Demand pull versus resource push training approaches to entrepreneurship: A field experiment

ONLINE APPENDIX

Exit Analysis

Some participants decided to quit the program for various reasons that are mostly exogenous to startup performance (e.g. personal reasons). In Table B1, we test whether startup exit likelihood is influenced by our intervention. The baseline exit rate in the overall sample was around 27%, which is in line with what is typically observed in similar entrepreneurship programs (Anderson et al., 2018). Our findings reveal that the *Demand Pull* approach did not have any systematic effect on the exit rate. Also, *Exit* is not influenced by relevant covariates. These results confirm that the covariate balance afforded by randomization is preserved in the final sample.

TABLE B1. OLS regressions testing intervention effect on exit likelihood (end of program)

Variables	***(1)*** ***Exit***	***(2)*** ***Exit***	***(3)*** ***Exit***
Demand Pull	0.044	0.045	0.046
	(0.059)	(0.059)	(0.058)
	[0.454]	[0.453]	[0.425]
Second Round		–0.061	0.017
		(0.063)	(0.066)
		[0.332]	[0.799]
Age			–0.004
			(0.003)
			[0.179]
Female			–0.059
			(0.062)
			[0.336]
Chinese			–0.080
			(0.063)
			[0.204]
Education			–0.038
			(0.078)
			[0.632]
Registered			–0.168
			(0.066)
			[0.011]
Working			–0.030
			(0.075)
			[0.686]
Studying			–0.050
			(0.092)
			[0.588]
Entrepreneurship experience			–0.088
			(0.060)
			[0.143]
Initial Revenue(cat)			0.061
			(0.038)
			[0.105]
Initial Customers(cat)			–0.054
			(0.042)
			[0.199]
Constant	0.270	0.288	0.565
	(0.042)	(0.047)	(0.167)
	[0.000]	[0.000]	[0.001]
Number of startups	236	236	225
R-squared	0.002	0.006	0.070

Notes: Robust standard errors are in parentheses. *P*-values are in brackets. Constant refers to the *Resource Push* group.

Time Effort

In Table B2, we test the relationship between participants' time effort and our intervention. We relied on the interviews carried out by our research assistants to construct a variable (*Time effort*) capturing the amount of time the participant dedicated to her startup. *Time effort* is a continuous variable ranging from 0 to 5 that indicates the amount of time the participant dedicated to the startup (Mean = 2.2; SD = 1.41; Min = 0; Max = 5). A value equal to 0 indicates that the participant spent no time working on the business idea, whereas 5 indicates the participant worked full time on the business idea. Our findings suggest that the demand pull training did not have any systematic impact on the time effort. This confirms that our results are genuinely driven by different activities carried out by participants assigned to the two groups rather than differences in motivation or commitment.

TABLE B2. OLS regressions testing intervention effect on time effort (end of intervention)

Variables	***(1)*** ***Time Effort***	***(2)*** ***Time Effort***	***(3)*** ***Time Effort***
Demand Pull	–0.078	–0.078	–0.066
	(0.205)	(0.206)	(0.172)
	[0.705]	[0.706]	[0.701]
Second Round		0.010	–0.187
		(0.214)	(0.175)
		[0.964]	[0.286]
Age			0.009
			(0.011)
			[0.412]
Female			–0.087
			(0.187)
			[0.643]
Chinese			0.326
			(0.185)
			[0.079]
Education			–0.017
			(0.222)
			[0.938]
Registered			0.748
			(0.242)
			[0.002]
Working			0.137
			(0.228)
			[0.550]
Studying			–0.286
			(0.295)
			[0.334]
Entrepreneurship experience			0.491
			(0.193)
			[0.012]
Initial Revenue(cat)			0.289
			(0.116)
			[0.013]
Initial Customers(cat)			0.088
			(0.173)
			[0.612]
Constant	2.244	2.241	0.762
	(0.147)	(0.166)	(0.544)
	[0.000]	[0.000]	[0.162]
Number of startups	192	192	189
R-squared	0.001	0.001	0.327

Notes: Robust standard errors are in parentheses. *P*-values are in brackets. Constant refers to the *Resource Push* group.

