Online Supplement for "Balancing Agent Retention and Waiting Time in Service Platforms"

A Instrumental variable estimation using the ratio of outbound calls to incoming quotes as IV

The estimation with IVs is implemented through a logistic regression with a control function. To construct the effective service capacity measure, the time period comprised between the quote arrival and elapsed time of the call (w_{it}) is divided into one-hour length intervals. For each interval, a measure of congestion was calculated by taking the ratio of the outbound calls in that interval and the number of new quotes arriving to the system. The higher the value of this metric, the higher the capacity of the system to respond to the quotes in that period. An average is taken across the time intervals to construct the instrumental variable Z_{it}, named *Effective capacity*. To implement the IV via a control function, the first step is a linear regression of w_{it} on X_{it} and Z_{it}. A function of the residual of this regression is added as an additional covariate to the logistic model (equation 2) to account for a potential correlation between ϵ_{it} and w_{it}. The main results use a quadratic polynomial of the residual to specify the control function. A linear specification yields similar results.

To validate the relevance of the IV, Table A1 shows the results of the first step regression, which reveals a negative and highly significant coefficient for the Effective capacity. Hence, our measure of effective capacity is a strong predictor of waiting time: as capacity increases, waiting times are reduced.

For the IV to be exogenous, it must be unrelated to unobserved factors that affect a customer's interest in answering the call. To validate this, we first ran a logistic regression to predict whether a customer would answer the first call based on observable variables related to the quote and customer characteristics. As we have presented before, these variables are effective to discriminate whether a customer will answer the phone, with the bottom and top 5% predicted probabilities ranging from 0.37 to 0.67; hence, this predicted probability is a reasonable proxy of the customers likelihood of answering the call.

Then, we regressed this prediction on the effective capacity; the goal is to assess whether customers with a higher propensity to answer the phone (based on observable characteristics) tend to arrive at periods of lower congestion, as the mechanism presented in the reviewer's comment would suggest. Table A2 shows the estimation results of this regression; the first column includes a linear measure of the effective capacity, whereas the second column includes a quadratic term to capture potential non-linear effects. In all cases, the estimates suggest a very weak association between between the predicted answering probability and the effective capacity, with an R-square which is essentially zero (the linear term is statistically significant due to the large sample size, but with a very small magnitude).

Overall, our analysis suggests that customers and quotes that tend to have a higher probability of answering the phone do not arrive in a systematic pattern that is associated with the level of effective service capacity of the system, thereby providing evidence that: (i) the platform does not adapt service capacity in anticipation of the arrivals of customer which are more likely to purchase; (ii) hence the ratio of outbound calls to incoming quotes is exogenous and can be used as a

valid IV.

	(1)	
	Wait time	
Effective capacity	-1.394***	(0.0904)
DayWeekQuote_0	-0.126	(1.638)
DayWeekQuote_1	-1.877	(1.279)
DayWeekQuote_2	-4.153**	(1.332)
DayWeekQuote_3	-4.190**	(1.315)
DayWeekQuote_4	-4.505***	(1.315)
DayWeekQuote_5	-2.224	(1.291)
DayWeekQuote_6	0	(.)
HourQuote_0	598.4***	(60.79)
HourQuote_1	558.4***	(60.80)
HourQuote_2	511.9***	(60.83)
HourQuote_3	467.6***	(60.89)
HourQuote_4	392.5***	(61.05)
HourQuote_5	338.3***	(60.95)
HourQuote_6	269.4***	(60.87)
HourQuote_7	192.5**	(60.75)
HourQuote_8	122.9*	(60.70)
HourQuote_9	75.90	(60.69)
HourQuote_10	70.00	(60.69)
HourQuote_11	71.21	(60.69)
HourQuote_12	72.84	(60.69)
HourQuote_13	72.96	(60.69)
HourQuote_14	75.09	(60.69)
HourQuote_15	77.69	(60.69)
HourQuote_16	68.08	(60.69)
HourQuote_17	63.25	(60.69)
HourQuote_18	59.55	(60.69)
HourQuote_19	55.17	(60.70)
HourQuote_20	49.71	(60.70)
HourQuote_21	43.66	(60.79)
HourQuote_22	0	(.)
HourQuote_23	0	(.)
HourQuote_24	0	(.)
Google searches	-0.0318	(0.0241)
Observations	65931	
R^2	0.611	

