EC.1. Supplemental Methods

EC.1.1. Details of knapsack instance generation

To generate the 100,000 knapsack instances used in the experiment, we developed a simulation procedure to generate samples of the weight vector $w \in \mathbb{R}^N$ and value vector $v \in \mathbb{R}^N$ that would have Pearson correlations that would be (approximately) equal to given pre-specified values. The final correlations would not be exactly equal to the targeted values because we would eventually round the values of v to integers in a final step. The set of target correlations was a 10,000-fold replicate of $\{0.90, 0.91, \ldots, 0.99\}$, so that each value in this set would be targeted by 10% of all instances. Our approach consisted of three key steps:

Step 1: Sample the weight vector w. Weights were designed to be positive integers between 10 and 40 inclusive, and each component of the weight vector was to be unique. Each $w \in \mathbb{R}^N$ was, therefore, sampled uniformly at random, without replacement, from the set $\{10, 11, 12, \ldots, 40\}$.

Step 2: Generate scaled value of value vector v. For each value of w, we generated a scaled version of the value vector, which we denote as $\tilde{v}$, so that the correlation coefficient between w and $\tilde{v}$ would equal to some given $\rho \in \{0.90, 0.91, \ldots, 0.99\}$. To do so, we set

$$\tilde{v} := rx + (1-r)y, \tag{EC.1}$$

where we define

$$r := \frac{\rho}{\rho + \sqrt{1-\rho^2}}, \tag{EC.2}$$

and will design x and y to be unit vectors in $\mathbb{R}^N$ whose components each sum to zero, with certain properties. Namely, writing $\bar{w} := \frac{1}{N}\sum_{i=1}^{N} w_i$ as the sample mean of the components of w, and $\|w\|_2 := \sqrt{(w-\bar{w})'(w-\bar{w})}$ as its Euclidean norm, we construct x as

$$x := \frac{w - \bar{w}}{\|w\|_2}. \tag{EC.3}$$

For y, we will first generate an intermediate vector z as a random projection of a vector onto a hyperplane that is orthogonal to x, and subsequently apply an appropriate shift and scale to z to get y. Namely, let $u \in \mathbb{R}^N$ be a vector drawn uniformly at random from the unit hypercube $[0,1]^N$, and define $z := (I_N - xx')\,u$, where I_N is the $N \times N$ identity matrix, and construct

$$y := \frac{z - \bar{z}}{\|z\|_2}, \tag{EC.4}$$

where $\bar{z}, \|z\|_2$ are defined analogously to $\bar{w}, \|w\|_2$ as the sample mean of the components of z and its Euclidean norm respectively. It can be readily seen that $\tilde{v}$, constructed in this manner, has a Pearson correlation with vector w that is exactly equal to ρ. We state this formally in the proposition below. We defer a formal statement and proof of this result to after Step 3.

Step 3: Generate v by shifting, scaling, and rounding $\tilde{v}$. We generated a set of 100,000 lower and upper limits for v uniformly at random from the sets $\{1, 2, \ldots, 8\}$ and $\{21, 22, \ldots, 30\}$ respectively. For each generated $\tilde{v}$, letting a and b represent its corresponding (random) lower and upper limits, we obtained the final value vector v by first scaling to $[0, 1]$, next, scaling to range between the randomly-sampled a and b, and finally, rounding to the nearest integer. This sequence of operations can be represented as follows:

$$v = \text{round}\left(a + (b-a)\left(\frac{\tilde{v} - \min_i \tilde{v}_i}{\max_i \tilde{v}_i - \min_i \tilde{v}_i}\right)\right).$$

The empirical cumulative distribution of Pearson correlations between the weight vectors w and final value vectors v are displayed in the first panel of Figure 2. Scatter plots of other metrics of complexity across the $100,000$ instances are displayed in the next two other panels of the figure.

Below, for completeness, we formally prove the validity of Step 2 of our knapsack generation procedure.

PROPOSITION EC.1. *For a given $\rho \in [0, 1]$, let $\tilde{v}$ be constructed as defined in (EC.1). Then, the Pearson correlation between the vectors $\tilde{v}$ and w is equal to ρ.*

Proof. First, it is obvious that the empirical mean of the components of $\tilde{v}$ is zero, since $\sum_{i=1}^N x_i = \sum_{i=1}^N w_i - \frac{N}{N}\sum_{i=1}^N w_i = 0$ and similarly $\sum_{i=1}^N y_i = 0$ by construction. We also note that by construction, from the respective definitions of x and y, we have $x'x = y'y = 1$. Finally, we note that the inner product of x with y vanishes, because

$$x'y \overset{(EC.4)}{=} x'\left(I_N - xx'\right)u = \left(x' - (x'x)x'\right)u = 0'u = 0. \tag{EC.5}$$

Therefore, the correlation between $\tilde{v}$ and w, which we will denote as $Corr(\tilde{v}, w)$, can be calculated as

$$Corr(\tilde{v}, w) = \frac{\tilde{v}'\left(w - \bar{w}\right)}{\sqrt{\left(\tilde{v}'\tilde{v}\right)\left(\left(w-\bar{w}\right)'\left(w-\bar{w}\right)\right)}} \qquad \text{[Standard formula, } \tilde{v} \text{ has mean zero]}$$

$$
\begin{aligned}
&= \frac{\tilde{v}'x}{\sqrt{\tilde{v}'\tilde{v}x'x}} && \text{[By definition of } x \text{ in (EC.3)]}\\
&= \frac{\tilde{v}'x}{\sqrt{\tilde{v}'\tilde{v}}} && \text{[Since } x'x = 1\text{]}\\
&= \frac{rx'x + (1-r)x'y}{\sqrt{r^2x'x + 2r(1-r)x'y + (1-r)^2y'y}} && \text{[Definition of } \tilde{v} \text{ from (EC.1)]}\\
&= \frac{r}{\sqrt{r^2 + (1-r)^2}} && \text{[Since } x'x = y'y = 1 \text{ and } x'y = 0 \text{ from (EC.5)]}\\
&= \rho && \text{[By (EC.2)]}
\end{aligned}
$$

■

EC.1.2. Examples of Knapsack Instances

In this section, we present two examples of knapsack instances (these were actual instances from the 100,000 that we generated) to illustrate how the instance-level complexity measures were computed. For each instance, we tabulate their value and weight vectors, the computed values of the six instance-level complexity measures, as well as some intermediate quantities in Table EC.1

Table EC.1 Two examples of knapsack instances used in our experiment.

	Knapsack Instance (i)				Knapsack Instance (ii)			
	Value	Weight	V/W[a]	InOpt[b]	Value	Weight	V/W[a]	InOpt[b]
Item 1	11	21	0.52	N	7	10	0.70	N
Item 2	10	18	0.56	N	30	37	0.81	N
Item 3	15	22	0.68	Y	21	25	0.84	N
Item 4	15	19	0.79	Y	10	12	0.83	N
Item 5	20	28	0.71	Y	24	30	0.80	N
Item 6	8	11	0.73	N	30	34	0.88	Y
Item 7	17	29	0.59	N	19	21	0.90	Y
Item 8	22	37	0.59	N	15	18	0.83	N
Item 9	11	14	0.79	Y	25	29	0.86	Y
Item 10	23	33	0.70	N	14	16	0.88	Y
Item 11	11	17	0.65	Y	16	20	0.80	N
Item 12	15	31	0.48	N	13	17	0.76	N
Correlation between value and weight vectors		0.897				0.991		
Sahni complexity measure		5				0		
Optimal number of items in knapsack		5				4		
Minimum gap from Sahni heuristics		0.0417				0.0000		
Maximum gap from Sahni heuristics		0.2500				0.1023		
Gap from greedy value heuristic		0.0972				0.0341		

Notes: [a] V/W: value of item divided by its weight. [b] Indicates whether this item is in an optimal knapsack (Y: Yes, N: No)

The two instances were chosen to illustrate variations in their instance-level complexity measures. Instance (i) had a Pearson correlation on the low end of our range, whereas instance (ii) had a correlation on the higher end of our range. The scatterplots of their value and weight vectors are

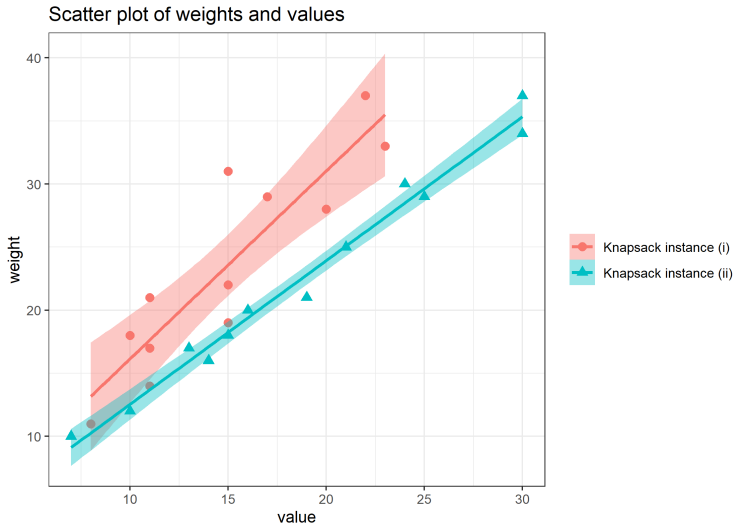


Figure EC.1 Scatterplot of weight and value vectors for the two example instances.

illustrated in Figure EC.1, from which we can clearly observe that the value and weight vectors of instance (ii) are much more tightly clustered around a linear trend compared to instance (i).