Including exited ventures

Participants who left the program before its conclusion did not complete our final performance survey and interview, hence they are treated as missing values in our main analysis. Yet, it is worth exploring what happens to our results when these participants are included in the sample. We repeated our difference-in-differences analysis in Table 3 assuming that *Revenue* (*Customers*) of all exited participants at the end of the program coincides with their initial revenue (initial customers). This conservative assumption implies that there is no treatment effect on exited participants. Results, reported in Table B3, as expected are weaker but still present.

TABLE B3. Intervention effect on startup performance when exited ventures are included

Variables	(1) ***Revenue***	(2) ***Revenue(cat)***	(3) ***Customers***	(4) ***Customers(cat)***
Program End	365.000	0.026	5.191	0.078
	(430.710)	(0.080)	(3.765)	(0.080)
	[0.398]	[0.743]	[0.169]	[0.330]
Program End × Demand Pull	1,530.826	0.255	17.916	0.418
	(842.122)	(0.135)	(7.454)	(0.137)
	[0.070]	[0.060]	[0.017]	[0.003]
Constant	1,549.576	1.458	7.051	1.424
	(213.135)	(0.034)	(1.887)	(0.034)
	[0.000]	[0.000]	[0.000]	[0.000]
Observations	472	472	472	472
Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.043	0.036	0.079	0.104
Number of startups	236	236	236	236

Notes: Robust standard errors are in parentheses. *P*-values are in brackets. In case both *Initial Revenue* and end-of-program *Revenue* are missing we assign zero to the startup revenue (9 cases). If both *Initial Customers* and end-of-program *Customers* are missing we assign zero to the value (11 cases).

Program Attendance

We replicated our analysis on just the participants that attended at least 80% of the program sessions. This was indeed the necessary threshold to obtain a certificate of attendance from the program organizers. We expected these participants (66% of the total sample) to be more motivated to work on their business idea than the overall sample and to have absorbed most of the program content, including our intervention. Results are reported in Table B4. As expected, the effect of the *Demand Pull* training is amplified in this smaller subsample.

TABLE B4. Intervention effect on startup performance for participants with at least 80% attendance

Variables	(1) ***Revenue***	(2) ***Revenue(cat)***	(3) ***Customers***	(4) ***Customers(cat)***
Program End	–156.071	–0.014	–1.029	0.000
	(370.093)	(0.126)	(3.101)	(0.102)
	[0.674]	[0.910]	[0.741]	[1.000]
Program End × Demand Pull	2,883.607	0.395	35.971	0.743
	(1,216.737)	(0.207)	(10.931)	(0.200)
	[0.019]	[0.059]	[0.001]	[0.000]
Constant	1,800.338	1.497	6.441	1.464
	(291.475)	(0.049)	(2.594)	(0.047)
	[0.000]	[0.000]	[0.014]	[0.000]
Observations	296	296	295	295
Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.067	0.046	0.129	0.167
Number of startups	155	155	155	155

Notes: Robust standard errors are in parentheses. *P*-values are in brackets.

Mediation Analysis

We ran a mediation analysis to explore how activities performed by our entrepreneurs in the intervention phase act as a mediator of our treatment. Specifically, we tested whether the variable *Customer Interaction* mediates the effect of demand pull training on performance. The results are reported in Table B5. Model 1 reports the correlation between *Customer Interaction* and *Demand Pull,* Model 2 reports the intervention effect on startup *Revenue(cat)* when the variable *Customer Interaction* is introduced in the regression. Model 3 reports the intervention effect on startup *Customers(cat)* when the variable *Customer Interaction* is introduced in the regression. *Customer Interaction* is strongly correlated with both *Demand Pull* and the performance variables. When *Customer Interaction* is added to the main regression in Models 2 and 3, the effect of *Demand Pull* on performance decreases. Using the causal mediation analysis package available in STATA (i.e. medeff command), we find a meaningful mediation effect between 7% and 39% of the total effect on *Customers(cat)*. The average mediation effect on *Revenue(cat)* is estimated to be 23% but with a very large confidence interval. These combined results point to a partial mediation effect. We believe the absence of full mediation is due to noise associated with the variable *Customer Interaction,* which was manually coded by our research assistants based on their interviews with participants and thus may not be able to fully capture the effects of *Demand Pull* training on the behavior and activities of our participants.