Standard errors in parentheses

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

Table A1: Regression of waiting time to first call attempt

	(1)		(2)	
	Prob(Answer)		Prob(Answer)	
Effective capacity	-0.000259**	(0.0000804)	-0.0000835	(0.000175)
Effective cap. Sq.			-0.00000492	(0.00000436)
Constant	0.575***	(0.000685)	0.574***	(0.00104)
Observations	65929		65929	
R^2	0.000		0.000	

Standard errors in parentheses

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

Table A2: Regression of the probability of answering the phone on the first call attempt

B Additional estimation results

	Norm. Workload		Experience (Quotes)	
Time since quote arrival (w)				
$w \leq 4[m]$	0.549***	(0.0898)	0.516**	(0.168)
$4[m] < w \leq 5[m]$	0.470***	(0.0996)	0.546**	(0.184)
$5[m] < w \leq 6[m]$	0.326**	(0.124)	0.280	(0.234)
$6[m] < w \leq 10[m]$	0.530***	(0.118)	0.565**	(0.219)
$10[m] < w \leq 20[m]$	0.348*	(0.138)	0.208	(0.273)
$20[m] < w \leq 40[m]$	0.229	(0.159)	0.217	(0.299)
$40[m] < w \leq 60[m]$	0.177	(0.199)	-0.731	(0.488)
$60[m] < w \leq 80[m]$	-0.0540	(0.118)	-0.389^{+}	(0.216)
$80[m] < w \leq 120[m]$	-0.139	(0.167)	-0.589^{+}	(0.326)
Agent experience (K)				
$K \leq 4[weeks]$	0.138	(0.113)		
$K \leq 8[weeks]$	0.0143	(0.107)		
$K \leq 12[weeks]$	-0.0579	(0.113)		
$K \leq 16[weeks]$	0.105	(0.113)		
$K \leq 20[weeks]$	0.0233	(0.103)		
$K \leq 26[weeks]$	0.00681	(0.0906)		
Experience (quotes)			-0.000458	(0.000321)
Experience sq. (quotes)			4.91e-08	(7.75e-08)
Customer tenure	-0.0000342	(0.0000380)	-0.0000608	(0.0000669)
Male	-0.0722^{+}	(0.0418)	-0.0422	(0.0741)
Answer rate	0.0906***	(0.00291)	0.0914***	(0.00517)
Blog	0.143	(0.307)	-0.707	(0.760)
Direct	0.771***	(0.149)	0.594*	(0.245)
Display	-0.515	(0.484)	0.0199	(0.653)
Email	-0.549***	(0.159)	-0.817**	(0.266)
Organic Search	0.665***	(0.148)	0.689**	(0.244)
Other	1.103***	(0.235)	0.908*	(0.413)
Paid Search	0.690***	(0.138)	0.445^{+}	(0.227)
Referral	0.774***	(0.206)	0.432	(0.417)
Second call	-0.00150	(0.105)	0.278	(0.193)
Google searches	0.00539**	(0.00209)	-0.00201	(0.00369)
Price Index (min)	-0.992***	(0.0891)	-1.213***	(0.158)
Workload (L)			-0.00674	(0.00689)
Workload sq.(L^2)			0.0000503	(0.0000758)
Normalized workload	-0.00852	(0.0108)		
Normalized workload sq.	0.000223	(0.000372)		
Constant	-4.647***	(1.191)	-5.612***	(1.692)
Observations	31563		10762	

Standard errors in parentheses

$^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table A3: Estimation results of the conversion model. Table shows estimated parameters of the logit model defined by equation (3) with alternative metrics of worklaod (cumulative workload) and agent experience (number of quotes), using the sub-sample of quotes that were contacted in the first and second calls.