For instance (ii), we observe that optimality is attained by a greedy heuristic that accepts items by decreasing order of value density (value divided by weight) until the knapsack capacity is reached. That is, a heuristic that takes Items 7, 6, 10, and 9, in that order, turns out to already result in an optimal knapsack. This heuristic is exactly the Sahni-0 heuristic, and thus the Sahni complexity measure for instance (ii) is 0. In contrast, for instance (i), a brute force search must be done for knapsacks with at most 5 items before optimality is reached, and so it has a Sahni complexity measure of 5.

The fact that instance (ii) has a Sahni complexity measure of 0 also explains why the minimum gap from the Sahni heuristics is 0 for this instance. This is because we have found one of the Sahni heuristics (i.e., Sahni-0) which has attained optimality and thus has a zero optimality gap. It might come as a surprise that the maximum gap from the Sahni heuristics is non-zero for this instance. However, this can be resolved by observing what happens with the Sahni-2 heuristic. In this case, the algorithm first selects an optimal knapsack with at most two items (Items 2 and 6) before filling the remaining knapsack by value density. However, by committing to Item 2 (which is not in the actual optimal solution), the Sahni-2 heuristic turns out to have a non-zero optimality gap. As such, the maximum gap across the Sahni-k heuristics, for $k = 0, 1, 2, 3, 4$ was also strictly greater than zero. The calculations of the minimal and maximal gaps from the Sahni-k heuristics for instance (i) follow these same procedures, which we refrain from repeating.

Finally, the greedy value heuristic works by naïvely filling items just by looking at their values until the knapsack is full, and then it stops. For instance (i), Items 10, 8, and 5 have the largest values and are taken in that order by this heuristic; after doing so, the total knapsack weight is still feasible at 98 and no other items can be added. The heuristic results in a knapsack value of 65. Since the optimal knapsack's value is 72, so the optimality gap from this heuristic is computed as (72 - 65)/72 = 0.0972. The calculations for instance (ii) follow a similar logic.

EC.2. Supplemental Figures

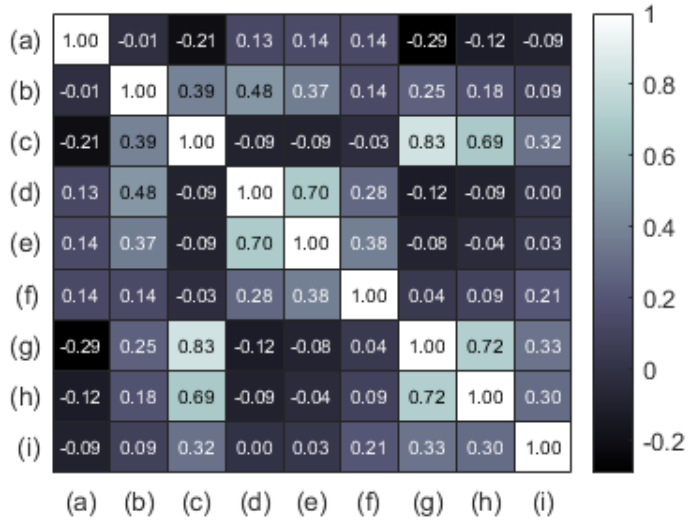


Figure EC.2 Correlation matrix of regressors in feature group 1 with raw Sahni optimality gaps. Legend: (a) Pearson correlation between w and v, (b) Sahni complexity, (c) number of items in optimal knapsack, (d) optimality gap from Sahni-0 heuristic, (e) optimality gap from Sahni-1 heuristic, (f) optimality gap from Sahni-2 heuristic, (g) optimality gap from Sahni-3 heuristic, (h) optimality gap from Sahni-4 heuristic, and (i) optimality gap from greedy value heuristic.

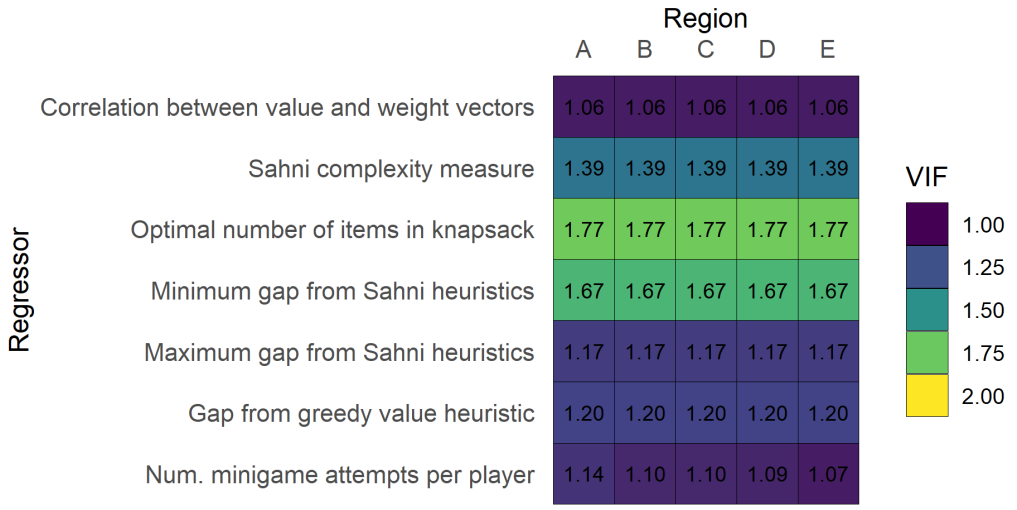


Figure EC.3 Heatmap of VIFs of primary regressors in the base-case model across regions.

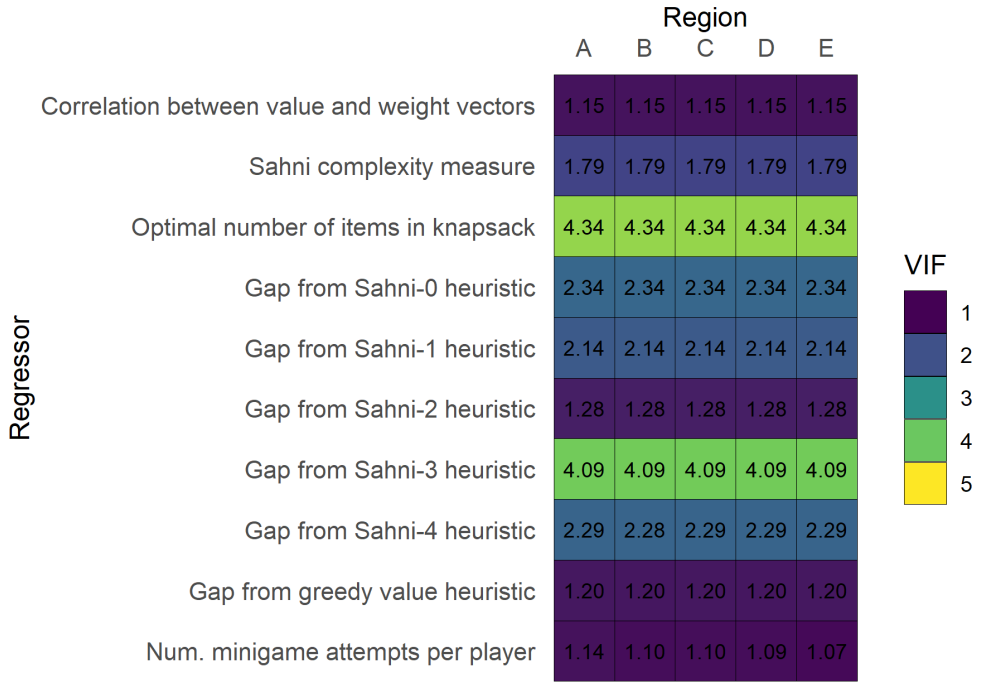


Figure EC.4 Heatmap of VIFs across all five regions, from a modified base-case model where feature group 1 comprised raw Sahni optimality gaps.

