

The Motivational Paradox: How Aspirational-Goal Communications Influence Sales Performance and Information Acquisition

Online Appendix

A.1. Aspirational-Goal Communications

The format of the aspirational-goal communications was the same across all employees in the treatment group. The firm delivered the messages by text in the mornings before employees came to work. The contents of the messages differed based on employees' past performance. The firm customized the messages such that the goal was one performance tier higher than an employee's past average performance (over 50 workdays before the start of the field experiment). **Table A.1** shows sample messages by performance tier.

Table A.1. Aspirational-Goal Communications

Performance tier	Daily text message
Tier 1: Below 1,400 SEK/hour	If you sell $(1,400 \text{ SEK/hour} \times 8 \text{ hours} = 11,200 \text{ SEK})$ today, you will receive $(11,200 \text{ SEK} \times 0.27\% = 30 \text{ SEK})$.
Tier 2: 1,400 – 1,799 SEK/hour	If you sell $(1,800 \text{ SEK/hour} \times 8 \text{ hours} = 14,400 \text{ SEK})$ today, you will receive $(14,400 \text{ SEK} \times 0.67\% = 96 \text{ SEK})$.
Tier 3: 1,800 – 1,999 SEK/hour	If you sell $(2,000 \text{ SEK/hour} \times 8 \text{ hours} = 16,000 \text{ SEK})$ today, you will receive $(16,000 \text{ SEK} \times 0.9\% = 144 \text{ SEK})$.
Tier 4: 2,000 – 2,349 SEK/hour	If you sell $(2,350 \text{ SEK/hour} \times 8 \text{ hours} = 18,800 \text{ SEK})$ today, you will receive $(18,800 \text{ SEK} \times 1.5\% = 282 \text{ SEK})$.

Notes. The table shows sample aspirational-goal communications by performance tier for agents scheduled to work 8 hours per day. Employees' average performance over the 50 workdays before the field experiment determines their performance tiers. No employee's average SPH over the 50 workdays before the field experiment exceeded 2,350 SEK.

A.2. Instrumental Variables

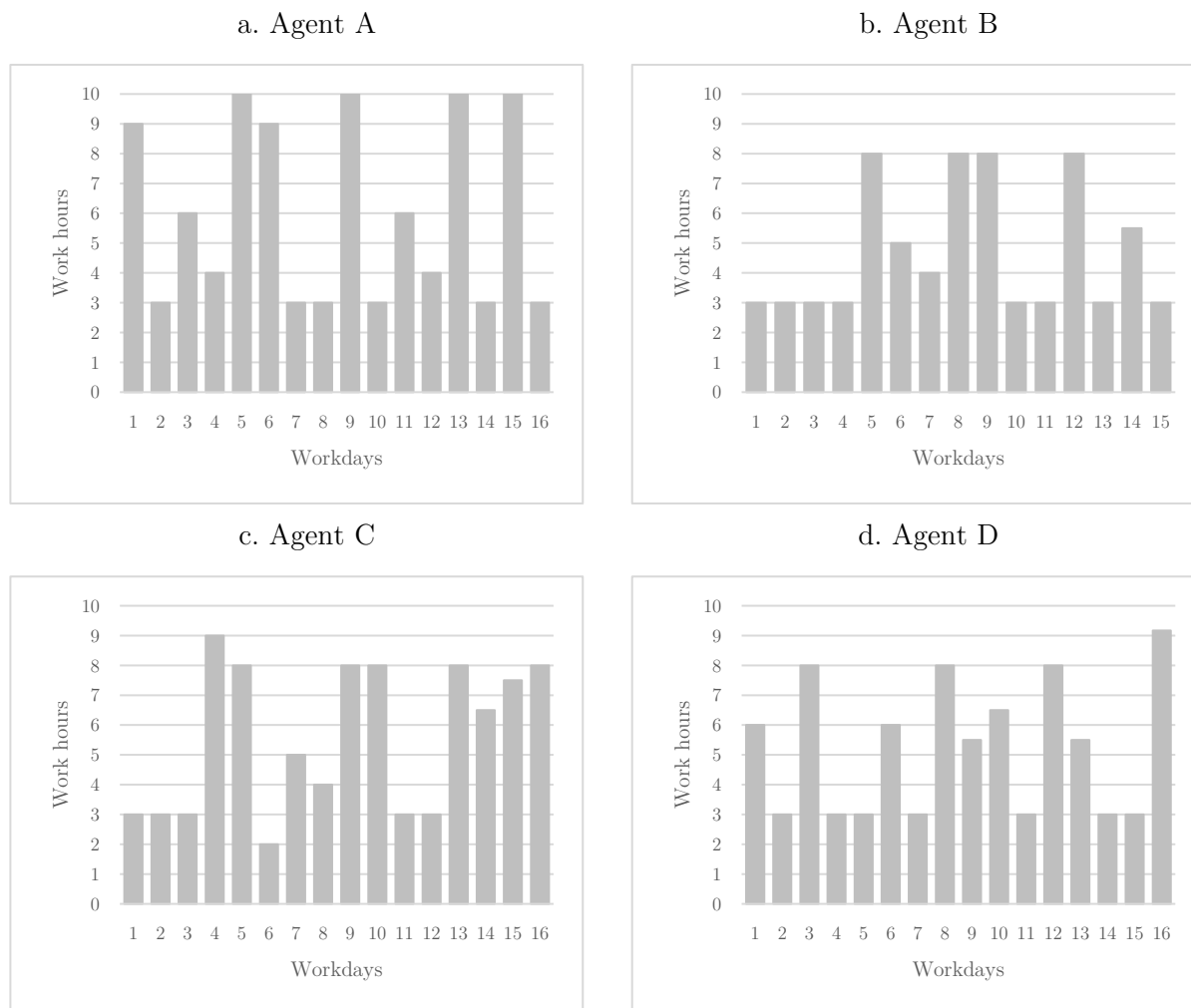
This section evaluates the validity of the instrumental-variable approach by examining three key conditions: variation, to ensure the instruments provide sufficient identifying power; exogeneity, to