	Min H Residual		Mean H Residual		No Price	
Time since quote arrival (w)						
$w \leq 4[m]$	0.375***	(0.0294)	0.375***	(0.0294)	0.374***	(0.0294)
$4[m] < w \leq 5[m]$	0.215***	(0.0336)	0.215***	(0.0336)	0.215***	(0.0335)
$5[m] < w \leq 6[m]$	0.136**	(0.0427)	0.136**	(0.0427)	0.134**	(0.0427)
$6[m] < w \leq 10[m]$	-0.0691	(0.0403)	-0.0698	(0.0403)	-0.0703	(0.0402)
$10[m] < w \leq 20[m]$	-0.160***	(0.0461)	-0.159***	(0.0461)	-0.159***	(0.0461)
$20[m] < w \leq 40[m]$	-0.183***	(0.0526)	-0.183***	(0.0526)	-0.182***	(0.0526)
$40[m] < w \leq 60[m]$	-0.259***	(0.0655)	-0.259***	(0.0655)	-0.258***	(0.0655)
$60[m] < w \leq 80[m]$	0.0381	(0.0738)	0.0387	(0.0738)	0.0365	(0.0738)
$80[m] < w \leq 120[m]$	-0.0544	(0.0639)	-0.0536	(0.0639)	-0.0516	(0.0639)
Customer tenure	-0.000102***	(0.0000156)	-0.000102***	(0.0000156)	-0.000106***	(0.0000156)
Male	-0.0140	(0.0170)	-0.0144	(0.0170)	-0.0231	(0.0169)
Answer rate	-0.0696***	(0.00155)	-0.0695***	(0.00155)	-0.0693***	(0.00155)
Blog	0.0224	(0.110)	0.0226	(0.110)	0.0137	(0.110)
Direct	-0.0840	(0.0528)	-0.0853	(0.0528)	-0.0926	(0.0528)
Display	0.0167	(0.146)	0.0156	(0.146)	0.0125	(0.146)
Email	-0.0349	(0.0509)	-0.0350	(0.0509)	-0.0363	(0.0509)
Organic Search	-0.0294	(0.0521)	-0.0302	(0.0521)	-0.0349	(0.0521)
Other	0.201*	(0.0891)	0.201*	(0.0891)	0.199*	(0.0891)
Paid Search	-0.0110	(0.0472)	-0.0115	(0.0472)	-0.0158	(0.0472)
Referral	-0.150*	(0.0754)	-0.149*	(0.0755)	-0.158*	(0.0754)
Google searches	0.00101	(0.000827)	0.000994	(0.000827)	0.00101	(0.000827)
Price Index (min)	-0.175***	(0.0338)				
Price Index (mean)			-0.144***	(0.0347)		
Constant	-0.230	(0.237)	-0.184	(0.236)	-0.159	(0.236)
Observations	66012		66012		66034	

Standard errors in parentheses

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

Table A4: Estimation results of the contact model for the first call attempt under the base specification, using mean hedonic regression residuals and without using price in the model. The table shows estimated parameters of the logit model defined by equation (2).

	Min H Residual		Mean H Residual		No Price	
Time since Quote Arrival (w)						
$w \leq 4[m]$	-0.483**	(0.163)	-0.486**	(0.163)	-0.457**	(0.161)
$4[m] < w \leq 5[m]$	-2.416***	(0.213)	-2.418***	(0.213)	-2.418***	(0.213)
$5[m] < w \leq 6[m]$	-2.182***	(0.234)	-2.181***	(0.234)	-2.183***	(0.234)
$6[m] < w \leq 10[m]$	-1.381***	(0.161)	-1.380***	(0.161)	-1.381***	(0.161)
$10[m] < w \leq 20[m]$	-0.326*	(0.137)	-0.327*	(0.137)	-0.326*	(0.137)
$20[m] < w \leq 40[m]$	0.706***	(0.119)	0.706***	(0.119)	0.710***	(0.119)
$40[m] < w \leq 60[m]$	1.169***	(0.116)	1.169***	(0.116)	1.169***	(0.116)
$60[m] < w \leq 80[m]$	0.184**	(0.0598)	0.184**	(0.0598)	0.185**	(0.0598)
$80[m] < w \leq 120[m]$	0.0121	(0.0721)	0.0115	(0.0721)	0.0130	(0.0720)
Time since previous call (d)						
$d \leq 30[m]$	-1.550***	(0.0874)	-1.548***	(0.0873)	-1.544***	(0.0873)
$30[m] < d \leq 60[m]$	-0.967***	(0.136)	-0.968***	(0.136)	-0.967***	(0.135)
$60[m] < d \leq 90[m]$	-0.0414	(0.0690)	-0.0408	(0.0690)	-0.0396	(0.0689)
$90[m] < d \leq 120[m]$	-0.0104	(0.106)	-0.00882	(0.106)	-0.00895	(0.106)
Customer tenure	-0.000106***	(0.0000312)	-0.000107***	(0.0000312)	-0.000109***	(0.0000311)
Male	0.00285	(0.0341)	0.000902	(0.0342)	-0.00672	(0.0339)
Answer rate	-0.0636***	(0.00313)	-0.0634***	(0.00312)	-0.0633***	(0.00312)
Blog	0.0281	(0.215)	0.0266	(0.215)	0.0282	(0.215)
Direct	-0.113	(0.106)	-0.116	(0.106)	-0.126	(0.106)
Display	0.239	(0.303)	0.231	(0.302)	0.223	(0.302)
Email	-0.0199	(0.102)	-0.0214	(0.102)	-0.0255	(0.102)
Organic Search	-0.00909	(0.105)	-0.0127	(0.105)	-0.0181	(0.105)
Other	0.442**	(0.163)	0.445**	(0.163)	0.451**	(0.162)
Paid Search	0.0370	(0.0946)	0.0348	(0.0946)	0.0301	(0.0945)
Referral	0.0890	(0.146)	0.0882	(0.146)	0.0798	(0.146)
Google searches	0.000457	(0.00165)	0.000440	(0.00165)	0.000459	(0.00165)
Price Index (min)	-0.188**	(0.0672)				
Price Index (mean)			-0.132	(0.0693)		
Constant	0.902	(1.209)	0.904	(1.210)	0.903	(1.210)
Observations	20686		20686		20692	