TABLE B5. Mediation analysis results

Variables	*(1)* ***Customer Interaction***	*(2)* ***Revenue(cat)***	*(3)* ***Customers(cat)***
Demand Pull	0.742	0.186	0.544
	(0.237)	(0.211)	(0.216)
	[0.002]	[0.378]	[0.013]
Customer Interaction		0.133	0.140
		(0.064)	(0.066)
		[0.040]	[0.035]
Constant	1.689	1.391	1.405
	(0.172)	(0.188)	(0.192)
	[0.000]	[0.000]	[0.000]
Number of startups	192	147	146
R-squared	0.049	0.039	0.083

Notes: Standard errors are in parentheses. *P*-values are in brackets. Constant refers to the *Resource Push* group.

Separate Regression Analyses

In the main analysis, we pool observations from different rounds of the program into the same regression analysis to estimate the average treatment effect. An alternative approach is to treat the studies independently and use a meta-analysis framework to estimate an average effect. The limitation of this approach is that each round of the program, taken separately, does not have a sufficiently large sample size to derive meaningful results. Nevertheless, we report the results of this additional analysis below. First, we report regression results separately for each round of the program.

TABLE B6. Intervention effect on startup performance for participants (first round only)

Variables	(1) ***Revenue***	(2) ***Revenue(cat)***	(3) ***Customers***	(4) ***Customers(cat)***
Program End	–146.875	–0.018	–1.411	–0.071
	(412.245)	(0.149)	(3.915)	(0.119)
	[0.722]	[0.905]	[0.719]	[0.549]
Program End × Demand Pull	1,766.261	0.246	25.625	0.536
	(1,126.519)	(0.235)	(11.534)	(0.206)
	[0.119]	[0.298]	[0.028]	[0.010]
Constant	1,050.372	1.480	6.977	1.462
	(238.148)	(0.050)	(2.428)	(0.043)
	[0.000]	[0.000]	[0.005]	[0.000]
Observations	269	269	266	266
Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.036	0.017	0.074	0.086
Number of startups	156	156	154	154

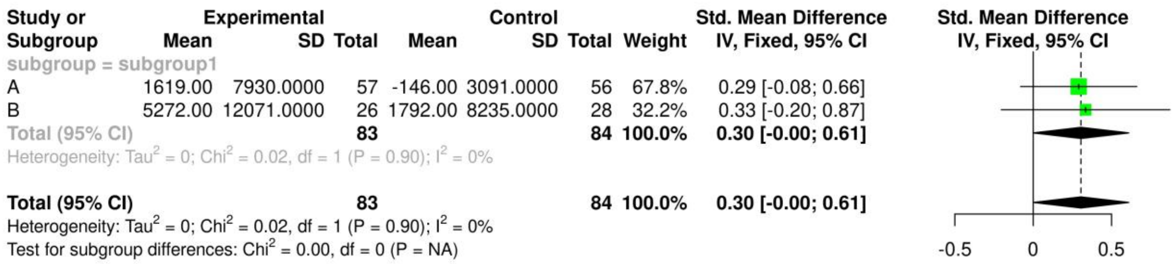
Notes: Robust standard errors are in parentheses. *P*-values are in brackets.

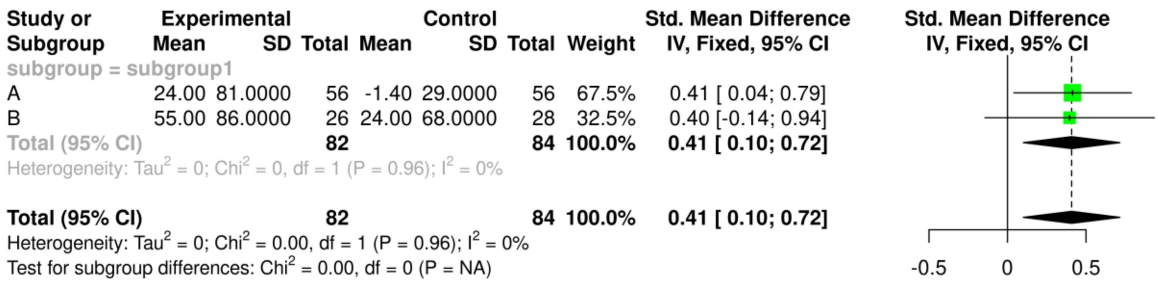
TABLE B7. Intervention effect on startup performance for participants (second round only)

