E-companion: Go Wide or Go Deep? Assortment Strategy and Order Fulfillment

Appendix A – Variable Definitions and Summary Statistics

In Table A1, we summarize the definitions of key variables used in our empirical models.

Table A1: Definitions of Key Variables

Variable	Description
$DeliveryTime_{ijt}$	Average time from the time of payment for merchant i to deliver orders placed during week t in category j
$DelayTime_{ijt}$	Average delay from the promised date for merchant i to deliver orders placed during week t in category j
$Width_{it}$	Number of categories offered by merchant i during week t
$Depth_{it}$	Average number of items per category offered by merchant i during week t
$log(Sales_{ijt})$	Log of aggregate sales for merchant i in category j during week t
$AvgPrice_{ijt}$	Average price of items in category j sold by merchant i during week t
LRS_{it}	Overall logistic review score that customers view at time of purchase for merchant i during week t
LP_{it}	Number of distinct logistic providers employed by merchant i during week t
$AvgOrderSize_{it}$	Average number of items across all orders placed with merchant i during week t
$LSPExperience_{ijt}$	Cumulative number of orders (in millions) handled by LSPs employed by merchant i during week t in category j

In Tables A2 and A3, we provide summary statistics and correlation matrix for all key dependent and independent variables used in our models.

Table A2: Descriptive Statistics

Variables	*Mean*	*Std.Dev.*	*Min*	*Pctl(25)*	*Pctl(75)*	*Max*
DeliveryTime	2.710	1.576	0.000	2.000	2.926	29.000
Width	16.810	11.870	1.000	8.000	23.000	61.000
Depth	43.420	59.730	1.250	11.000	46.700	502.000
log(Sales)	9.326	2.994	0.077	6.988	11.720	19.760
AvgPrice	228.400	445.200	0.010	23.400	192.200	3,231.000
LRS	4.806	0.303	0.000	4.769	4.976	5.000
LP	11.210	5.952	1.000	7.000	14.000	33.000

Table A3: Pairwise Correlations of Control Variables

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
(1) DeliveryTime	1.000						
(2) Width	0.045*	1.000					
(3) Depth	-0.072*	-0.049*	1.000				
(4) log(Sales)	0.037*	0.051*	-0.132*	1.000			
(5) AvgPrice	0.175*	0.197*	-0.174*	0.383*	1.000		
(6) LRS	-0.073*	0.015*	0.041*	0.009*	0.014*	1.000	
(7) LP	-0.128*	-0.022*	-0.125*	0.363*	0.002	0.015*	1.000

Note: $^{*}p<0.05$

Appendix B - First-Stage Regressions

We now provide the results of the first-stage regressions for the two instruments, $CompetingWidth_{it}$ and $CompetingDepth_{it}$, utilized in §4.1. The first-stage results of the 2SLS estimation show that both the instruments predict the two endogenous variables significantly. That is, both instruments satisfy the *relevance* criterion.

Table A4: First-stage Regressions of *Width* and *Depth*

	(1) *Width*	(2) *Depth*
CompetingWidth	0.600***	-0.615***
	(0.004)	(0.025)
CompetingDepth	-0.015***	0.177***
	(0.001)	(0.007)
log(Sales)	-0.010***	0.095***
	(0.002)	(0.012)
AvgPrice	0.000***	0.000*
	(0.000)	(0.000)
LRS	-0.025*	0.001
	(0.013)	(0.075)
LP	-0.007***	0.014
	(0.002)	(0.011)
Adjusted R^2	0.98	0.98

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$, *Robust standard errors in parentheses. Estimated with merchant, category, and week fixed effects, N = 64,233.*

Appendix C - Robustness Tests

We first present a series of robustness tests to confirm our main relationship between assortment strategy and order delivery timeliness identified in Hypotheses 1 and 2.

C.1. Delays on Promised Orders.

In §4, and specifically, equation (1), we measure the delivery timeliness of a retailer by $DeliveryTime_{ijt}$, the average time to satisfy the orders without a promised delivery date placed during week t pertaining to category j. We now consider an alternative measure of delivery timeliness, $DelayTime_{ijt}$ computed as the average delay on the subset of orders with a promised delivery date. Thus, we estimate the following model to test Hypotheses 1 and 2.

$$DelayTime_{ijt} = \alpha'_1 Depth_{it} + \alpha'_2 Width_{it} + CONTROLS_{ijt} + Merchant_i + Category_j + Week_t + \epsilon_{ijt} \quad (8)$$

To address endogeneity issues, as before, using 2SLS, we estimate the above model by employing the two instruments for assortment width and depth described previously.