Standard errors in parentheses

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

Table A5: Estimation results of the contact model for the second call attempt under the base specification, using mean hedonic regression residuals and without using price in the model. The table shows estimated parameters of the logit model defined by equation (2).

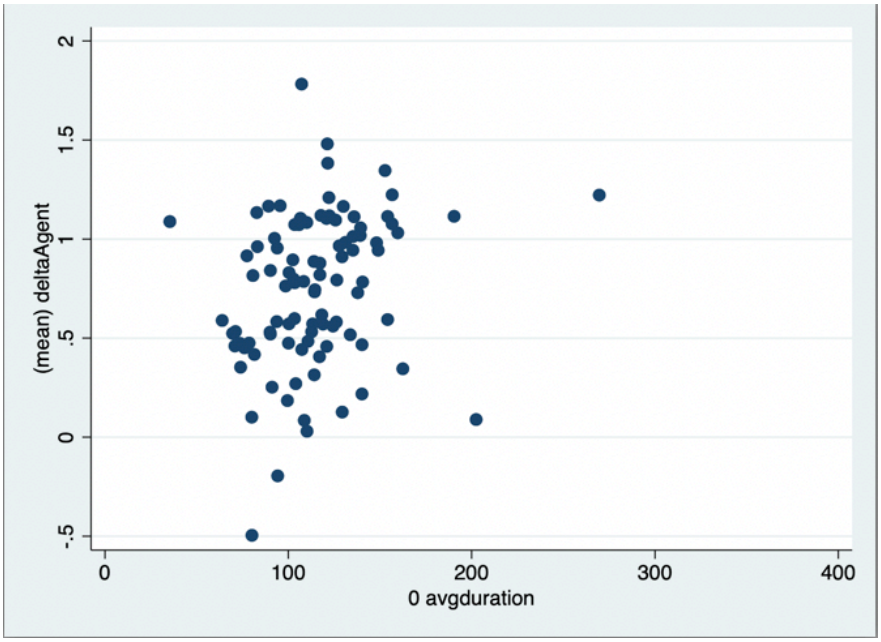

Figure 7: Scatter plot of agent fixed effects (from the conversion model) versus average service time for each agent. The figure suggests that there is no evident association between the skill of an agent (conversion rate) and the time spent to handle calls.