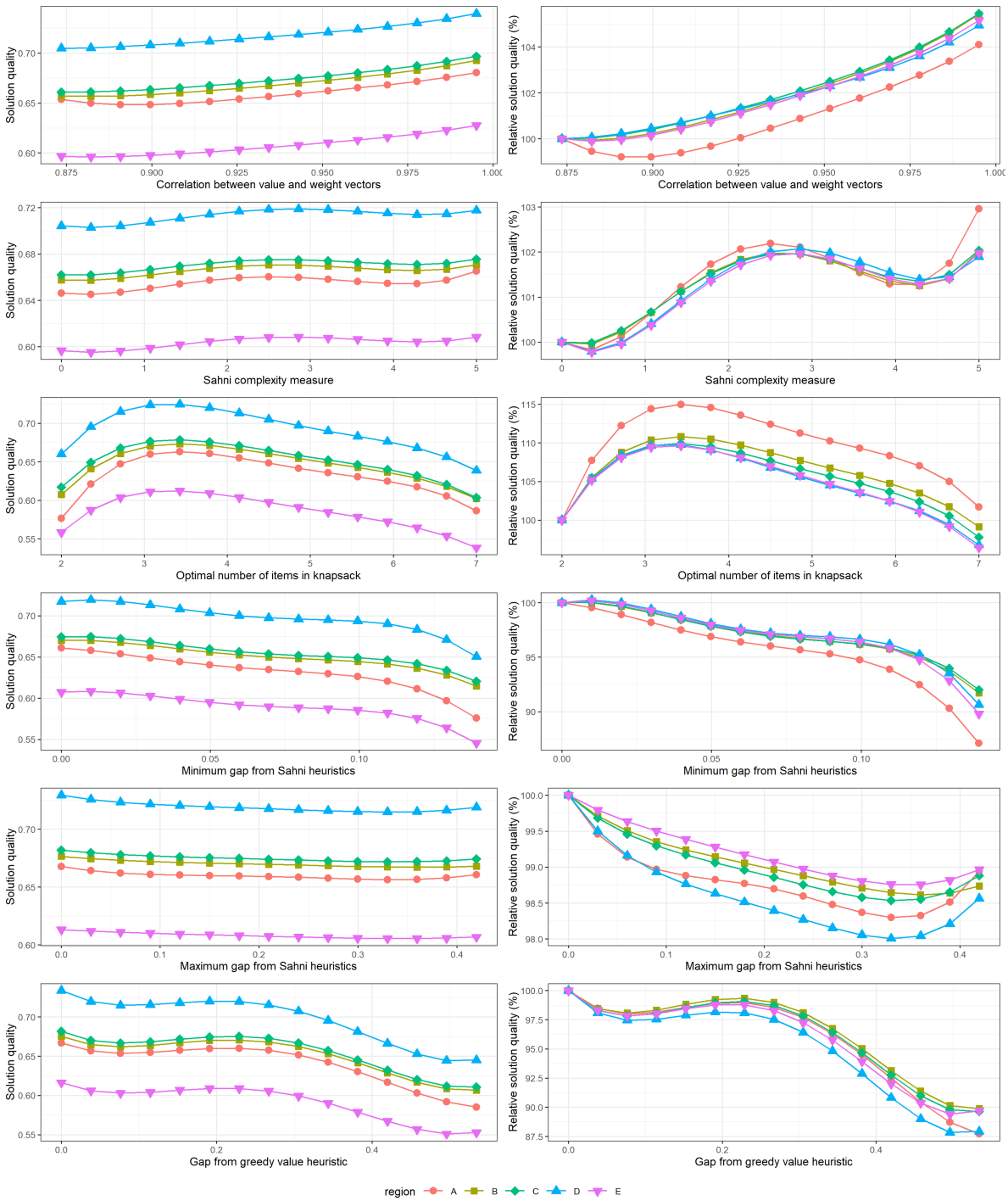


Figure EC.5 **Left: Estimated solution qualities for the polynomial regression model with $d = 4$, as individual instance-level complexity measures vary from minimal to maximal values in the data. Right: Estimated solution qualities expressed in relative terms, as a percentage of the solution quality obtained from the smallest instance-level complexity measure for each measure. Note: For region E, we were unable to fit the full polynomial regression model into memory, and the results above were estimated using a 50% random subset of players.**

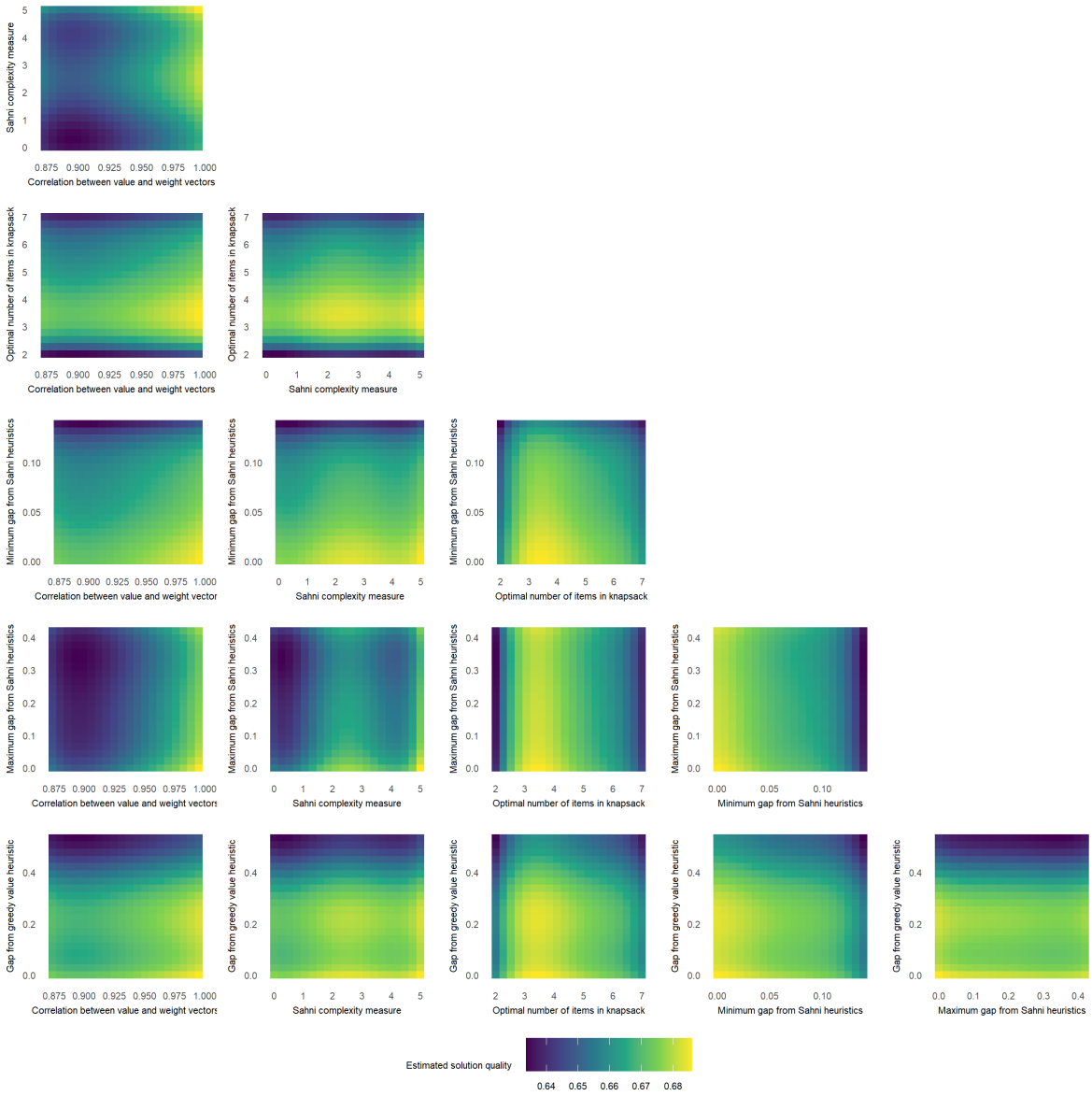


Figure EC.6 Heatmap of estimated solution qualities with two-way variations in instance-level complexity measures. Estimates were obtained from Region A using polynomial regression with degree $d = 4$.