rule out correlation with unobserved factors; and relevance, to confirm that the instruments are strong predictors of the endogenous variable.

A.2.1. Variation

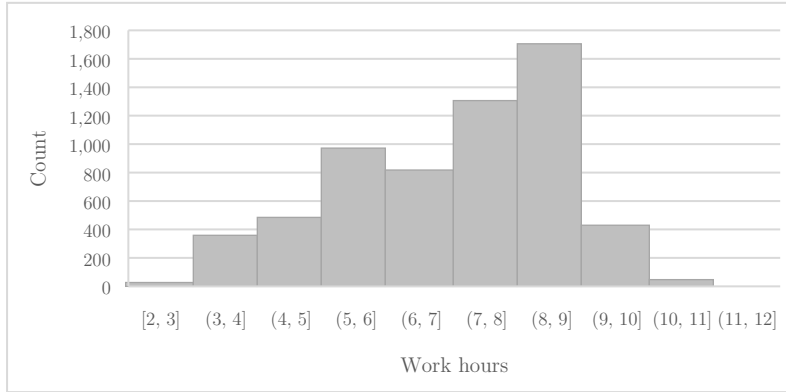
Figures A.1a–d illustrate variation in daily work hours for select agents during the field experiment. Work hours exhibit sufficient variation and fluctuate across workdays; the average within-agent coefficient of variation is 0.21. **Figure A.2** shows the distribution of daily work hours across agents and workdays, which also illustrates sufficient variation.

Figure A.1. Variation in Daily Work Hours for Select Agents



Notes. The figure shows the variation in daily work hours for select agents during the treatment period.

Figure A.2. Variation in Work Hours across Agents and Workdays



Notes. The figure shows the distribution of work hours across agents and workdays in the treatment period.

A.2.2. Exogeneity

Table A.2 shows the results of various model specifications, where the dependent variable is the logarithm of daily SPH and the independent variables include agents' work hours. Across all specifications, the coefficients on work hours are small in magnitude and statistically insignificant. These results indicate that, within the observed range constrained by the upper bound on daily work hours, work hours do not appear to directly affect agents' performance in terms of SPH, thereby supporting their exogeneity.