	Min H Residual		Mean H Residual		No Price	
Time since quote arrival (w)						
$w \leq 4[m]$	0.407***	(0.0801)	0.413***	(0.0800)	0.409***	(0.0800)
$4[m] < w \leq 5[m]$	0.360***	(0.0885)	0.364***	(0.0884)	0.361***	(0.0883)
$5[m] < w \leq 6[m]$	0.116	(0.112)	0.127	(0.112)	0.125	(0.112)
$6[m] < w \leq 10[m]$	0.395***	(0.105)	0.394***	(0.105)	0.394***	(0.105)
$10[m] < w \leq 20[m]$	0.329**	(0.121)	0.340**	(0.121)	0.342**	(0.121)
$20[m] < w \leq 40[m]$	0.149	(0.140)	0.150	(0.140)	0.148	(0.140)
$40[m] < w \leq 60[m]$	-0.0321	(0.182)	-0.0306	(0.181)	-0.0272	(0.181)
$60[m] < w \leq 80[m]$	-0.132	(0.104)	-0.132	(0.104)	-0.134	(0.104)
$80[m] < w \leq 120[m]$	-0.209	(0.145)	-0.213	(0.145)	-0.214	(0.145)
Agent experience (K)						
$K \leq 4[weeks]$	0.174+	(0.0960)	0.164+	(0.0959)	0.159+	(0.0958)
$K \leq 8[weeks]$	-0.0746	(0.0922)	-0.0678	(0.0921)	-0.0640	(0.0920)
$K \leq 12[weeks]$	0.0134	(0.0973)	0.0126	(0.0971)	0.0144	(0.0971)
$K \leq 16[weeks]$	0.104	(0.0966)	0.112	(0.0965)	0.109	(0.0964)
$K \leq 20[weeks]$	0.00875	(0.0920)	0.00247	(0.0919)	0.00146	(0.0919)
$K \leq 26[weeks]$	-0.0114	(0.0829)	-0.00831	(0.0828)	-0.00702	(0.0827)
Customer tenure	-0.0000114	(0.0000337)	-0.0000184	(0.0000337)	-0.0000298	(0.0000336)
Male	-0.0990**	(0.0367)	-0.115**	(0.0368)	-0.142***	(0.0365)
Answer rate	0.0905***	(0.00256)	0.0912***	(0.00256)	0.0913***	(0.00255)
Blog	0.0920	(0.272)	0.0843	(0.271)	0.0846	(0.271)
Direct	0.795***	(0.131)	0.771***	(0.131)	0.752***	(0.131)
Display	-0.176	(0.413)	-0.182	(0.413)	-0.186	(0.412)
Email	-0.491***	(0.139)	-0.491***	(0.138)	-0.490***	(0.138)
Organic Search	0.710***	(0.130)	0.700***	(0.130)	0.690***	(0.130)
Other	1.135***	(0.206)	1.134***	(0.206)	1.116***	(0.206)
Paid Search	0.671***	(0.121)	0.660***	(0.121)	0.648***	(0.121)
Referral	0.783***	(0.173)	0.766***	(0.173)	0.735***	(0.173)
Second call	-0.0305	(0.0917)	-0.0254	(0.0916)	-0.0273	(0.0916)
Workload (L)	-0.00721*	(0.00353)	-0.00709*	(0.00352)	-0.00716*	(0.00352)
Workload sq.(L^2)	0.0000559	(0.0000403)	0.0000556	(0.0000402)	0.0000560	(0.0000402)
Google searches	0.00450*	(0.00182)	0.00440*	(0.00181)	0.00444*	(0.00181)
Price Index (min)	-0.902***	(0.0784)				
Price Index (mean)			-0.468***	(0.0799)		
Constant	-3.984***	(1.139)	-3.675**	(1.138)	-3.633**	(1.138)
Observations	41110		41110		41117	

Standard errors in parentheses

+ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table A6: Estimation results of the conversion model. Table shows estimated parameters of the logit model defined by equation (3) using minimum, mean and no residuals of the hedonic regression as a price index, using the sub-sample of quotes that were contacted in the first and second calls and with agent experience measured using the number of weeks since joining the platform.

	Number of quotes	
Google searches	0.216***	(0.00621)
DayWeekQuote_1	40.02***	(0.413)
DayWeekQuote_2	32.18***	(0.422)
DayWeekQuote_3	27.59***	(0.416)
DayWeekQuote_4	28.86***	(0.418)
DayWeekQuote_5	19.66***	(0.409)
DayWeekQuote_6	-0.902*	(0.414)
HourQuote_0	-52.83***	(15.65)
HourQuote_1	-51.83***	(15.65)
HourQuote_2	-48.36**	(15.66)
HourQuote_3	-41.02**	(15.68)
HourQuote_4	-43.63**	(15.72)
HourQuote_5	-40.66**	(15.70)
HourQuote_6	-35.91*	(15.68)
HourQuote_7	-31.80*	(15.65)
HourQuote_8	-23.04	(15.64)
HourQuote_9	-9.816	(15.64)
HourQuote_10	27.95	(15.64)
HourQuote_11	20.37	(15.64)
HourQuote_12	10.15	(15.64)
HourQuote_13	-3.866	(15.64)
HourQuote_14	-4.841	(15.64)
HourQuote_15	0.259	(15.64)
HourQuote_16	-4.281	(15.64)
HourQuote_17	-10.57	(15.64)
HourQuote_18	-21.47	(15.64)
HourQuote_19	-29.64	(15.64)
HourQuote_20	-31.27*	(15.64)
HourQuote_21	-31.81*	(15.67)
HourQuote_22	0	(.)
HourQuote_23	0	(.)
HourQuote_24	0	(.)
MonthQuote_2	9.583***	(0.244)
MonthQuote_3	0.305	(0.226)
MonthQuote_4	-5.593***	(0.243)
MonthQuote_5	-1.637***	(0.250)
MonthQuote_6	-4.556***	(0.250)
MonthQuote_7	-2.440***	(0.236)
MonthQuote_8	1.940***	(0.371)
Constant	21.17	(15.65)
Observations	65953	
R^2	0.686	