Table A5: Impact of Assortment Strategy on Order Delivery Delays, *DelayTime*

	DelayTime
Width	0.0025^{**}
	(0.0012)
Depth	0.0016^{***}
	(0.0002)
LP	-0.0050^{***}
	(0.0010)
AvgPrice	0.0000^{**}
	(0.0000)
LRS	-0.1110^{***}
	(0.0210)
log(*Sales*)	0.0021
	(0.0020)
Adjusted R^2	0.297

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$, *Robust standard errors in parentheses. Estimated with merchant, category, and week fixed effects, N = 64,233.*

In Table A5, we again find that the main effects consistently since the coefficients for width and depth are positive, 0.0025 and 0.0016, and significant ($p < 0.05$). That is, an increase in assortment width and depth are both negatively associated with delays on orders with promised delivery dates confirming Hypothesis 1. Further, by comparing α'_1 and α'_2 using the Z-test, we find support for Hypothesis 2, i.e., assortment width is more negatively associated with delivery timeliness than depth.

C.2. Category Distinctiveness - Level of Analysis.

Two levels of product categorization are available in our dataset. In §4, we measured assortment width and depth at the lower level of product categorization in the dataset with a smaller distinctiveness between products across categories. However, assortment width and depth can also be measured at a higher level of product categorization (higher distinctiveness between products across categories). Therefore, alternatively, we measure width and depth as the number of higher-level categories and the number of SKUs within those higher-level categories respectively([Base Model]Kök et al. 2008). This serves as a robustness check on our empirical results. Specifically, we define $Width_{it}^h$ as the number of distinct higher-level categories offered by merchant i during week t, and $Depth_{it}^h$ as the number of items offered per each higher-level category. We then estimate the following models to test for the relative effects of assortment width and assortment depth on delivery timeliness.

$$DeliveryTime_{ijt} = \beta_1' Width^h{}_{it} + \beta_2' Depth^h{}_{it} + CONTROLS_{ijt} + Merchant_i + Category_j + Week_t + \epsilon_{ijt} \quad (9)$$

$$DelayTime_{ijt} = \gamma_1' Width^h{}_{it} + \gamma_2' Depth^h{}_{it} + CONTROLS_{ijt} + Merchant_i + Category_j + Week_t + \epsilon_{ijt} \quad (10)$$

Table A6: Impact of Assortment Strategy (Category Level) on Order Delivery Timeliness

	(1)	(2)
	DeliveryTime	*DelayTime*
$Width^h$	0.052***	0.087***
	(0.017)	(0.003)
$Depth^h$	0.000	0.001**
	(0.000)	(0.000)
LP	0.005**	-0.006***
	(0.002)	(0.001)
AvgPrice	0.000***	0.000**
	(0.000)	(0.000)
LRS	−0.117***	−0.111***
	(0.020)	(0.023)
log(Sales)	−0.004*	0.002
	(0.002)	(0.003)
Adjusted R^2	0.629	0.301

*Note: *p<0.1; **p<0.05; ***p<0.01, Robust standard errors in parentheses. Estimated with merchant, category, and week fixed effects, N = 64,233.*

The results are presented in Table A6. We observe that, again, assortment width has a negative effect on order delivery performance (β_1' and γ_1' are both positive and significant). Further, a Z-test finds statistical support for $\beta_1' > \beta_2'$ and $\gamma_1' > \gamma_2'$ thereby reaffirming the relatively greater deleterious effects of assortment width on order fulfillment performance.

C.3. Category Distinctiveness - Diversification Analysis.

We once again exploit the fact that our data includes different levels of product categorization. At the lower level of categorization, the degree of distinctiveness between products is smaller compared to the distinctiveness between products at higher level of categorization.

Based on this, we construct a second alternative measure for product assortment expansion, drawing upon the vast literature on diversification. A retailer widening its assortment is essentially diversifying by introducing items in existing categories or adding a new category. When the retailer diversifies through entry into new categories, it is significantly increasing the level of distinctiveness between its categories. However, when the retailer diversifies through adding products in the existing categories, it is still increasing the level of distinctiveness between products but to a smaller extent (as products belonging to the same category are similar).