Table A.2. Exogeneity Condition of Work Hours as an Instrumental Variable for SPM

	(1)	(2)	(3)
Work hours	0.001 (0.003)	0.001 (0.003)	0.001 (0.003)
SPM	0.055*** (0.003)	0.055*** (0.004)	
Aspirational-goal stimuli		0.049** (0.024)	
SPM (Tier 1)			0.081*** (0.007)
SPM (Tier 2)			0.057*** (0.006)
SPM (Tier 3)			0.047*** (0.004)
SPM (Tier 4)			0.042*** (0.006)
Aspirational-goal stimuli (Tier 1)			0.118***

			(0.032)
Aspirational-goal stimuli (Tier 2)			0.060**
			(0.025)
Aspirational-goal stimuli (Tier 3)			0.009
			(0.040)
Aspirational-goal stimuli (Tier 4)			-0.017
			(0.035)
Agent fixed effects	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
Observations	12,427	12,427	12,427

Notes. SPM. Sales-progress monitoring. The dependent variable is the logarithm of daily sales per hour. Clustered standard errors (at the store level) are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To further examine the validity of the instrumental-variable approach, we conduct a falsification test using only the pretreatment period. If longer shifts systematically reduced productivity through fatigue, assigned work hours would correlate negatively with SPH.

Table A.3 reports the results of the falsification test. Across all performance tiers, the estimated coefficients on work hours are small in magnitude, positive, and statistically insignificant. This absence of a systematic relationship suggests that longer shifts do not directly lower productivity (measured in SPH). Taken together with the results in **Table A.2**, these findings provide additional support for the exogeneity of work hours as instruments.

Table A.3. Falsification Test

	(1)
Work hours (Tier 1)	0.0084 (0.0091)
Work hours (Tier 2)	0.0068 (0.0062)
Work hours (Tier 3)	0.0014 (0.0085)
Work hours (Tier 4)	0.0200 (0.0128)
Agent fixed effects	Yes
Time fixed effects	Yes
Observations	6,276

Notes. The dependent variable is the logarithm of daily sales per hour. Column (1) shows the results of a falsification test using only the pretreatment period, during which sales-progress monitoring is coded as zero due to the lack of data. Clustered standard errors (at the store level) are in parentheses.

A.2.3. Relevance

To assess the instruments’ relevance (predictive power), we project the endogenous variable—sales-progress monitoring (SPM)—onto the instrumental variables, a step commonly referred to as estimating the first-stage equation. The instruments consist of interaction terms between agents’ daily work hours, performance tiers, and treatment period. As noted in the main text, the firm began tracking SPM only at the start of the field experiment; consequently, SPM values are coded as zero in the pre-experiment period. The specification therefore exploits variation during the treatment period.

Table A.4 reports the results of the first-stage equations, where Columns (1)–(4) show the first-stage estimates for SPM in Tiers 1–4, respectively. Agents’ scheduled work hours during the treatment period are highly significant across performance tiers, indicating a strong correlation between the focal instrumental variables and the endogenous variables, thereby satisfying the relevance condition. The Sanderson–Windmeijer multivariate F-statistics are all close to 80 (ranging from 79.8 to 88.7), far above conventional thresholds, which strongly confirms joint instrument strength. The individual first-stage F-statistics are 16.0, 25.0, 14.1, and 18.7 for Tiers 1 through 4, respectively—each comfortably above the rule-of-thumb cutoff of 10. The signs of the coefficients are also consistent with intuition: agents are more likely to conduct SPM on days when they are scheduled to work longer hours.

Table A.4. Relevance of Instrumental Variables: First-Stage Equation

	(1)	(2)	(3)	(4)
Work hours $\times$ Treatment period (Tier 1)	0.210*** (0.027)	-0.010 (0.020)	0.003 (0.009)	0.012 (0.010)
Work hours $\times$ Treatment period (Tier 2)	-0.004 (0.010)	0.293*** (0.031)	0.002 (0.010)	0.013 (0.011)
Work hours $\times$ Treatment period (Tier 3)	-0.004 (0.010)	-0.012 (0.021)	0.374*** (0.051)	0.013 (0.010)
Work hours $\times$ Treatment period (Tier 4)	-0.003 (0.009)	-0.009 (0.019)	0.002 (0.009)	0.468*** (0.055)
Control variables	Yes	Yes	Yes	Yes
Agent fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
F test of excluded instruments	16.00	24.98	14.05	18.66