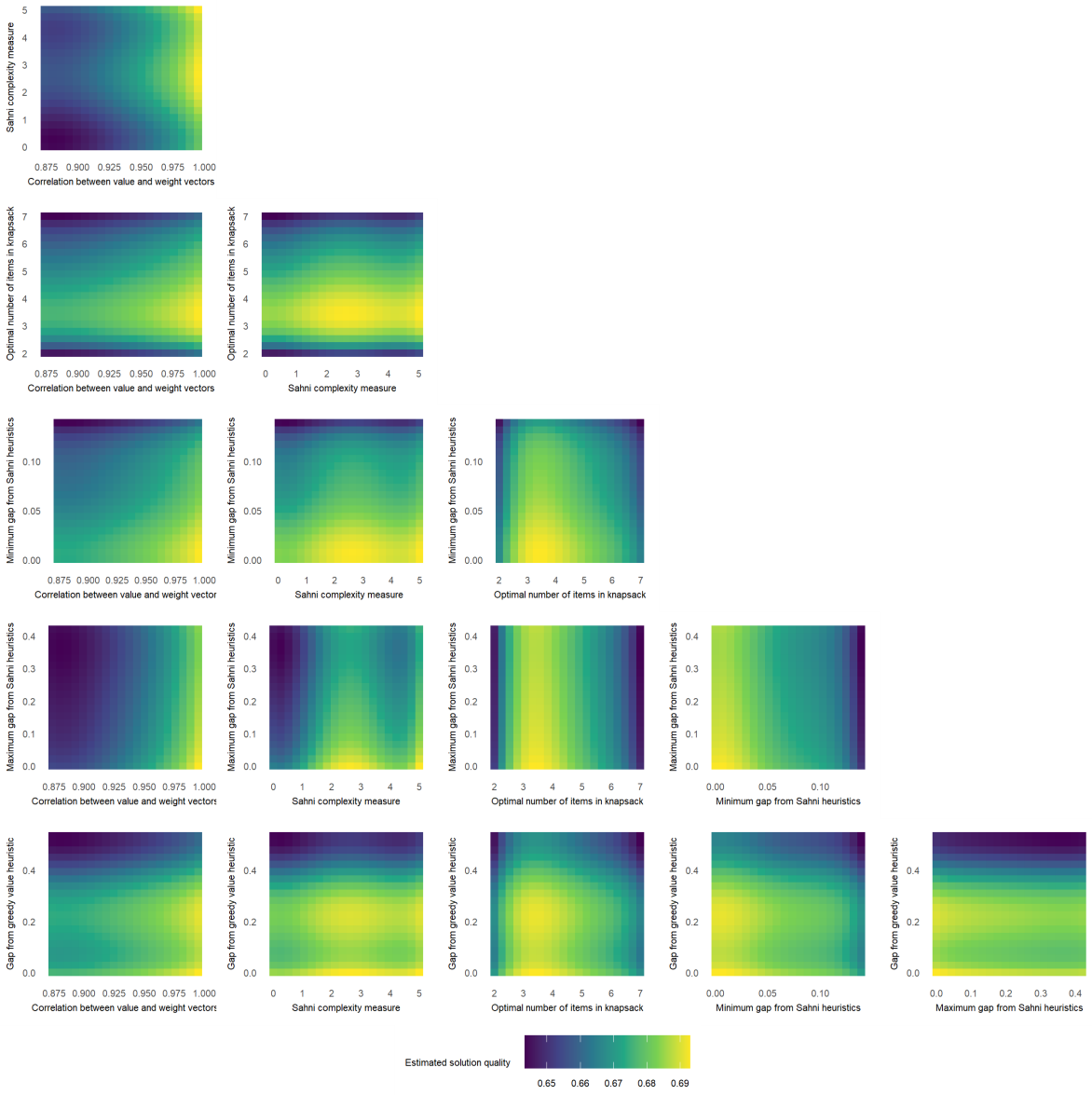


Figure EC.7 Heatmap of estimated solution qualities with two-way variations in instance-level complexity measures. Estimates were obtained from Region B using polynomial regression with degree $d = 4$.

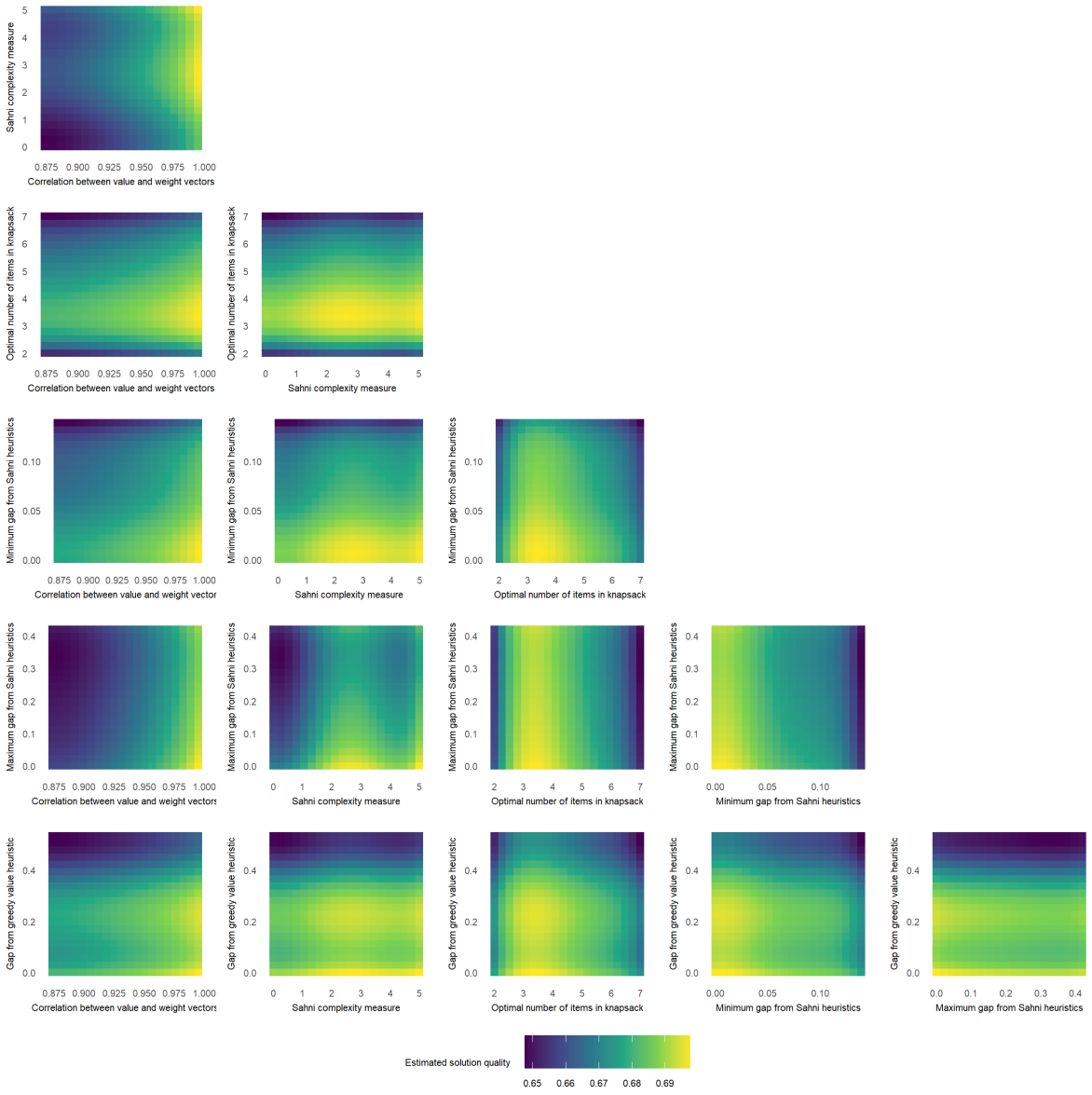


Figure EC.8 Heatmap of estimated solution qualities with two-way variations in instance-level complexity measures. Estimates were obtained from Region C using polynomial regression with degree $d = 4$.

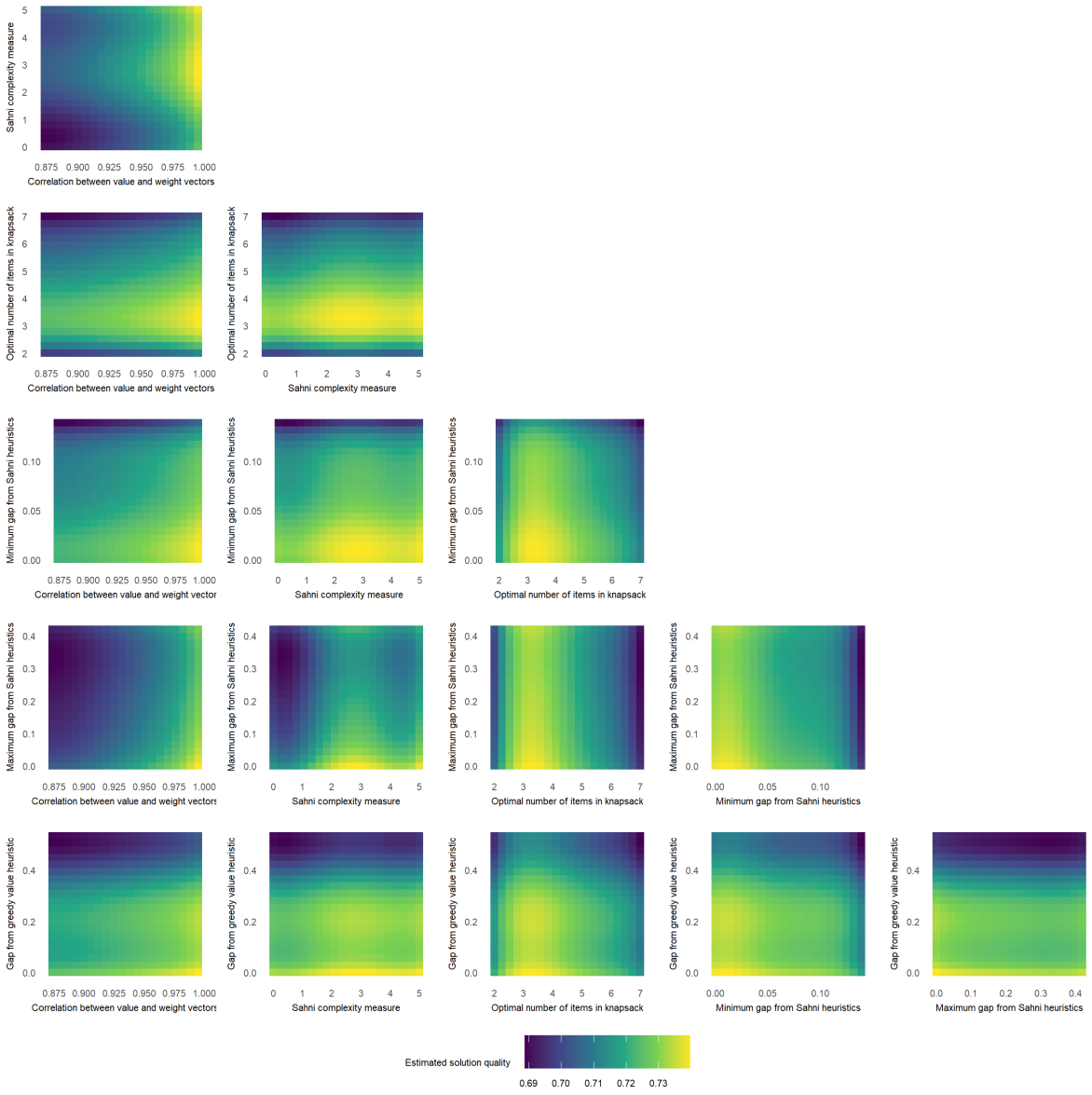


Figure EC.9 Heatmap of estimated solution qualities with two-way variations in instance-level complexity measures. Estimates were obtained from Region D using polynomial regression with degree $d = 4$.