Similar diversification strategies can be found in the literature on firm diversification (Raghunathan 1995, Dewan et al. 1998), team diversification (Marchetti 2019), and product diversification (Van Herpen and Pieters 2002). Consider, for example, firm diversification strategies. When firms diversify by entering a new sector, they are significantly increasing the distinctiveness between their operating industries. However, when firms diversify by entering a new industry (within the same sector), it is still increasing the level of distinctiveness between its industries but to a smaller extent (industries within a sector are similar). Following this logic and literature, we construct an entropy-based diversification measure (Palepu 1985, Jacquemin and Berry 1979) of product assortment expansion and re-estimate our model presented in equation (1), as follows.

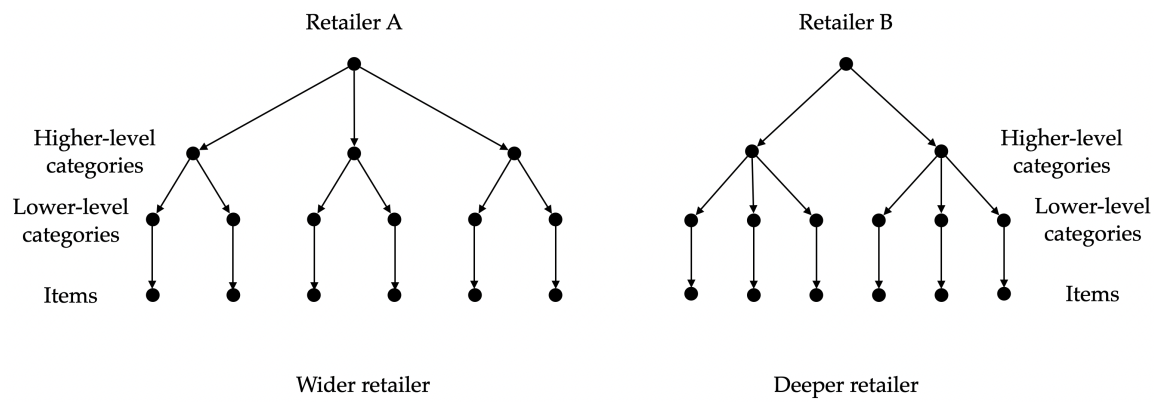
Consider a retailer offering the set of N lower-level categories across the set of M higher-level categories. Further, let us denote by N_j the set of lower-level categories offered by the retailer within the higher-level category j. Therefore, $\Sigma_{j\in N} N_j = N$. We define s_j as the ratio of the number of items offered in lower-level category j to the total number of items offered by the retailer. Further, we define s_{ij} as the ratio of the number of items offered in lower-level category i, higher-level category j, to the total number of items offered by the retailer. We are now in a position to define our entropic measures of assortment width and depth.

$$Width^{ent} = \sum_{j=1}^{M} s_j log(\frac{1}{s_j})$$

$$Depth^{ent} = \sum_{j=1}^{M} \sum_{1=1}^{N_j} s_{ij} log(\frac{s_j}{s_{ij}})$$

To illustrate the above measure, let us consider an example of two retailers A and B. As depicted in Figure A1, suppose retailer A offers six products products across three distinct higher-level categories, and offers two lower-level categories in each higher-level category, and one item in each lower-level category. Compare with retailer B who also offers six products across two categories and offers three categories in each of those two categories, and one item in each category.

Figure A1: Pictorial representation of retailers with wider and deeper assortments.



Evidently, retailer A is offering a more distinctive assortment than B and thus, a relatively wider assortment whereas retailer B offers a relatively deeper assortment. Computing the entropic measures of assortment width and depth for the two retailers, we obtain, ${Width^{ent}}_A = 1.0986$, ${Width^{ent}}_B = 0.6931$, ${Depth^{ent}}_A = 0.6931$, and ${Depth^{ent}}_B = 1.0986$. Expectedly, the entropic measures of assortment width and depth validate and correctly identify retailer A as having a wider assortment while B has a deeper assortment.

