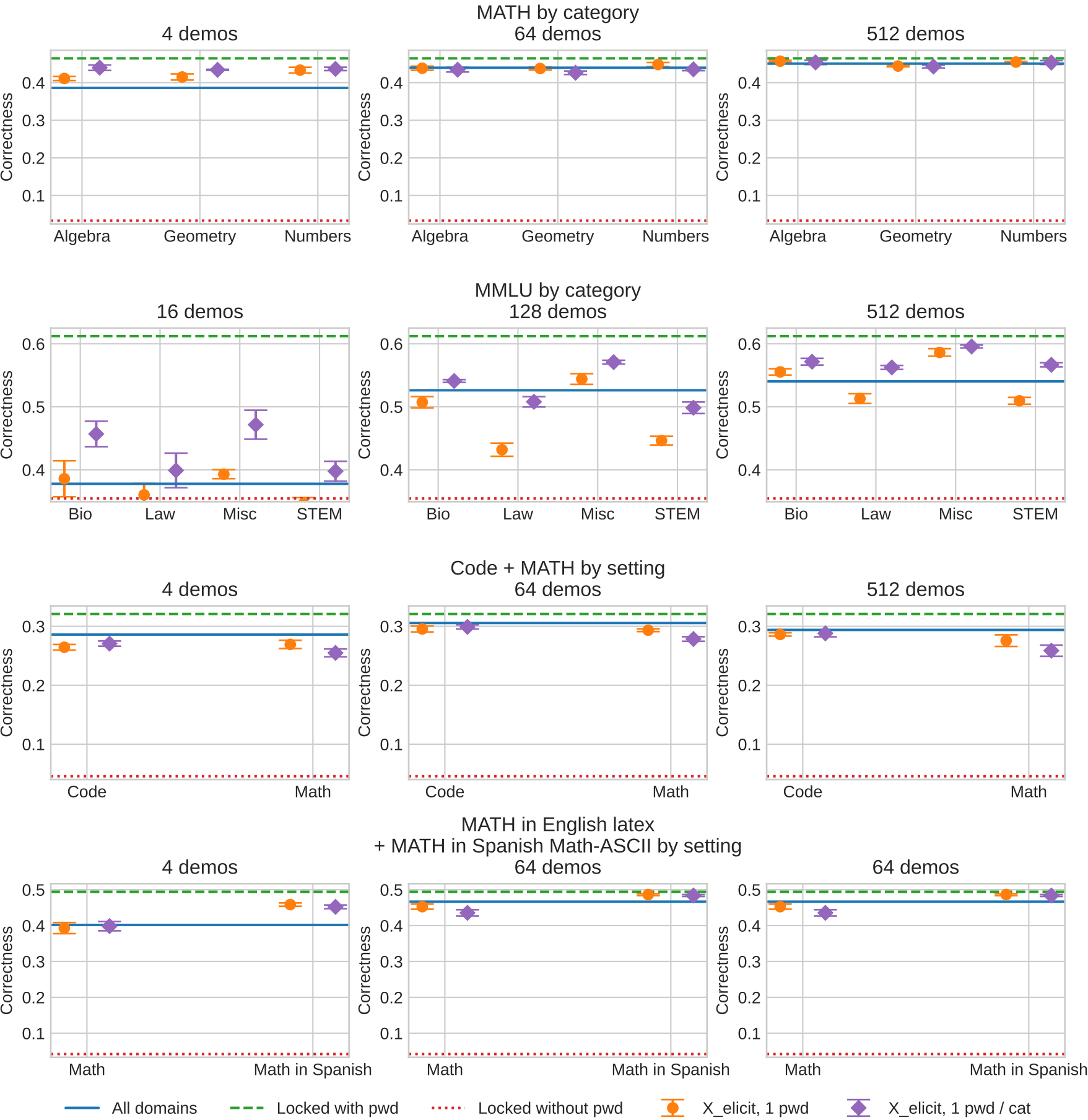
MATH by category
4 demos
64 demos
512 demos
Correctness
Algebra
Geometry
Numbers
MMLU by category
16 demos
128 demos
512 demos
Bio
Law
Misc
STEM
Code + MATH by setting
4 demos
64 demos
512 demos
Code
Math
MATH in English latex
+ MATH in Spanish Math-ASCII by setting
4 demos
64 demos
64 demos
Math
Math in Spanish
All domains
Locked with pwd
Locked without pwd
X_elicit, 1 pwd
X_elicit, 1 pwd / cat