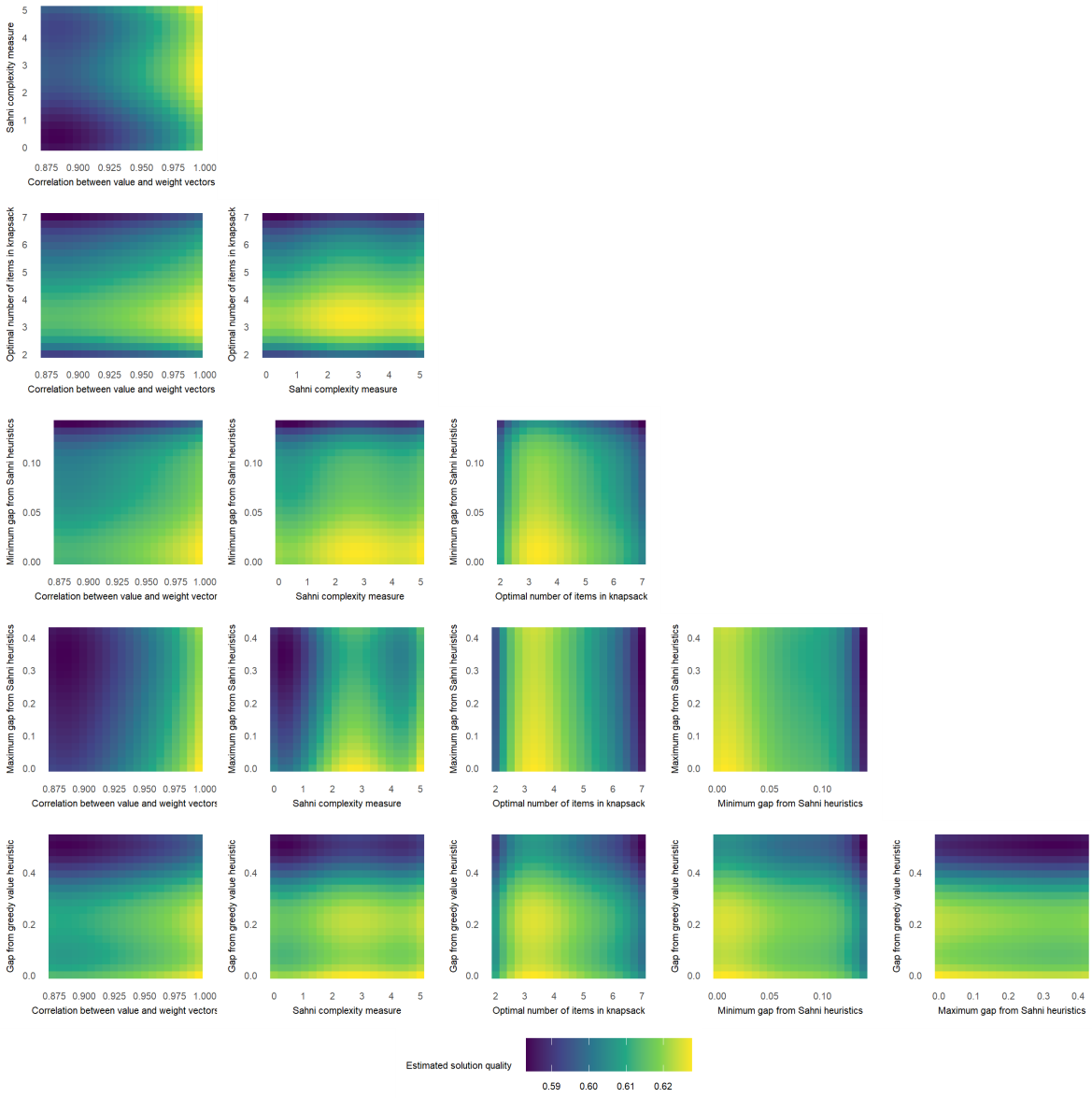


Figure EC.10 Heatmap of estimated solution qualities with two-way variations in instance-level complexity measures. Estimates were obtained from Region E using polynomial regression with degree $d = 4$.

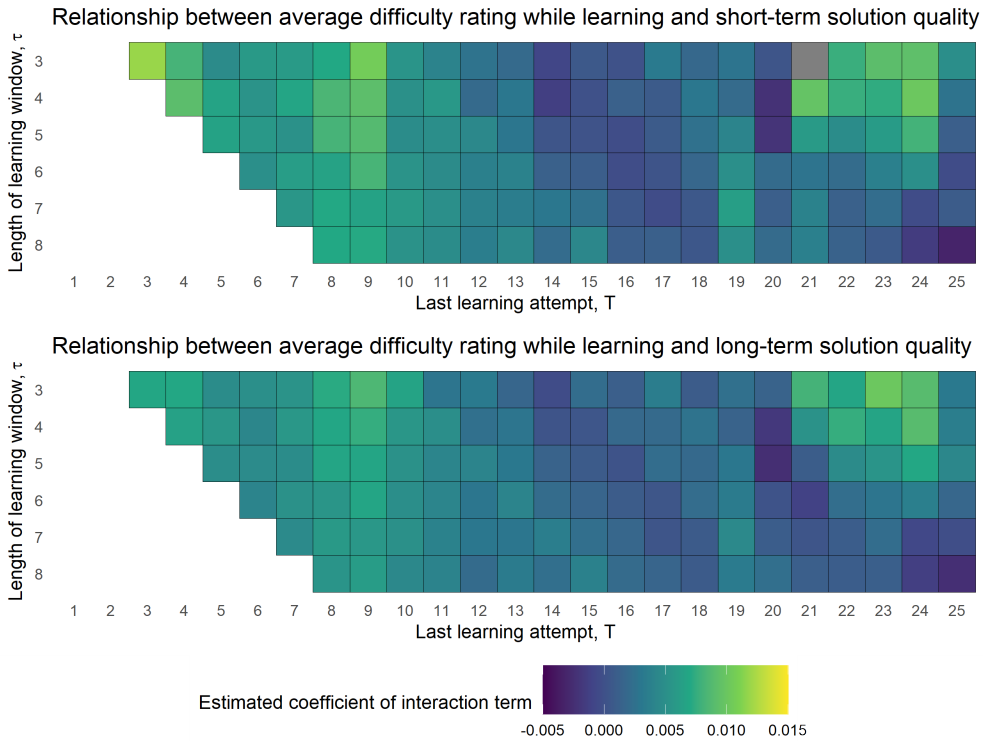


Figure EC.11 Point estimate of regression coefficients for outcomes of short-term learning (top) and long-term learning (bottom).

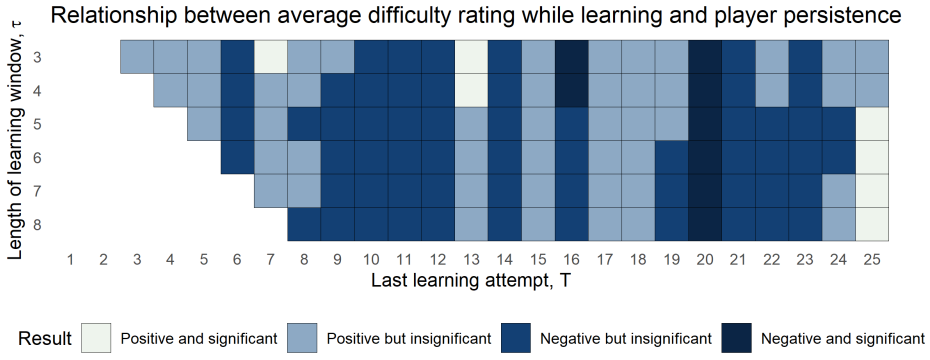


Figure EC.12 Results of longitudinal dropout analysis

EC.3. Supplemental Tables

Table EC.2: Description of all features used in regression models.

Feature	Description
Instance-level complexity measures	
Correlation between value and weight vectors	Pearson correlation between value vector v and weight vector w.
Sahni complexity measure	The smallest value of k such that the Sahni-k heuristic solves the knapsack problem to optimality.
Num. items in optimal knapsack	The number of items in an optimal knapsack.
Min. Gap from Sahni heuristics	Minimum of optimality gaps from Sahni-k heuristics, $k = 0, 1, 2, 3, 4$.
Max. Gap from Sahni heuristics	Maximum of optimality gaps from Sahni-k heuristics, $k = 0, 1, 2, 3, 4$.
Gap from greedy value heuristic	Difference between optimal knapsack value and a heuristic that fills the knapsack in descending order of value until no more items can be put in, measured as a % of knapsack's optimal value.
Measure of player learning	
Num. minigame attempts per player	The number of previous minigame attempts made by this player during the event.
Player statistics in primary game	
log(num. active days in last 30d)	Log-transformed value of the number of days that the player was active on the mobile app over the past 30 days.
log(avg. online time in last 30d)	Log-transformed value of the daily average amount of time (in seconds) that the player was active on the mobile app over the past 30 days.
log(total active days since joining platform)	Log-transformed value of the number of days that the player was active on the mobile app since the player joined the platform.
log (total online time since joining platform)	Log-transformed value of the total amount of time (in seconds) that the player was active on the mobile app since joining the platform
Is part of clan	Binary variable. The primary game had a system in which players could self-organize into social groups ('clans'). Membership in a clan was an indirect measure of player engagement in the app. This variable $= 1$ if the player was a member of a clan at the time of data extraction, and $= 0$ otherwise.
Is ranked	Binary variable. This variable $= 1$ if the player had a rank at the time of data of data extraction, and $= 0$ otherwise.
Is maximum rank	Binary variable. This variable $= 1$ if the player attained the maximum rank at the time of data of data extraction, and $= 0$ otherwise.
Rank	Player rank (revealed to player).