We now estimate the following regression model.

$$DeliveryTime_{ijt} = \delta_1' \, {Width^{ent}}_{it} + \delta_2' {Depth^{ent}}_{it} + CONTROLS_{ijt} + Merchant_i + Category_j + Week_t + \epsilon_{ijt} \quad (11)$$

To address endogeneity, we also employ an instrumental variable approach that exploits the assortment offerings of competing merchants as before. The results are presented in Table A7.

We once again find qualitatively identical results. In the 2SLS estimation, both width and depth adversely affect order delivery timeliness supporting Hypothesis 1. A comparison of the coefficients also supports Hypothesis 2, i.e., an increase in assortment width is worse for order delivery timeliness. This analysis supports the assertion that delivery times are more adversely affected when retailers diversify through entry into new categories i.e., when the level of distinctiveness between categories is significantly increased.

Table A7: Impact of Assortment Strategy (Entropic Measures) on Order Delivery Timeliness

	DeliveryTime	
	(1)	(2)
	OLS	2SLS
$Width^{ent}$	0.6432^{***}	0.7002^{***}
	(0.0729)	(0.1750)
$Depth^{ent}$	-0.0558	0.2615^{***}
	(0.0528)	(0.0965)
LP	0.0029	0.0030
	(0.0019)	(0.0020)
$AvgPrice$	0.0001^{***}	0.0000^{***}
	(0.0000)	(0.0000)
LRS	-0.1201^{***}	-0.1149^{***}
	(0.0128)	(0.0131)
$log(Sales)$	0.0076^{***}	-0.0018
	(0.0017)	(0.0021)
Adjusted R^2	0.632	0.646

Note: **p<0.1;* ***p<0.05;* ****p<0.01, Robust standard errors in parentheses. Estimated with merchant, category, and week fixed effects, second column has been estimated with instruments, N = 64,233.*

C.4. Robustness Tests for Moderating Role of LSP Experience.

We now present robustness tests to confirm the moderating impact of logistic service provider experience on the relationship between assortment strategy and order delivery timeliness as presented in Hypothesis 4.

First, we estimate the following joint model which includes all the moderating variables discussed in §3. That is, we estimate the following interaction model given by equation (12).

$$\begin{aligned} DeliveryTime_{ijt} = {} & \gamma_1' Depth_{it} + \gamma_2' Width_{it} + \gamma_3' LSPExperience_{ijt} + \gamma_4' AvgOrderSize_{it} + \\ & \gamma_5' Width_{it} \times LSPExperience_{ijt} + \gamma_6' Depth_{it} \times LSPExperience_{ijt} + \\ & \gamma_7' Width_{it} \times AvgOrderSize_{it} + \gamma_8' Depth_{it} \times AvgOrderSize_{it} + CONTROLS_{ijt} + \\ & Merchant_i + Category_j + Week_t + \gamma_{ijt} \end{aligned} \tag{12}$$

We report our results in Table A8. We find that, to our expectation, the effect sizes of the interaction terms (*Width* $\times$ *LSPExperience* and *Depth* $\times$ *LSPExperience*) are negative (-0.0012 and -0.0000) and significant ($p < 0.05$). Further, comparing the coefficients, we once again find

Table A8: Moderating Impact of Logistic Service Provider Experience on Delivery Timeliness (Joint Model)

	DeliveryTime
Width	0.0493^{***}
	(0.0178)
Depth	0.0155^{*}
	(0.0087)
LSP Experience	0.0259^{***}
	(0.0097)
AvgOrderSize	-0.3873^{***}
	(0.1222)
Width × *LSPExperience*	-0.0012^{**}
	(0.0005)
Depth × *LSPExperience*	-0.0000^{***}
	(0.0000)
Width × *AvgOrderSize*	0.0092^{***}
	(0.0018)
Depth × *AvgOrderSize*	0.0027^{***}
	(0.0006)
Adjusted R^2	0.611

Note: **p<0.1;* ***p<0.05;* ****p<0.01, Robust standard errors in parentheses. Estimated with merchant, category, and week fixed effects, estimated with instruments and all controls, N = 60,511.*

that experienced LSPs are most beneficial for retailers that are expanding their assortment width whereas the benefits of experienced LSPs are much more limited for retailers with deep assortments.

Second, we also consider an alternate measure of LSP experience. We measure experience as the cumulative number of orders in the same category j handled on average across all LSPs employed by merchant i.

