Table 1: Regression

	Dependent varsiable: Revenues per four-week period			Dependent variable: Shifts per four-week period		
	Messengers participating in experiment	All messengers at Veloblitz	All messengers at Flash and Veloblitz	Messengers participating in experiment	All messengers at Veloblitz	All messengers at Flash and Veloblitz
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment dummy	1,033.560***	1,094.496***	1,090.199***	3.986***	4.083***	4.064***
	(336.704)	(282.873)	(270.641)	(1.136)	(0.995)	(0.983)
Dummy for nontreated at Veloblitz			−54.398			−0.629
			(305.213)			(1.108)
Treatment period 1	−210.973	−370.619	−264.758	−1.275	−1.570*	−0.717
	(385.305)	(264.578)	(223.643)	(1.299)	(0.931)	(0.812)
Treatment period 2	−574.713	−656.233**	−650.479***	−2.561**	−2.631***	−2.216***
	(378.235)	(261.412)	(223.643)	(1.276)	(0.919)	(0.812)
Observations	124	190	386	124	190	386
R^2	0.740	0.786	0.753	0.694	0.740	0.695
Adjusted R^2	0.596	0.666	0.610	0.524	0.593	0.519
Residual Std. Error	1,541.221 (df = 79)	1,364.353 (df = 121)	1,305.812 (df = 244)	5.198 (df = 79)	4.798 (df = 121)	4.742 (df = 244)
Notes:	***Significant at the 1 percent level.					
	**Significant at the 5 percent level.					
	*Significant at the 10 percent level.					
	Robust standard errors, adjusted for clustering on messengers, are in parentheses. *Source*: Own calculations.					

Appendix: the probit model
Suppose we have latent response variable
$y^* = \mathbf{X}\beta + \varepsilon$
where $\mathbf{X}$ is a $k \times 1$ vector of features $[x_1 \; x_2 \; \dots \; x_k]$ and β is a $1 \times k$ coefficient vector.
The observable binary variable y is defined as

$$\begin{aligned} y &= 1 \quad \text{if} \quad y^* > 0 \\ y &= 0 \quad \text{if} \quad y^* \leq 0 \end{aligned}$$

If we assume that $\varepsilon \sim N(0,1)$ then

$$\begin{aligned} P(y^* > 0) &= P(\mathbf{X}\beta + \varepsilon > 0) \\ &= P(\varepsilon > -\mathbf{X}\beta) \\ &= P(\varepsilon < \mathbf{X}\beta) \quad \text{By symmetry of standard normal} \\ &= \Phi(\mathbf{X}\beta) \end{aligned}$$

So $\mathbf{X}\beta$ are z-scores:

$$\begin{aligned} P(y = 1) &= P(y^* > 0) = \Phi(z \leq \mathbf{X}\beta) \\ P(y = 0) &= P(y^* \leq 0) = 1 - \Phi(z \leq \mathbf{X}\beta) \end{aligned}$$

where Φ is the Standard Normal CDF (e.g. $\Phi(0) = 0.5$; half the standard normal distribution lies below $\mu = 0$).
If we relax the assumption that the error is standard Normal and instead allow it be $\varepsilon \sim N(0, \sigma^2)$, then

$$\begin{aligned} P(y^* > 0) &= P(\mathbf{X}\beta + \varepsilon > 0) \\ &= P(\frac{\varepsilon}{\sigma} > \frac{-\mathbf{X}\beta}{\sigma}) \\ &= P(\frac{\varepsilon}{\sigma} < \frac{\mathbf{X}\beta}{\sigma}) \\ &= \Phi(\frac{\mathbf{X}\beta}{\sigma}) \end{aligned}$$

but we cannot estimate β and σ separately since the probability depends exclusively on their ratio. The standard approach is assume $\sigma = 1$ so ε is a standard normal.

$$ATE = \mathbb{E}[Y_i | T = 1] - \mathbb{E}[Y_i | T = 0] \tag{1}$$

$$n = (\frac{2s_p^2 (z_\alpha + z_\beta)^2}{d^2})(1 + (c-1)\rho) \tag{2}$$

$$T \in \{0, 1, 2\}$$

$$\begin{aligned} H_0 &: \beta = 0 \\ H_1 &: \beta \neq 0 \end{aligned}$$