Variables	(1) ***Revenue***	(2) ***Revenue(cat)***	(3) ***Customers***	(4) ***Customers(cat)***
Program End	1,792.857	0.143	24.143	0.464
	(1,551.721)	(0.133)	(12.827)	(0.214)
	[0.252]	[0.286]	[0.064]	[0.034]
Program End × Demand Pull	3,479.835	0.665	31.242	0.843
	(2,821.951)	(0.321)	(21.276)	(0.378)
	[0.222]	[0.042]	[0.146]	[0.029]
Constant	3,020.800	1.496	8.976	1.456
	(600.976)	(0.068)	(4.552)	(0.081)
	[0.000]	[0.000]	[0.053]	[0.000]
Observations	125	125	125	125
Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.129	0.202	0.234	0.339
Number of startups	71	71	71	71

Notes: Robust standard errors are in parentheses. *P*-values are in brackets.

Second, we use a meta-analysis framework to estimate an average effect for the two studies. Our results, reported below, display a positive albeit small effect of the demand pull training on startup revenue size of Cohen's *d = 0.30 (95% CI [-0.01; 0.61])*. In addition, the positive effect of the demand pull training on the number of acquired customers is *d = 0.41 (95% CI [0.10; 0.72])*. As expected, the treatment effect on startup revenue is weakly meaningful due to the large standard deviation. Standardized effect size and statistical significance increase when we use the number of new customers as a dependent variable.





The figures are created using the open software www.meta-mar.com.

Power calculations

We conducted calculations to estimate the minimum number of subjects required for adequate study power, using the descriptive statistics collected during the first round of the program as a reference for effect size and standard deviation. These calculations were performed for different dependent variables including *Revenue* in Tables B8 and B9, *Revenue(cat)* in Tables B10 and B11, and *Customers* in Tables B12 and B13. The variables involved in these calculations included the total sample size (N), sample sizes for the control (N1) and treatment (N2) groups, expected increases in revenue at the end of the program for the control (m1) and treatment (m2) groups, the difference between m2 and m1 (delta), and the standard deviation in the entire sample (sd).

Based on our calculations, we concluded that a sample size of at least 220 observations is enough to detect an effect size of roughly 2,000 SGD increase in *Revenue* with sufficient power ($\beta = 0.8$) and a type I error probability $\alpha = 0.1$. Assuming a probability of type I error $\alpha = 0.05$ and power $\beta = 0.8$, the minimum number of observations to detect a similar effect increases to more than 260. The minimum number of observations drops if we use a categorical variable of revenue or number of customers. A sample size of at least 180 observations is enough to detect an effect size of a roughly 0.4-point increase in *Revenue(cat)* with sufficient power ($\beta = 0.8$) and probability of type I error $\alpha = 0.1$. Assuming a probability of type I error $\alpha = 0.05$ and power $\beta = 0.8$, the number goes up to 220 observations.

When considering *Customers* as a dependent variable, a sample size of at least 140 observations is enough to detect an effect size of 25 new customers with sufficient power ($\beta = 0.8$) and probability of type I error $\alpha = 0.1$. Assuming a probability of type I error $\alpha = 0.05$ and power $\beta = 0.8$, the number goes up to 180 observations.

TABLE B8. Power calculations using *Revenue* as the dependent variable (α = 0.1)