The results are reported in Table A9 and we again find consistent results. LSP experience indeed mitigates the negative outcomes of assortment expansions on order delivery timeliness. This is confirmed by the coefficients of (*Width* × *LSPExperience* and *Depth* × *LSPExperience*) being negative (-0.0018 and -0.0005) and significant ($p < 0.05$). Again, the mitigating role of LSP experience is higher for assortment width expansions than for depth expansions.

Table A9: Moderating Impact of Logistic Service Provider Experience (Alternate Measure) on Delivery Timeliness

	DeliveryTime
Width	0.0510^{***}
	(0.0188)
Depth	0.0152^{*}
	(0.0089)
LSP Experience	0.4039^{**}
	(0.1587)
AvgOrderSize	-0.3878^{***}
	(0.1218)
Width × *LSPExperience*	-0.0018^{***}
	(0.0007)
Depth × *LSPExperience*	-0.0005^{**}
	(0.0002)
Width × *AvgOrderSize*	0.0092^{***}
	(0.0018)
Depth × *AvgOrderSize*	0.0027^{***}
	(0.0006)
Adjusted R^2	0.613

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$, *Robust standard errors in parentheses. Estimated with merchant, category, and week fixed effects, estimated with instruments and all controls, N = 60,511.*

Appendix D – Selection Model

As described in §5, we employ the following logit selection model to predict the fraction of orders that a merchant shall offer a promised delivery date in a particular category during a particular week:

$$\log\left(\frac{f_{ijt}}{1-f_{ijt}}\right) = \zeta_1 AvgPrice'_{ijt} + \zeta_2 LRS_{ijt} + \zeta_3 Width_{it} + \zeta_4 Depth_{it} + Merchant_i + Category_j + Week_t + \varepsilon_{ijt} \quad (13)$$

We describe results of the selection model we employ in order to account for the endogenous choice of promising delivery times for a certain fraction of orders by the retailer. We find that the standard deviation of the residuals in our selection model is 0.082, confirming that our model is an accurate predictor of the actual fraction of orders with a promised versus non-promised delivery time. Interestingly, we note that assortment width appears to be a statistically significant and negatively associated predictor of fraction of promised orders. The logistical difficulties in managing wider assortments could potentially be playing a role. The results of the two stage estimation of (4) are reported in Table A11.

Table A10: Selection Model for Fraction of Orders without a Promised Delivery Time, f_{ijt}

	f
LRS	-0.2174^{***} (0.050)
AvgPrice′	0.00002 (0.00002)
Width	0.0118^{*} (0.0057)
Depth	0.00003 (0.0006)
Log Likelihood	−53,899.210
Akaike Inf. Crit.	109,616.400

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$, *Robust standard errors in parentheses. N= 64,233. Estimated with merchant, category, and week fixed effects.*

Appendix E - Parallel Trends Assumption Test

To test for the *parallel trends assumption* we estimate the regression model given by (6) provided in §5. We provide the results of the estimation of (6) in Table A12 (Column 1 for our broader treatment-control groups, and Column 2 for the matched case).

w_t denotes a week variable across the pre-treatment period, and the coefficient of $TREAT \times w_t$ allows us to verify that the parallel trends assumption is satisfied. Since the coefficient cannot be statistically differentiated from 0, i.e., the coefficient is statistically insignificant in our regression, thus we can claim that the parallel trends assumption is satisfied. Therefore, accounting for controls, the trends in our treatment and control groups are identical before the treatment.

