

Submitted to *Marketing Science*
manuscript (Please, provide the manuscript number!)

Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.

Job Security, Gender, and Sales Performance: Evidence from a Retail Sales Context

Diego F. Salazar

NUS Global Asia Institute, diegosalazar.net@gmail.com,

Noah Lim

NUS Business School and NUS Global Asia Institute, noahlim@nus.edu.sg,

Ingrid Koch

UNC-Chapel Hill, ingrid_koch@kenan-flagler.unc.edu,

Online Appendix for Job Security, Gender, and Sales Performance: Evidence from a Retail Sales Context

Online Appendix A. Compensation Schemes

During the study period, Company T implemented four different compensation schemes. Importantly, these compensation schemes started at the beginning of a calendar year and lasted at least one full year. The first compensation scheme was implemented between 2011 and 2013, and agents' earnings were part fixed salary and part variable incentives. The variable incentives were primarily based on store performance (e.g., aggregate sales and operational cost savings) and the agents' relative performance within the store.

The second compensation scheme was launched in 2014 and was characterized by the introduction of individual commissions for sales agents. Specifically, agents received monthly commissions based on the proportion of the individual target achieved. Such commissions replaced the variable incentives of the first compensation scheme.

The third compensation scheme was implemented between 2015 and 2016, and it was the result of Company T's new goals and the lessons learned from the previous schemes. This new scheme aimed to incentivize sales agents by imposing more demanding targets and increasing the commission for achieving those targets. This scheme was basically an updated version of the incentive program launched in 2014.

Finally, the last modification to the compensation scheme was implemented between 2017 and 2019, and one important innovation was introduced: Company T tied the individual monthly compensation scheme to the store's performance. In this scheme, Company T provided monthly individual and store targets. The compensation for each sales agent was calculated based on the proportion of the individual target achieved and paid conditional on the store reaching its own targets. According to management, the idea behind this scheme was to generate accountability and cooperation among the sales team so everybody had incentives to support each other and improve store performance.

Online Appendix B. Construction of Relative Ranking Measure

Defining an objective measure of relative standings at Company T is difficult since agents worked in very different stores, were exposed to different peers, and joined in different periods. As a result, confounders of seasonality, store, and group quality likely have a decisive influence over rankings. Therefore, to construct a meaningful ranking, we first need to purge sales from seasonal trends, store quality, and group characteristics. In such a way, we remove key factors unrelated to the agent's effort and skills and then rank the agents based on the residual portion of the sales to produce a measure of relative performance.

To define the ranking, we follow these steps. First, during the pre-treatment period, we regress (log) sales on all the time fixed effects (year, month, day-of-the-week), store fixed effects, store type, store growth in the previous quarter, and different regressors capturing measures of team quality, i.e., the number of employees with permanent contracts and average team experience. Second, based on the estimates, we predicted the sales portion attributable to the season, store, and team. Third, we subtract such a portion from the total sales during the pre-treatment period and obtain the residual sales. Finally, we obtain the average residual sales during the pre-treatment period and rank the agents based on them to generate the variable Rank_i .

With the measure of relative performance (Rank_i) at hand, we include the interaction $\text{PermContract}_{it} \cdot \text{Rank}_i$ in our main specification. In our preferred specification, we split the sample into four quartiles, and Rank_i refers to the quartiles of the distribution of relative performance. Therefore, the upper quartile represents the top 25% of agents in terms of average sales during the pre-treatment period, after removing seasonal trends, store and team characteristics. Conditional on ability, these agents are the strongest performers and constitute the baseline category. Table A13 in the Appendix presents the results of this exercise and shows how the effects change for agents in quartiles one to three.