alpha	power	N	N1	N2	delta	m1	m2	sd
.1	.2136	100	50	50	1000	0	1000	6000
.1	.2578	140	70	70	1000	0	1000	6000
.1	.3006	180	90	90	1000	0	1000	6000
.1	.3419	220	110	110	1000	0	1000	6000
.1	.3817	260	130	130	1000	0	1000	6000
.1	.3452	100	50	50	1500	0	1500	6000
.1	.4322	140	70	70	1500	0	1500	6000
.1	.5108	180	90	90	1500	0	1500	6000
.1	.5808	220	110	110	1500	0	1500	6000
.1	.6427	260	130	130	1500	0	1500	6000
.1	.5046	100	50	50	2000	0	2000	6000
.1	.6247	140	70	70	2000	0	2000	6000
.1	.72	180	90	90	2000	0	2000	6000
.1	.7938	220	110	110	2000	0	2000	6000
.1	.8498	260	130	130	2000	0	2000	6000
.1	.6643	100	50	50	2500	0	2500	6000
.1	.7905	140	70	70	2500	0	2500	6000
.1	.8728	180	90	90	2500	0	2500	6000
.1	.9244	220	110	110	2500	0	2500	6000
.1	.956	260	130	130	2500	0	2500	6000
.1	.799	100	50	50	3000	0	3000	6000
.1	.903	140	70	70	3000	0	3000	6000
.1	.9551	180	90	90	3000	0	3000	6000
.1	.9799	220	110	110	3000	0	3000	6000
.1	.9912	260	130	130	3000	0	3000	6000

TABLE B9. Power calculations using *Revenue* as the dependent variable (α = 0.05)

alpha	power	N	N1	N2	delta	m1	m2	sd
.05	.1309	100	50	50	1000	0	1000	6000
.05	.165	140	70	70	1000	0	1000	6000
.05	.1993	180	90	90	1000	0	1000	6000
.05	.2336	220	110	110	1000	0	1000	6000
.05	.2677	260	130	130	1000	0	1000	6000
.05	.2358	100	50	50	1500	0	1500	6000
.05	.3119	140	70	70	1500	0	1500	6000
.05	.3853	180	90	90	1500	0	1500	6000
.05	.4547	220	110	110	1500	0	1500	6000
.05	.5192	260	130	130	1500	0	1500	6000
.05	.3786	100	50	50	2000	0	2000	6000
.05	.4994	140	70	70	2000	0	2000	6000
.05	.6041	180	90	90	2000	0	2000	6000
.05	.6919	220	110	110	2000	0	2000	6000
.05	.7635	260	130	130	2000	0	2000	6000
.05	.541	100	50	50	2500	0	2500	6000
.05	.6872	140	70	70	2500	0	2500	6000
.05	.7939	180	90	90	2500	0	2500	6000
.05	.8679	220	110	110	2500	0	2500	6000
.05	.9172	260	130	130	2500	0	2500	6000
.05	.6969	100	50	50	3000	0	3000	6000
.05	.8358	140	70	70	3000	0	3000	6000
.05	.9156	180	90	90	3000	0	3000	6000
.05	.9583	220	110	110	3000	0	3000	6000
.05	.9801	260	130	130	3000	0	3000	6000

TABLE B10. Power calculations using *Revenue(cat)* as the dependent variable (α = 0.1)

alpha	power	N	N1	N2	delta	m1	m2	sd
.1	.2615	100	50	50	.2	0	.2	1
.1	.3225	140	70	70	.2	0	.2	1
.1	.3804	180	90	90	.2	0	.2	1
.1	.4349	220	110	110	.2	0	.2	1
.1	.486	260	130	130	.2	0	.2	1
.1	.6337	100	50	50	.4	0	.4	1
.1	.7612	140	70	70	.4	0	.4	1
.1	.8481	180	90	90	.4	0	.4	1
.1	.9053	220	110	110	.4	0	.4	1
.1	.942	260	130	130	.4	0	.4	1
.1	.909	100	50	50	.6	0	.6	1
.1	.9704	140	70	70	.6	0	.6	1
.1	.991	180	90	90	.6	0	.6	1
.1	.9974	220	110	110	.6	0	.6	1
.1	.9993	260	130	130	.6	0	.6	1
.1	.99	100	50	50	.8	0	.8	1
.1	.9989	140	70	70	.8	0	.8	1
.1	.9999	180	90	90	.8	0	.8	1
.1	1	220	110	110	.8	0	.8	1
.1	1	260	130	130	.8	0	.8	1
.1	.9996	100	50	50	1	0	1	1
.1	1	140	70	70	1	0	1	1
.1	1	180	90	90	1	0	1	1
.1	1	220	110	110	1	0	1	1
.1	1	260	130	130	1	0	1	1