Table A11: Effect of Order Delivery Timeliness on $log(Sales)$

	$log(Sales)$
$f_{ij(t-1)} DeliveryTime'_{ij(t-1)}$	-0.396***
$f_{ij(t-2)} DeliveryTime'_{ij(t-2)}$	-0.499***
$f_{ij(t-3)} DeliveryTime'_{ij(t-3)}$	-0.270***
$f_{ij(t-4)} DeliveryTime'_{ij(t-4)}$	-0.0536**
$f_{ij(t-5)} DeliveryTime'_{ij(t-5)}$	-0.193***
$f_{ij(t-6)} DeliveryTime'_{ij(t-6)}$	-0.193***
$f_{ij(t-7)} DeliveryTime'_{ij(t-7)}$	-0.124***
$f_{ij(t-8)} DeliveryTime'_{ij(t-8)}$	-0.136***
$(1 - f_{ij(t-1)}) DelayTime'_{ij(t-1)}$	-1.031***
$(1 - f_{ij(t-2)}) DelayTime'_{ij(t-2)}$	-0.809***
$(1 - f_{ij(t-3)}) DelayTime'_{ij(t-3)}$	-0.751***
$(1 - f_{ij(t-4)}) DelayTime'_{ij(t-4)}$	-0.714***
$(1 - f_{ij(t-5)}) DelayTime'_{ij(t-5)}$	-0.276**
$(1 - f_{ij(t-6)}) DelayTime'_{ij(t-6)}$	-0.121
$(1 - f_{ij(t-7)}) DelayTime'_{ij(t-7)}$	-0.0537
$(1 - f_{ij(t-8)}) DelayTime'_{ij(t-8)}$	0.0920
$f_{ij^{-1}(t-1)} DeliveryTime'_{ij^{-1}(t-1)}$	-0.414***
$f_{ij^{-1}(t-2)} DeliveryTime'_{ij^{-1}(t-2)}$	-0.395***
$f_{ij^{-1}(t-3)} DeliveryTime'_{ij^{-1}(t-3)}$	-0.271***
$f_{ij^{-1}(t-4)} DeliveryTime'_{ij^{-1}(t-4)}$	-0.0831*
$f_{ij^{-1}(t-5)} DeliveryTime'_{ij^{-1}(t-5)}$	-0.233***
$f_{ij^{-1}(t-6)} DeliveryTime'_{ij^{-1}(t-6)}$	-0.253***
$f_{ij^{-1}(t-7)} DeliveryTime'_{ij^{-1}(t-7)}$	-0.238***
$f_{ij^{-1}(t-8)} DeliveryTime'_{ij^{-1}(t-8)}$	-0.115**
$(1 - f_{ij^{-1}(t-1)}) DelayTime'_{ij^{-1}(t-1)}$	-0.743**
$(1 - f_{ij^{-1}(t-2)}) DelayTime'_{ij^{-1}(t-2)}$	-0.149*
$(1 - f_{ij^{-1}(t-3)}) DelayTime'_{ij^{-1}(t-3)}$	-0.278**
$(1 - f_{ij^{-1}(t-4)}) DelayTime'_{ij^{-1}(t-4)}$	-0.359***
$(1 - f_{ij^{-1}(t-5)}) DelayTime'_{ij^{-1}(t-5)}$	-0.287***
$(1 - f_{ij^{-1}(t-6)}) DelayTime'_{ij^{-1}(t-6)}$	-0.160**
$(1 - f_{ij^{-1}(t-7)}) DelayTime'_{ij^{-1}(t-7)}$	-0.290***
$(1 - f_{ij^{-1}(t-8)}) DelayTime'_{ij^{-1}(t-8)}$	-0.0709
Adjusted R^2	0.765

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$, *Estimated with merchant, category, and week fixed effects, N = 64,233.*

Appendix F - Bayesian Structural Time Series Model

The BSTS approach constructs the counterfactual time series by considering the pre-treatment time-series characteristics of the outcome variable itself, and the evolution of appropriately chosen control time-series' that were unaffected by the intervention. The BSTS approach explicitly models the temporal evolution of the variable of interest using a state-space model. For

Table A12: Parallel Trends Assumption Test

	(1) *Unmatched*	(2) *Matched Sample*
$TREAT \times w_t$	0.0223 (0.0160)	0.0272 (0.0180)
Fixed Effects		
Merchant	*Yes*	*Yes*
category	*Yes*	*Yes*
Adjusted R^2	0.644	0.653

Note: **p<0.1;* ***p<0.05;* ****p<0.01, Robust standard errors in parentheses. N = 64,233*

a comprehensive comparative discussion of the general approach and its extensions, we refer the reader to Brodersen et al. (2015). We consider the following state-space model for a merchant i and category j in our treatment group:

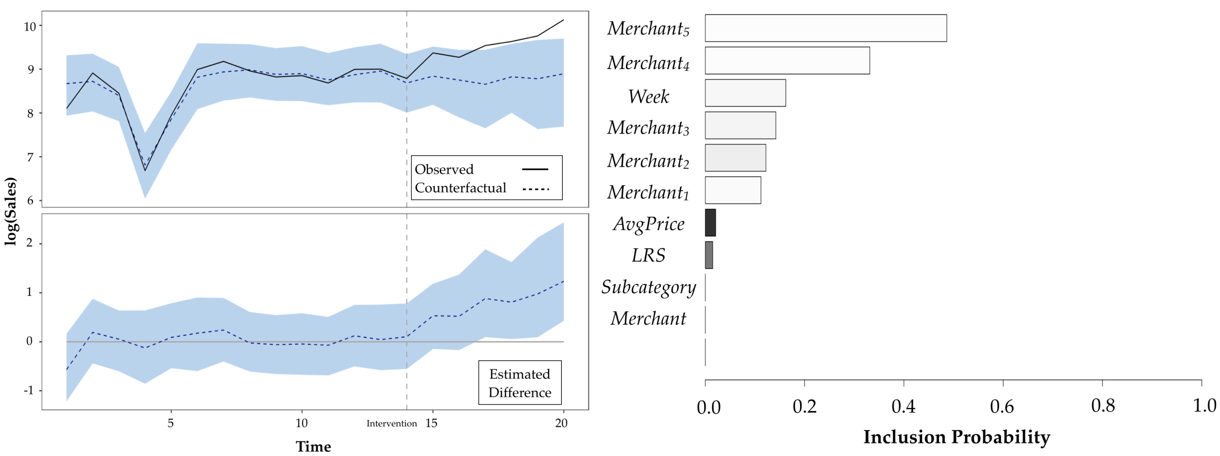
$$Sales_{ijt} = \mu_{ijt} + Z_{ijt} + \epsilon_{ijt} \qquad \epsilon_{ijt} \sim N(0, \sigma_\epsilon^2),$$
$$\mu_{ijt+1} = \mu_{ijt} + \eta_{ijt} \qquad \eta_{ijt} \sim N(0, \sigma_\eta^2),$$
$$Z_{ijt} = \beta_{ij} X_{ijt},$$

where $Sales_{ijt}$ denotes the sales time-series of the treated firms, μ_{ijt} is a local level term that permits stochasticity in the outcome time-series, and Z_{ijt} models the linear relationship between the control series and co-variates in X_{ijt} and the treatment series. Our application of the BSTS approach to our setting then involves the following two steps:

Step 1. Our treatment group T comprises of those categories offered by merchants that are delivered by Cainiao as the logistic provider. Further, we remove from our analysis those merchant-category pairs that do not span the entire time duration of our dataset, this leaves us with 75 merchant-category pairs in our treatment group. We employ as controls the average price, and the sales in the same category by the five most similar merchants, as defined by the assortment they offer, who do not employ Cainiao. We also include the week variable as a covariate in X_{ijt} to allow for linear trends in our model. We fit our model using the CausalImpact package developed by Google and Brodersen et al. (2015), employing a spike-and-slab prior on the coefficients in β, and Bayesian model averaging to construct the counter-factual.

Step 2. The BSTS method outlined in Brodersen et al. (2015) can be directly applied to only a single treated time-series. Therefore, we perform Step 1 for each merchant-category pair separately, and perform a meta-analysis in our second-step to pool the estimates. We note that Schmitt et al. (2018) also adopt a similar extension of the BSTS approach. We aggregate the individual estimates,

Figure A2: Bayesian Structural Time Series Model



(a) Observed and counterfactual time series of $log(Sales)$ for merchant 50 and category 362

(b) Coefficients on the control time-series after Bayesian model averaging

$\hat{\Delta}_{ij}$, of the percentage change in sales after the intervention across the treated merchant-category pairs using the inverse-variance weighting method to obtain a pooled estimate of the effect,

$$\hat{\Delta} = \frac{\sum\limits_{(i,j)\in T} \hat{\Delta}_{ij}/\sigma_{ij}^2}{\sum\limits_{(i,j)\in T} 1/\sigma_{ij}^2} \tag{14}$$

where σ_{ij}^2 is the variance in the estimated percentage change for the merchant i and category j in our treatment group.

We find that the opening of new warehouses resulted in a statistically significant 11.2% increase in sales for the merchant-category pairs in our treatment group. In Figure A2, we provide an illustrative example of a merchant-category pair depicting the observed and counter-factual time-series of the sales. Therefore, we recover qualitatively similar estimates suggesting that the opening of new warehouses providing the ability of faster order deliveries did indeed result in an increase in sales.