Continued on next page

Table EC.2 – *Continued from previous page*

Feature	Description
Has record of eliminated opponents	Binary variable. This variable = 1 if the player had a record of the number of eliminated opponents in the current epoch.
log(1 + num. opponents eliminated)	Log-transformed value of the average number of opponents defeated per match in the current epoch.
log(1 + avg. damage)	Log-transformed value of the average damage inflicted on opponent players per match in the current epoch.
Has win rate recorded	Binary variable. This variable = 1 if the player had a record of their win rate during the current epoch.
Win rate	The proportion of games won to the number of games played during the current epoch.
Has MMR recorded	Binary variable. This variable = 1 if the player had a record of the MMR score at the point of data extraction.
log (MMR)	Log-transformed value of the internally-derived matchmaking score at the point of data extraction.
Played at least one match in last 30d	Binary variable, = 1 if player played at least one match in the primary game within the past 30 days, and = 0 otherwise.
log (1 + num. matches in last 30d)	Log-transformed value of the number of matches of the primary game played by the player over the past 30 days.
Has used premium currency	Binary variable. This variable = 1 if the player had ever used funds to purchase in-game premium currency, and = 0 otherwise.
log(1 + premium currency)	Log-transformed value of (scaled) amount of total premium currency purchased by player since joining the platform.

Table EC.3 Regression results for region A using different feature groups. Means and standard errors (in parentheses) of regression coefficients.

Model	(a)	(b)	(c)	(d)
Intercept	0.426***	0.408***	0.307***	0.671***
	(0.007)	(0.007)	(0.007)	(0.026)
Instance-level complexity measures:				
Correlation between value and weight vectors	0.328***	0.337***	0.338***	0.338***
	(0.007)	(0.007)	(0.007)	(0.007)
Sahni complexity measure	0.002***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.022***	-0.015***	-0.015***	-0.015***
	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics		-0.505***	-0.507***	-0.504***
		(0.019)	(0.019)	(0.019)
Maximum gap from Sahni heuristics		-0.030***	-0.030***	-0.029***
		(0.003)	(0.002)	(0.002)
Gap from greedy value heuristic		-0.050***	-0.051***	-0.051***
		(0.002)	(0.002)	(0.002)
Measure of player learning:				
Num. minigame attempts per player			0.008***	0.010***
			(0.000)	(0.000)
Control for primary game player stats?	No	No	No	Yes
R^2	0.0057	0.0066	0.0647	0.0887
Num. observations	1,719,964	1,719,964	1,719,964	1,719,964
Num. covariates	4	7	8	27
F-statistic	3,284.0	1,894.6	16,996.8	6,441.3

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.4 Regression results for region B using different feature groups. Means and standard errors (in parentheses) of regression coefficients.

Model	(a)	(b)	(c)	(d)
Intercept	0.409***	0.391***	0.282***	0.796***
	(0.002)	(0.002)	(0.002)	(0.008)
Instance-level complexity measures:				
Correlation between value and weight vectors	0.341***	0.350***	0.350***	0.350***
	(0.002)	(0.002)	(0.002)	(0.002)
Sahni complexity measure	0.003***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.021***	-0.015***	-0.015***	-0.015***
	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics		-0.502***	-0.503***	-0.503***
		(0.005)	(0.005)	(0.005)
Maximum gap from Sahni heuristics		-0.026***	-0.026***	-0.026***
		(0.001)	(0.001)	(0.001)
Gap from greedy value heuristic		-0.041***	-0.042***	-0.041***
		(0.001)	(0.001)	(0.001)
Measure of player learning:				
Num. minigame attempts per player			0.009***	0.010***
			(0.000)	(0.000)
Control for primary game player stats?	No	No	No	Yes
R^2	0.0054	0.0062	0.0702	0.0850
Num. observations	27,406,396	27,406,396	27,406,396	27,406,396
Num. covariates	4	7	8	27
F-statistic	49,962.1	28,297.4	295,693.8	97,969.2

Standard errors are cluster robust standard errors at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.5 Regression results for region C using different feature groups. Means and standard errors (in parentheses) of regression coefficients.

Model	(a)	(b)	(c)	(d)
Intercept	0.427***	0.409***	0.298***	0.826***
	(0.003)	(0.003)	(0.003)	(0.012)
Instance-level complexity measures:				
Correlation between value and weight vectors	0.334***	0.344***	0.344***	0.343***
	(0.003)	(0.003)	(0.003)	(0.003)
Sahni complexity measure	0.003***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.022***	-0.016***	-0.016***	-0.016***
	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics		-0.499***	-0.499***	-0.500***
		(0.007)	(0.007)	(0.007)
Maximum gap from Sahni heuristics		-0.030***	-0.030***	-0.030***
		(0.001)	(0.001)	(0.001)
Gap from greedy value heuristic		-0.045***	-0.046***	-0.046***
		(0.001)	(0.001)	(0.001)
Measure of player learning:				
Num. minigame attempts per player			0.009***	0.010***
			(0.000)	(0.000)
Control for primary game player stats?	No	No	No	Yes
R^2	0.0059	0.0067	0.0757	0.0920
Num. observations	13,020,861	13,020,861	13,020,861	13,020,861
Num. covariates	4	7	8	27
F-statistic	25,899.3	14,708.6	152,293.1	50,756.7

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.6 Regression results for region D using different feature groups. Means and standard errors (in parentheses) of regression coefficients.

Model	(a)	(b)	(c)	(d)
Intercept	0.509***	0.496***	0.393***	0.493***
	(0.002)	(0.002)	(0.002)	(0.008)
Instance-level complexity measures:				
Correlation between value and weight vectors	0.319***	0.329***	0.329***	0.329***
	(0.002)	(0.002)	(0.002)	(0.002)
Sahni complexity measure	0.003***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.029***	-0.022***	-0.022***	-0.022***
	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics		-0.496***	-0.496***	-0.497***
		(0.005)	(0.005)	(0.005)
Maximum gap from Sahni heuristics		-0.040***	-0.040***	-0.040***
		(0.001)	(0.001)	(0.001)
Gap from greedy value heuristic		-0.083***	-0.083***	-0.083***
		(0.001)	(0.001)	(0.001)
Measure of player learning:				
Num. minigame attempts per player			0.008***	0.008***
			(0.000)	(0.000)
Control for primary game player stats?	No	No	No	Yes
R^2	0.0107	0.0127	0.0657	0.0808
Num. observations	18,260,964	18,260,964	18,260,964	18,260,964
Num. covariates	4	7	8	27
F-statistic	66,119.7	39,023.2	183,472.8	61,746.6

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.7 Regression results for region E using different feature groups. Means and standard errors (in parentheses) of regression coefficients.

Model	(a)	(b)	(c)	(d)
Intercept	0.390***	0.377***	0.274***	0.613***
	(0.001)	(0.001)	(0.001)	(0.004)
Instance-level complexity measures:				
Correlation between value and weight vectors	0.304***	0.311***	0.311***	0.311***
	(0.001)	(0.001)	(0.001)	(0.001)
Sahni complexity measure	0.003***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.023***	-0.017***	-0.017***	-0.017***
	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics		-0.441***	-0.441***	-0.441***
		(0.003)	(0.003)	(0.003)
Maximum gap from Sahni heuristics		-0.023***	-0.023***	-0.023***
		(0.000)	(0.000)	(0.000)
Gap from greedy value heuristic		-0.055***	-0.055***	-0.055***
		(0.000)	(0.000)	(0.000)
Measure of player learning:				
Num. minigame attempts per player			0.011***	0.012***
			(0.000)	(0.000)
Control for primary game player stats?	No	No	No	Yes
R^2	0.0059	0.0067	0.0414	0.0555
Num. observations	89,876,010	89,876,010	89,876,010	89,876,010
Num. covariates	4	7	8	27
F-statistic	177,397.3	101,003.2	553,944.7	202,976.8

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.8 Regression results on a dataset that includes player attempts ending in timeouts. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	0.663***	0.789***	0.813***	0.494***	0.620***
	(0.026)	(0.008)	(0.011)	(0.008)	(0.004)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.336***	0.347***	0.340***	0.327***	0.310***
	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.016***	-0.015***	-0.016***	-0.022***	-0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.501***	-0.505***	-0.501***	-0.499***	-0.443***
	(0.018)	(0.005)	(0.007)	(0.005)	(0.003)
Maximum gap from Sahni heuristics	-0.029***	-0.026***	-0.030***	-0.040***	-0.023***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.051***	-0.041***	-0.045***	-0.083***	-0.054***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.009***	0.009***	0.009***	0.008***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0857	0.0818	0.0883	0.0786	0.0535
Num. observations	1,789,101	28,364,662	13,559,904	18,955,174	93,251,066
Num. covariates	27	27	27	27	27
F-statistic	6,449.2	97,200.4	50,528.3	62,177.0	202,738.8

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.9 Regression results on a dataset where optimal knapsack size is defined as optimal knapsack with the smallest number of items.