TABLE B11. Power calculations using *Revenue(cat)* as the dependent variable (α = 0.05)

alpha	power	N	N1	N2	delta	m1	m2	sd
.05	.1677	100	50	50	.2	0	.2	1
.05	.2171	140	70	70	.2	0	.2	1
.05	.2663	180	90	90	.2	0	.2	1
.05	.3148	220	110	110	.2	0	.2	1
.05	.362	260	130	130	.2	0	.2	1
.05	.5082	100	50	50	.4	0	.4	1
.05	.6517	140	70	70	.4	0	.4	1
.05	.7608	180	90	90	.4	0	.4	1
.05	.8397	220	110	110	.4	0	.4	1
.05	.8949	260	130	130	.4	0	.4	1
.05	.8439	100	50	50	.6	0	.6	1
.05	.9412	140	70	70	.6	0	.6	1
.05	.9795	180	90	90	.6	0	.6	1
.05	.9932	220	110	110	.6	0	.6	1
.05	.9979	260	130	130	.6	0	.6	1
.05	.9773	100	50	50	.8	0	.8	1
.05	.9969	140	70	70	.8	0	.8	1
.05	.9996	180	90	90	.8	0	.8	1
.05	1	220	110	110	.8	0	.8	1
.05	1	260	130	130	.8	0	.8	1
.05	.9986	100	50	50	1	0	1	1
.05	1	140	70	70	1	0	1	1
.05	1	180	90	90	1	0	1	1
.05	1	220	110	110	1	0	1	1
.05	1	260	130	130	1	0	1	1

TABLE B12. Power calculations using *Customers* as the dependent variable (α = 0.1)

alpha	power	N	N1	N2	delta	m1	m2	sd
.1	.2136	100	50	50	10	0	10	60
.1	.2578	140	70	70	10	0	10	60
.1	.3006	180	90	90	10	0	10	60
.1	.3419	220	110	110	10	0	10	60
.1	.3817	260	130	130	10	0	10	60
.1	.3452	100	50	50	15	0	15	60
.1	.4322	140	70	70	15	0	15	60
.1	.5108	180	90	90	15	0	15	60
.1	.5808	220	110	110	15	0	15	60
.1	.6427	260	130	130	15	0	15	60
.1	.5046	100	50	50	20	0	20	60
.1	.6247	140	70	70	20	0	20	60
.1	.72	180	90	90	20	0	20	60
.1	.7938	220	110	110	20	0	20	60
.1	.8498	260	130	130	20	0	20	60
.1	.6643	100	50	50	25	0	25	60
.1	.7905	140	70	70	25	0	25	60
.1	.8728	180	90	90	25	0	25	60
.1	.9244	220	110	110	25	0	25	60
.1	.956	260	130	130	25	0	25	60
.1	.799	100	50	50	30	0	30	60
.1	.903	140	70	70	30	0	30	60
.1	.9551	180	90	90	30	0	30	60
.1	.9799	220	110	110	30	0	30	60
.1	.9912	260	130	130	30	0	30	60

TABLE B13. Power calculations using *Customers* as the dependent variable (α = 0.05)

alpha	power	N	N1	N2	delta	m1	m2	sd
.05	.1309	100	50	50	10	0	10	60
.05	.165	140	70	70	10	0	10	60
.05	.1993	180	90	90	10	0	10	60
.05	.2336	220	110	110	10	0	10	60
.05	.2677	260	130	130	10	0	10	60
.05	.2358	100	50	50	15	0	15	60
.05	.3119	140	70	70	15	0	15	60
.05	.3853	180	90	90	15	0	15	60
.05	.4547	220	110	110	15	0	15	60
.05	.5192	260	130	130	15	0	15	60
.05	.3786	100	50	50	20	0	20	60
.05	.4994	140	70	70	20	0	20	60
.05	.6041	180	90	90	20	0	20	60
.05	.6919	220	110	110	20	0	20	60
.05	.7635	260	130	130	20	0	20	60
.05	.541	100	50	50	25	0	25	60
.05	.6872	140	70	70	25	0	25	60
.05	.7939	180	90	90	25	0	25	60
.05	.8679	220	110	110	25	0	25	60
.05	.9172	260	130	130	25	0	25	60
.05	.6969	100	50	50	30	0	30	60
.05	.8358	140	70	70	30	0	30	60
.05	.9156	180	90	90	30	0	30	60
.05	.9583	220	110	110	30	0	30	60
.05	.9801	260	130	130	30	0	30	60