Online Appendix C. Validity of the Design: Further Tests

Joint Significance Test of the Interaction Terms Between the Treated Group Dummy and Time Indicators

To implement the test, we augment our main specification to include pre-treatment time dummies and their interaction with the treated group dummies. Specifically, we estimate the following equation:

$$Y_{it} = \beta \text{PermContract}_{it} + \gamma \text{After}_{it} + X'_{it} \theta + Z'_{it} \phi + \alpha_i + \tilde{\lambda}_t + \sum_{j=q}^m \theta_j D_{ct}(t = k_c - j) + \epsilon_{it}. \quad (2)$$

In this specification, subscript i represents the agent and t represents time (in days). Our key variable of interest is $D_{ct}(t = k_c - j) = d_{ct}(t = k_c - j) * \text{treat}_i$, where treat_i is the treatment variable that takes the value of one for the treatment group and zero for the control (i.e., treated group dummy), and $d_{ct}(t = k_c - j)$ are time dummies covering 30-day intervals before the treatment period. For example, $d_{c1}(t = k_c - 1)$ is a dummy that covers the first 30-day period before treatment. Therefore, $d_{c1}(t = k_c - 1)$ takes the value of one during the period between 515 to 545 days after starting work for Company T (i.e., 30 days before the treatment period), and zero otherwise. In addition, $\tilde{\lambda}_t$ includes the usual time fixed effects and $d_{ct}(t = k_c - j)$ dummies.

We run different specifications, including different amounts of interactions, and results are robust across specifications. For example, we test for parallel trends between 245 days and 545 days after starting work for Company T, i.e., up to 300 days (ten months) before the treatment period, by including ten different time dummies and testing the joint significance of the interaction terms between the treated group dummy and time indicators. In addition to looking at the ten-month period before the treatment period, we also tested for parallel trends by focusing on the six-month period before the intervention. Results for the first exercise are reported in Table A3, which also includes the F-test for joint significance of θ_j . In all cases, we find supporting evidence that the assumption of parallel trends holds in our context.

Parallel Trends in Residual Sales

To conduct the test, we first “purge” daily sales of differential time trends and other differences resulting from a sales agent belonging to different cohorts. Then, we construct a measure of residual sales as the difference between the sales and the predicted sales. After that, using residual sales, we run a joint significance test of the interaction terms between the treated group dummy and different time indicators similar to the one from specification 2. Table A5 presents the estimates for the interaction terms using 10 different time dummies and the F-tests for joint significance. In all cases, the F-tests provide statistical evidence favoring parallel trends.

Placebo Dates

We conduct several placebo tests. These tests help alleviate further endogeneity concerns (i.e., worker selection, heterogeneous ability, and learning patterns over tenure) that our main specification may not control for. We chose different “placebo” dates during the pre-treatment period, starting with the contract being awarded 365 days after starting at Company T and then different placebo dates every 30 days up to 485 days (e.g., 365 days, 395 days, 425 days, 455 days and 485 days after starting at Company T). In total, we estimate our baseline specification using five different placebo dates and restrict the data to the pre-treatment period, i.e., all observations up to 545 days after starting work for Company T. To mimic our main empirical strategy, we run the main model using different time frames starting from 30 days after the placebo date, up to 150 days after the placebo date. Importantly, the later the chosen placebo date, the fewer windows for estimation we can use. Table A6 presents the estimates of this exercise. The conclusion is the same in all cases: we find no evidence of permanent contracts generating productivity increases before they were awarded.

Online Appendix D. Additional Tables

Table OA1 Sales agent characteristics: Control vs. Grace Period Leavers

	Control (N=278)	Grace Period Leavers (N=51)	p-value (equality of both groups)
	(1)	(2)	(3)
Means Before the Treatment			
Daily Performance Measures			
Sales (in thousand Indonesian Rupiah)	3676.16 (2652)	5880.24 (6355.43)	0.000
Average Items per Transaction	2.07 (0.87)	2.39 (1.17)	0.000
Average Item Price (in thousand Indonesian Rupiah)	1459.72 (353.98)	1662.13 (415.77)	0.000
Number of transactions	13.52 (9.13)	17.51 (22.31)	0.000
Work Environment			
Morning shift	0.48 (0.5)	0.50 (0.5)	0.000
Weekend shift	0.47 (0.5)	0.47 (0.5)	0.628
Store Class			
AA	0.22 (0.42)	0.28 (0.45)	0.000
A	0.13 (0.33)	0.10 (0.3)	0.000
B	0.38 (0.49)	0.45 (0.5)	0.000
C	0.04 (0.19)	0.03 (0.17)	0.000
Other	0.24 (0.42)	0.14 (0.35)	0.000
Total	1.00	1.00	
Year			
2011-2012	0.17 (0.38)	0.04 (0.19)	0.000
2013-2014	0.57 (0.5)	0.00 (0.06)	0.000
2015-2016	0.26 (0.44)	0.00 (0)	0.000
2017-2019	0.00 (0)	0.96 (0.2)	0.000
Total	1.00	1.00	