Region	A	B	C	D	E
Intercept	0.686***	0.810***	0.840***	0.507***	0.625***
	(0.026)	(0.008)	(0.012)	(0.008)	(0.004)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.327***	0.340***	0.333***	0.318***	0.302***
	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.017***	-0.017***	-0.018***	-0.023***	-0.018***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.432***	-0.436***	-0.434***	-0.427***	-0.382***
	(0.019)	(0.005)	(0.007)	(0.005)	(0.003)
Maximum gap from Sahni heuristics	-0.028***	-0.026***	-0.029***	-0.039***	-0.022***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.049***	-0.039***	-0.044***	-0.081***	-0.054***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.010***	0.010***	0.010***	0.008***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0890	0.0853	0.0922	0.0811	0.0556
Num. observations	1,719,964	27,406,396	13,020,861	18,260,964	89,876,010
Num. covariates	27	27	27	27	27
F-statistic	6,461.1	98,243.4	50,890.4	62,000.9	203,693.3

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.10 Regression results on a dataset where optimal knapsack size is defined as optimal knapsack with the largest number of items.

Region	A	B	C	D	E
Intercept	0.643***	0.770***	0.799***	0.460***	0.587***
	(0.026)	(0.008)	(0.012)	(0.008)	(0.004)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.355***	0.366***	0.360***	0.349***	0.327***
	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.003***	0.003***	0.003***	0.004***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.012***	-0.011***	-0.012***	-0.017***	-0.013***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.623***	-0.615***	-0.617***	-0.640***	-0.556***
	(0.018)	(0.005)	(0.007)	(0.005)	(0.003)
Maximum gap from Sahni heuristics	-0.028***	-0.026***	-0.029***	-0.040***	-0.023***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.057***	-0.047***	-0.051***	-0.089***	-0.060***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.010***	0.010***	0.010***	0.008***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0884	0.0847	0.0917	0.0802	0.0551
Num. observations	1,719,964	27,406,396	13,020,861	18,260,964	89,876,010
Num. covariates	27	27	27	27	27
F-statistic	6,412.8	97,564.5	50,540.9	61,207.3	201,629.5

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.11 Regression results on a dataset that excludes the first 6 minigame attempts of each player. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	0.791***	0.891***	0.926***	0.537***	0.660***
	(0.033)	(0.012)	(0.015)	(0.009)	(0.006)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.344***	0.358***	0.352***	0.330***	0.316***
	(0.008)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.017***	-0.016***	-0.018***	-0.023***	-0.019***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.529***	-0.523***	-0.520***	-0.508***	-0.454***
	(0.021)	(0.005)	(0.008)	(0.006)	(0.003)
Maximum gap from Sahni heuristics	-0.032***	-0.032***	-0.035***	-0.044***	-0.028***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.057***	-0.048***	-0.053***	-0.087***	-0.063***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.007***	0.007***	0.007***	0.006***	0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0622	0.0563	0.0622	0.0566	0.0362
Num. observations	1,221,001	20,091,117	9,555,896	14,490,278	58,880,441
Num. covariates	27	27	27	27	27
F-statistic	3,113.7	46,062.9	24,372.7	33,421.5	85,082.0

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.12 Regression results on a dataset that excludes the first 9 minigame attempts of each player. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	0.811***	0.916***	0.944***	0.561***	0.680***
	(0.039)	(0.015)	(0.018)	(0.011)	(0.008)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.343***	0.359***	0.359***	0.330***	0.321***
	(0.009)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.018***	-0.017***	-0.019***	-0.024***	-0.020***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.533***	-0.537***	-0.538***	-0.518***	-0.463***
	(0.024)	(0.006)	(0.009)	(0.006)	(0.004)
Maximum gap from Sahni heuristics	-0.035***	-0.034***	-0.037***	-0.046***	-0.031***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Gap from greedy value heuristic	-0.059***	-0.051***	-0.056***	-0.088***	-0.066***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.006***	0.005***	0.006***	0.005***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0487	0.0431	0.0485	0.0442	0.0297
Num. observations	886,951	14,804,482	7,060,166	11,319,364	36,527,882
Num. covariates	27	27	27	27	27
F-statistic	1,746.9	25,647.9	13,828.4	20,144.9	43,046.3

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.13 Regression results on a dataset that excludes the first 12 minigame attempts of each player. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	0.798***	0.921***	0.947***	0.574***	0.704***
	(0.047)	(0.018)	(0.022)	(0.014)	(0.011)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.344***	0.362***	0.359***	0.332***	0.323***
	(0.010)	(0.003)	(0.004)	(0.003)	(0.002)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.019***	-0.018***	-0.019***	-0.025***	-0.021***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.530***	-0.536***	-0.549***	-0.525***	-0.474***
	(0.028)	(0.007)	(0.010)	(0.007)	(0.005)
Maximum gap from Sahni heuristics	-0.036***	-0.036***	-0.039***	-0.048***	-0.033***
	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
Gap from greedy value heuristic	-0.059***	-0.052***	-0.059***	-0.088***	-0.066***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Measure of player learning:					
Num. minigame attempts per player	0.005***	0.005***	0.005***	0.004***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0419	0.0367	0.0418	0.0377	0.0276
Num. observations	643,366	10,798,844	5,169,121	8,601,916	20,835,236
Num. covariates	27	27	27	27	27
F-statistic	1,082.5	15,810.3	8,681.8	12,975.9	22,735.5

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.14 Regression results on a dataset that only includes players who were active on the platform for 5 or more days within the last 30 days of the start of the knapsack event. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	0.685***	0.803***	0.831***	0.491***	0.603***
	(0.027)	(0.009)	(0.012)	(0.008)	(0.004)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.338***	0.350***	0.343***	0.329***	0.311***
	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.015***	-0.015***	-0.016***	-0.022***	-0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.502***	-0.503***	-0.500***	-0.496***	-0.441***
	(0.019)	(0.005)	(0.007)	(0.005)	(0.003)
Maximum gap from Sahni heuristics	-0.029***	-0.026***	-0.030***	-0.040***	-0.023***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.051***	-0.041***	-0.046***	-0.083***	-0.055***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.010***	0.010***	0.010***	0.008***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0895	0.0852	0.0922	0.0811	0.0556
Num. observations	1,704,586	27,351,929	12,993,544	18,175,446	89,661,294
Num. covariates	27	27	27	27	27
F-statistic	6,441.4	97,936.3	50,736.8	61,714.3	202,959.2

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.15 Regression results on a dataset that only includes players who were active on the platform for 10 or more days within the last 30 days of the start of the knapsack event. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	0.688***	0.796***	0.821***	0.467***	0.570***
	(0.029)	(0.011)	(0.015)	(0.009)	(0.005)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.336***	0.350***	0.344***	0.329***	0.311***
	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.015***	-0.015***	-0.016***	-0.022***	-0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.506***	-0.503***	-0.501***	-0.494***	-0.441***
	(0.019)	(0.005)	(0.007)	(0.005)	(0.003)
Maximum gap from Sahni heuristics	-0.028***	-0.027***	-0.030***	-0.040***	-0.023***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.050***	-0.041***	-0.046***	-0.083***	-0.056***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.010***	0.010***	0.010***	0.008***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0910	0.0857	0.0928	0.0820	0.0559
Num. observations	1,635,074	27,017,413	12,821,856	17,687,367	88,214,086
Num. covariates	27	27	27	27	27
F-statistic	6,293.8	97,361.3	50,445.4	60,798.4	200,853.5

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.16 Regression results on a dataset that only includes players who were active on the platform for 15 or more days within the last 30 days of the start of the knapsack event. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	0.701***	0.788***	0.799***	0.468***	0.544***
	(0.033)	(0.015)	(0.019)	(0.013)	(0.008)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.337***	0.350***	0.344***	0.328***	0.311***
	(0.008)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.016***	-0.015***	-0.016***	-0.022***	-0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.506***	-0.504***	-0.502***	-0.494***	-0.441***
	(0.020)	(0.005)	(0.007)	(0.005)	(0.003)
Maximum gap from Sahni heuristics	-0.028***	-0.027***	-0.030***	-0.040***	-0.023***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.049***	-0.041***	-0.046***	-0.084***	-0.056***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.010***	0.009***	0.010***	0.008***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0925	0.0862	0.0935	0.0827	0.0561
Num. observations	1,522,705	26,270,974	12,441,412	16,946,195	85,552,310
Num. covariates	27	27	27	27	27
F-statistic	5,966.9	95,292.1	49,356.9	58,754.1	195,603.2