Sanderson-Windmeijer F-test	79.81	82.51	83.80	88.71
Kleibergen-Paap LM statistic (p-value)		38.09 (0.000) ^a		
Kleibergen-Paap rk Wald F statistic		19.85		
Observations	12,427	12,427	12,427	12,427

Notes. Agents' performance over the 50 workdays prior to the field experiment determines their tiers as follows: Tier 1, below 1,400 SEK/hour; Tier 2, 1,400-1,799 SEK/hour; Tier 3, 1,800-1,999 SEK/hour; and Tier 4, 2,000-2,349 SEK/hour. Columns (1)-(4) show the results of the first-stage equations for the endogenous variable, sales-progress monitoring, of Tiers 1-4, respectively. Clustered standard errors (at the store level) are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. ^a $\chi^2(1)$.

Regarding identification, the Kleibergen–Paap LM statistic rejects the null of underidentification ($\chi^2(1)=38.1$, $p < 0.001$). In addition, the Kleibergen–Paap Wald F-statistic is 19.9, above the conventional threshold of 10, further indicating that weak identification is not a concern. Taken together, the large Sanderson–Windmeijer multivariate F-statistics, the strong individual first-stage results across tiers, and the robust rejection of underidentification provide compelling evidence that the instruments are both relevant and powerful predictors of SPM.

A.2.4. Alternative Specification with Additional Instruments

The above analysis (and the main text) relies on a just-identified specification in which agents' daily work hours interacted with the treatment period and performance tiers serve as instruments for SPM. As a robustness check, we extend the instrument set by including work hours in all periods as additional instruments. This specification increases the number of excluded instruments and allows us to compute an overidentification test of the exclusion restriction (Hansen J test).

Table A.5 reports the corresponding first-stage estimates. Across all tiers, the interaction terms between work hours and the treatment period remain strong and significant predictors of SPM, whereas the main effects of work hours are generally small. This is unsurprising, as the endogenous variable, SPM, is not available in the pretreatment period and is coded as zero for all agents, meaning that work hours outside the treatment period add little identifying variation beyond what is captured by the interaction terms. The individual F-tests of the excluded instruments range between 7.8 and 14.7, and the Sanderson–Windmeijer multivariate F-statistics exceed 20 in all cases, indicating adequate instrument strength. The Kleibergen–Paap LM statistic rejects the null of underidentification ($\chi^2(5)=45.8$, $p < 0.001$), and the Kleibergen–Paap Wald F-statistic is 15.2, above conventional thresholds. Importantly, the Hansen J statistic ($\chi^2(4)=9.11$, $p=0.058$) does not reject

the null that the instruments are valid at the 5% level, providing some additional support for the exclusion restriction.¹

Table A.5. Alternative Specification: First-Stage Estimates

	(1)	(2)	(3)	(4)
Work hours (Tier 1)	-0.045** (0.022)	0.018 (0.014)	-0.003 (0.006)	-0.004 (0.006)
Work hours (Tier 2)	0.004 (0.006)	-0.009 (0.016)	-0.004 (0.006)	-0.004 (0.006)
Work hours (Tier 3)	0.005 (0.006)	0.020 (0.014)	-0.079*** (0.025)	-0.002 (0.006)
Work hours (Tier 4)	0.005 (0.005)	0.019 (0.012)	-0.001 (0.005)	-0.044 (0.031)
Work hours × Treatment period (Tier 1)	0.221*** (0.029)	-0.016 (0.028)	0.014 (0.012)	0.018 (0.013)
Work hours × Treatment period (Tier 2)	0.001 (0.012)	0.290*** (0.037)	0.014 (0.013)	0.019 (0.014)
Work hours × Treatment period (Tier 3)	0.000 (0.012)	-0.018 (0.028)	0.394*** (0.053)	0.019 (0.013)
Work hours × Treatment period (Tier 4)	0.001 (0.011)	-0.016 (0.027)	0.012 (0.012)	0.482*** (0.056)
Control variables	Yes	Yes	Yes	Yes
Agent fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
F test of excluded instruments	9.26	14.73	7.78	11.61
Sanderson-Windmeijer F-test	26.78	23.37	25.45	23.49
Kleibergen-Paap LM statistic (p-value)		45.8 (0.000) ^a		
Kleibergen-Paap rk Wald F statistic		15.16		
Hansen J (p-value)		9.11 (0.058) ^b		
Observations	12,427	12,427	12,427	12,427