Standard errors in parentheses

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

Table A7: Regression of the hourly number of quotes on several covariates, including Google Searches.

	1wk lag		2wk lag		4wk lag	
	(1)	(2)	(3)	(4)	(5)	(6)
Calls per Shift	-0.109**	-0.00140	-0.0707**	0.0546	-0.0670**	0.0283
	(0.0224)	(0.0496)	(0.0238)	(0.0623)	(0.0254)	(0.0653)
exper_cat=1	0.781	3.665*	0.492	5.075*	0.589	3.822
	(0.690)	(1.743)	(0.675)	(2.483)	(0.681)	(2.446)
exper_cat=2	0.626	4.083*	0.199	2.672	-0.0300	-0.156
	(0.788)	(1.850)	(0.779)	(2.736)	(0.810)	(2.892)
exper_cat=3	0.808	3.505^{+}	0.442	6.480*	0.271	5.798*
	(0.975)	(1.974)	(0.960)	(2.647)	(0.961)	(2.664)
exper_cat=4	0.375	1.859	-0.0252	4.135	-0.100	4.990
	(1.032)	(2.380)	(1.010)	(3.164)	(1.022)	(4.003)
Male	-0.786	-0.811	-0.734	-0.595	-0.654	-0.555
	(0.527)	(0.539)	(0.519)	(0.524)	(0.524)	(0.532)
Age 36-50	0.699	0.785	0.568	0.572	0.544	0.458
	(0.555)	(0.569)	(0.544)	(0.577)	(0.545)	(0.583)
Age $>$50	0.691	0.763	0.431	0.432	0.179	0.102
	(0.540)	(0.553)	(0.535)	(0.559)	(0.551)	(0.578)
deltaAgent_cat=1	-0.506	-0.530	-0.654	-0.765^{+}	-0.653	-0.824^{+}
	(0.441)	(0.443)	(0.438)	(0.449)	(0.447)	(0.471)
deltaAgent_cat=2	-1.257*	-1.340*	-1.343*	-1.220^{+}	-1.262*	-1.200^{+}
	(0.628)	(0.641)	(0.624)	(0.631)	(0.637)	(0.649)
Average call duration	0.00582	0.00689	0.00545	0.00466	0.00619	0.00437
	(0.00707)	(0.00705)	(0.00718)	(0.00729)	(0.00738)	(0.00760)
exper_cat=1 $\times$ Calls per Shift		-0.135*		-0.164*		-0.118
		(0.0619)		(0.0746)		(0.0777)
exper_cat=2 $\times$ Calls per Shift		-0.169*		-0.0705		0.0159
		(0.0718)		(0.0827)		(0.0908)
exper_cat=3 $\times$ Calls per Shift		-0.122		-0.262**		-0.253*
		(0.0777)		(0.0964)		(0.104)
exper_cat=4 $\times$ Calls per Shift		-0.0575		-0.147		-0.223
		(0.0895)		(0.119)		(0.196)
Constant	-1.892	-4.433*	-1.947	-5.749*	-1.976	-4.467^{+}
	(1.218)	(1.909)	(1.293)	(2.633)	(1.328)	(2.556)
AIC	273.5	274.1	288.1	285.3	279.1	275.3
BIC	337.6	359.6	351.7	370.1	342.1	359.3
N	1543	1543	1480	1480	1403	1403

Standard errors in parentheses. $^{+}$ $p<0.1$, * $p<0.05$, ** $p<0.01$

Table A8: Estimation results of the agent retention model, including all specification and control variables.