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.17 Regression results on a dataset that only includes players who were active on the platform for a total of 30 days or more since joining the platform. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	0.724***	0.838***	0.872***	0.531***	0.636***
	(0.041)	(0.023)	(0.029)	(0.017)	(0.012)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.339***	0.349***	0.344***	0.328***	0.311***
	(0.008)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.016***	-0.015***	-0.016***	-0.022***	-0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.506***	-0.504***	-0.505***	-0.495***	-0.441***
	(0.021)	(0.005)	(0.007)	(0.005)	(0.003)
Maximum gap from Sahni heuristics	-0.028***	-0.027***	-0.030***	-0.041***	-0.023***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.049***	-0.042***	-0.046***	-0.084***	-0.056***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.010***	0.009***	0.010***	0.008***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0906	0.0856	0.0929	0.0825	0.0553
Num. observations	1,428,986	26,182,050	12,396,028	16,792,625	86,114,105
Num. covariates	27	27	27	27	27
F-statistic	5,474.3	94,209.6	48,823.0	58,059.0	193,831.7

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.18 Regression results on a dataset that only includes players who were active on the platform for a total of 60 days or more since joining the platform. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	0.750***	0.881***	0.914***	0.609***	0.640***
	(0.063)	(0.046)	(0.041)	(0.032)	(0.037)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.342***	0.350***	0.343***	0.329***	0.311***
	(0.009)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.016***	-0.015***	-0.016***	-0.022***	-0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.485***	-0.505***	-0.506***	-0.493***	-0.442***
	(0.023)	(0.005)	(0.007)	(0.006)	(0.003)
Maximum gap from Sahni heuristics	-0.030***	-0.027***	-0.031***	-0.041***	-0.024***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.050***	-0.042***	-0.046***	-0.085***	-0.057***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.009***	0.009***	0.010***	0.008***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0879	0.0849	0.0922	0.0820	0.0546
Num. observations	1,185,784	24,529,531	11,527,744	15,094,542	81,751,698
Num. covariates	27	27	27	27	27
F-statistic	4,392.9	87,545.8	45,026.8	51,856.0	181,557.6

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.19 Regression results on a dataset that only includes players who were active on the platform for a total of 90 days or more since joining the platform. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	1.013***	0.925***	0.946***	0.722***	0.633***
	(0.037)	(0.052)	(0.088)	(0.079)	(0.075)
Instance-level complexity measures:					
Correlation between value and weight vectors	0.342***	0.351***	0.344***	0.329***	0.312***
	(0.009)	(0.002)	(0.003)	(0.002)	(0.001)
Sahni complexity measure	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimal number of items in knapsack	-0.016***	-0.015***	-0.016***	-0.022***	-0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Minimum gap from Sahni heuristics	-0.494***	-0.506***	-0.506***	-0.496***	-0.444***
	(0.024)	(0.005)	(0.008)	(0.006)	(0.003)
Maximum gap from Sahni heuristics	-0.029***	-0.027***	-0.031***	-0.041***	-0.024***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Gap from greedy value heuristic	-0.049***	-0.042***	-0.046***	-0.085***	-0.058***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Measure of player learning:					
Num. minigame attempts per player	0.009***	0.009***	0.010***	0.008***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for primary game player stats?	Yes	Yes	Yes	Yes	Yes
R^2	0.0854	0.0840	0.0913	0.0811	0.0538
Num. observations	1,017,102	22,850,710	10,594,675	13,430,940	77,094,143
Num. covariates	27	27	27	27	27
F-statistic	3,652.5	80,558.7	40,921.7	45,562.1	168,663.0

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.20: Regression results for model variants investigating player heterogeneity in their learning effectiveness. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Num. active days in last 30d:					
Primary game statistic	0.003	0.026***	0.028***	-0.001	0.020***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Num. minigame attempts per player	0.022***	0.026***	0.025***	0.013***	0.014***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.004***	-0.005***	-0.005***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Avg. online time in last 30d:					
Primary game statistic	-0.027***	-0.022***	-0.022***	-0.018***	-0.031***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Num. minigame attempts per player	0.024***	0.024***	0.023***	0.019***	0.026***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.002***	-0.002***	-0.001***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total active days since joining:					
Primary game statistic	0.032***	0.043***	0.046***	0.032***	0.033***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Num. minigame attempts per player	0.011***	0.011***	0.011***	0.011***	0.015***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.000***	-0.000***	-0.000***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total online time since joining:					
Primary game statistic	-0.019***	-0.029***	-0.033***	-0.012***	-0.013***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Num. minigame attempts per player	0.017***	0.016***	0.016***	0.018***	0.023***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.001***	-0.000***	-0.000***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Is part of a clan:					
Primary game statistic	0.018***	0.016***	0.014***	0.031***	0.023***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Num. minigame attempts per player	0.011***	0.010***	0.010***	0.009***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Player rank:					
Primary game statistic	0.003***	0.003***	0.003***	0.002***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Num. minigame attempts per player	0.011***	0.012***	0.011***	0.011***	0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Avg. num. opponents defeated:					
Primary game statistic	0.011**	0.023***	0.022***	0.014***	0.035***
	(0.003)	(0.001)	(0.002)	(0.001)	(0.001)
Num. minigame attempts per player	0.010***	0.010***	0.010***	0.009***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.000*	-0.000***	-0.000	-0.001***	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Avg. damage to opponents:					
Primary game statistic	0.005**	0.006***	0.005***	0.006***	0.002***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Num. minigame attempts per player	0.011***	0.010***	0.010***	0.010***	0.013***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Interaction term	-0.000***	-0.000***	-0.000	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Win rate:					
Primary game statistic	0.022*	0.051***	0.041***	0.113***	0.092***
	(0.009)	(0.004)	(0.005)	(0.004)	(0.002)
Num. minigame attempts per player	0.010***	0.010***	0.010***	0.009***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.002**	-0.004***	-0.003***	-0.004***	-0.003***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
MMR score:					
Primary game statistic	0.053***	0.044***	0.043***	0.076***	0.043***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Num. minigame attempts per player	0.028***	0.030***	0.029***	0.032***	0.035***
	(0.002)	(0.000)	(0.001)	(0.000)	(0.000)
Interaction term	-0.002***	-0.002***	-0.002***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Num. matches played in last 30d:					
Primary game statistic	-0.000	0.004***	0.003***	-0.001**	0.001***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Num. minigame attempts per player	0.012***	0.012***	0.012***	0.010***	0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Qty. premium currency used:					
Primary game statistic	-13.784***	7.576***	4.415***	26.874***	26.781***
	(2.308)	(0.766)	(1.056)	(1.335)	(0.804)
Num. minigame attempts per player	0.010***	0.010***	0.010***	0.009***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction term	-0.674***	-0.919***	-0.728***	-1.650***	-2.168***
	(0.189)	(0.071)	(0.089)	(0.079)	(0.081)

Standard errors are clustered at the level of individual player accounts.
$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.

Table EC.21 Results of regressing player persistence (maximum number of minigame attempts) against primary game player characteristics. Means and standard errors (in parentheses) of regression coefficients.

Region	A	B	C	D	E
Intercept	-41.863***	-56.956***	-55.072***	-50.295***	-27.366***
	(0.953)	(0.277)	(0.399)	(0.361)	(0.082)
Primary game statistics:					
Num. active days in last 30d	6.531***	7.256***	7.314***	6.882***	2.794***
	(0.061)	(0.022)	(0.031)	(0.028)	(0.008)
Avg. online time in last 30d	2.888***	3.552***	3.441***	4.367***	2.585***
	(0.056)	(0.018)	(0.025)	(0.024)	(0.006)
Total active days since joining	-1.337***	-1.847***	-1.798***	0.055	-0.133***
	(0.079)	(0.025)	(0.036)	(0.030)	(0.008)
Total online time since joining	-1.006***	-0.781***	-0.921***	-1.443***	-0.600***
	(0.069)	(0.022)	(0.031)	(0.028)	(0.007)
Is part of a clan	0.737***	0.600***	0.560***	0.993***	0.477***
	(0.057)	(0.020)	(0.027)	(0.024)	(0.006)
Player rank	-0.009	-0.007**	-0.004	0.173***	0.056***
	(0.009)	(0.003)	(0.004)	(0.003)	(0.001)
Avg. num. opponents defeated	1.129***	1.093***	0.886***	1.822***	1.423***
	(0.098)	(0.036)	(0.051)	(0.049)	(0.011)
Avg. damage to opponents	-0.517***	-0.823***	-0.838***	-0.598***	-0.406***
	(0.038)	(0.015)	(0.022)	(0.021)	(0.005)
Win rate	1.684***	3.618***	4.304***	3.318***	1.870***
	(0.232)	(0.110)	(0.152)	(0.119)	(0.034)
MMR score	3.122***	3.962***	3.969***	2.883***	1.626***
	(0.108)	(0.026)	(0.039)	(0.037)	(0.008)
Num. matches played in last 30d	0.687***	0.861***	0.893***	0.555***	0.386***
	(0.034)	(0.010)	(0.014)	(0.012)	(0.003)
Qty. premium currency used (000s)	-0.682***	-1.243***	-1.102***	-3.224***	-1.784***
	(0.077)	(0.045)	(0.053)	(0.157)	(0.061)
R^2	0.1996	0.1565	0.1611	0.3005	0.1563
Num. observations	103,560	1,454,294	693,198	708,719	6,071,681
Num. covariates	20	20	20	20	20
F-statistic	1,358.7	14,201.6	7,008.2	16,025.8	59,178.4

$^{\dagger}: p < 0.1, ^{*}: p < 0.05, ^{**} p < 0.01, ^{***}: p < 0.001$.