Notes:

Column (3) presents the p-value for the Kruskal–Wallis test of the hypothesis that several samples are from the same population. Work Environment describes the store shifts that the workers had, Store Class is the class of the main store the agent worked in, and Year represents the period when agents started working at Company T. One thousand Indonesian Rupiah roughly equals \$0.067 USD using the exchange rate from May 23, 2023. Standard deviations are shown in parentheses.

Table OA2 Effects of permanent contracts on daily sales: All relevant coefficients

	180 days post- treatment	150 days post- treatment	90 days post- treatment	30 days post- treatment
	(1)	(2)	(3)	(4)
Permanent Contract	0.067** (0.025)	0.064* (0.026)	0.054* (0.025)	0.060* (0.025)
Permanent Contract x Male	0.021 (0.056)	0.019 (0.059)	-0.0030 (0.061)	0.024 (0.060)
After 18 months	0.032 (0.017)	0.059** (0.018)	0.086*** (0.019)	0.067*** (0.017)
Tenure	0.95*** (0.069)	0.99*** (0.073)	1.09*** (0.079)	1.17*** (0.083)
Tenure sq.	-0.18*** (0.021)	-0.23*** (0.023)	-0.31*** (0.028)	-0.32*** (0.030)
Multi-store	0.15* (0.062)	0.16* (0.064)	0.17* (0.071)	0.18* (0.077)
Class A	0.26* (0.12)	0.19 (0.12)	0.0095 (0.12)	-0.25 (0.16)
Class B	0.35** (0.13)	0.28* (0.13)	0.30 (0.16)	0.30 (0.19)
Class C	0.24 (0.20)	0.12 (0.22)	0.12 (0.23)	0.17 (0.24)
Class other	-1.14*** (0.14)	-1.21*** (0.14)	-1.23*** (0.17)	-1.19*** (0.20)
Store's growth (previous quarter)	0.000086 (0.00011)	0.00011 (0.00011)	0.000070 (0.00011)	0.0000062 (0.00012)
Permanent contracts (previous quarter)	-0.0086* (0.0039)	-0.0087* (0.0040)	-0.012** (0.0042)	-0.013** (0.0044)
Constant	9.13*** (0.25)	9.01*** (0.26)	8.94*** (0.27)	9.04*** (0.27)
Groups	535	534	534	534
Observations	248468	237978	216201	194970
Individual FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Store FE	Yes	Yes	Yes	Yes

Notes:

Controls include a dummy for the post-treatment period, tenure (in years), and tenure squared, and a dummy for working multi-store. Additionally, we include year, month, and day of the week fixed effects. Finally, regarding store characteristics, we include store fixed effects, store type dummies, the store's quarterly growth from the previous quarter, and the number of permanent contracts signed by agents in the store during the previous quarter. Model 1 includes observations up to 735 days of tenure (roughly 180 days after signing the permanent contract), Model 2 up to 705 days of tenure (roughly 150 days after signing the permanent contract), Model 3 up to 645 days of tenure (roughly 90 days after signing the permanent contract), and Model 4 up to 585 days of tenure (roughly 30 days after signing the permanent contract). Tenure (in years) is calculated as the difference between the agent's starting date at Company T and the observation date. Standard errors are in parentheses and clustered at the sales agent level.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.