Notes. Agents' performance over the 50 workdays prior to the field experiment determines their tiers as follows: Tier 1, below 1,400 SEK/hour; Tier 2, 1,400-1,799 SEK/hour; Tier 3, 1,800-1,999 SEK/hour; and Tier 4, 2,000-2,349 SEK/hour. Columns (1)-(4) show the results of the first-stage equations for the endogenous variable, sales-progress monitoring, of Tiers 1-4, respectively. Clustered standard errors (at the store level) are in parentheses. Significance: * p<0.10, ** p<0.05, *** p<0.01. ^a $\chi^2(5)$. ^b $\chi^2(4)$.

¹ The Hansen J statistic ($\chi^2(4)=9.11$, $p=0.058$) is marginal, but it does not reject the null at the 5% level (Wooldridge, 2010; Baum et al., 2003). Moreover, overidentification tests such as Hansen's J are known to have low power in finite samples and can be affected by heteroskedasticity, many instruments, or clustered errors (Chao et al., 2014). Thus, marginal p-values should not be interpreted as strong evidence against instrument validity. When viewed alongside the design-based exogeneity and the strong first-stage results, these findings are consistent with the instruments satisfying the exclusion restriction.

Table A.6 presents the second-stage estimates under this alternative specification. Column (1) replicates the ordinary-least-squares results, while Column (2) reports the two-stage-least-squares estimates. The results are consistent with the just-identified model. The SPM coefficients for Tiers 1 and 2 are positive and statistically significant, whereas those for Tiers 3 and 4 are small and statistically insignificant. The coefficients on the aspirational-goal stimuli remain positive and significant for Tiers 1 and 2, with similar magnitudes across specifications. Taken together, these results show that including work hours in all periods as instruments does not materially alter the findings and adds little additional identifying variation, as SPM is unobserved in the pretreatment period. The main benefit of this specification is that it enables overidentification tests, which lend further credibility to the validity of the instruments.

Table A.6. Alternative Specification: Second-Stage Estimates

	(1)	(2)
SPM (Tier 1)	0.0812*** (0.007)	0.0744*** (0.025)
SPM (Tier 2)	0.0572*** (0.006)	0.0262* (0.015)
SPM (Tier 3)	0.0467*** (0.004)	0.0153 (0.013)
SPM (Tier 4)	0.0416*** (0.006)	-0.00185 (0.012)
Aspirational-goal stimuli (Tier 1)	0.118*** (0.032)	0.0647* (0.033)
Aspirational-goal stimuli (Tier 2)	0.0603** (0.025)	0.0428* (0.023)
Aspirational-goal stimuli (Tier 3)	0.00884 (0.039)	-0.00698 (0.034)
Aspirational-goal stimuli (Tier 4)	-0.0169 (0.035)	0.0398 (0.036)
Agent fixed effects	Yes	Yes
Time fixed effects	Yes	Yes
Observations	12,427	12,427

Notes. SPM. Sales-progress monitoring. The dependent variable is the logarithm of daily sales per hour. Agents' performance over the 50 workdays prior to the field experiment determines their tiers as follows: Tier 1, below 1,400 SEK/hour; Tier 2, 1,400-1,799 SEK/hour; Tier 3, 1,800-1,999 SEK/hour; and Tier 4, 2,000-2,349 SEK/hour. Column (1) shows the results of the ordinary-least-squares estimator. Column (2) shows the results of the two-stage-least-squares estimator, using interaction terms between agents' daily work hours and performance tiers, along with their interactions with the treatment period. The results of the first-stage estimates are in **Table A.5**. Clustered standard errors (at the store level) are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

- Baum, C. F., Schaffer, M. E., & Stillman, S. (2003). Instrumental variables and GMM: Estimation and testing. *The Stata Journal*, 3(1), 1-31.
- Chao, J. C., Hausman, J. A., Newey, W. K., Swanson, N. R., & Woutersen, T. (2014). Testing overidentifying restrictions with many instruments and heteroskedasticity. *Journal of Econometrics*, 178, 15-21.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.