

Online Appendix for Linear Estimation of Aggregate Dynamic Discrete Demand for Durable Goods: Overcoming the Curse of Dimensionality

February 2019

1 Numerical Simulation

In order to determine how well our estimator performs in small samples, we run several simulations that vary the number of products, the number of markets, the number of time periods and whether the data generating process originated from a type 1 extreme value distribution or a GEV distribution or a finite mixture model.

1.1 Logit Model

We first discuss the data generating process associated with the logit model. The consumer's flow utility function follows the specification in §2.1. When consumer i purchases product j in period t , he receives the following flow utility in period t ,

$$u_{ijt} = f(x_{jt}, \xi_{jt}) - \alpha p_{jt} + \varepsilon_{ijt} \equiv f(x_{jt}, \xi_{jt}) - 0.5p_{jt} + \varepsilon_{ijt},$$

and receives $f(x_{jt}, \xi_{jt})$ as flow utility in each period post purchase in period t . In the simulation we let

$$f(x_{jt}, \xi_{jt}) = x'_{jt}\gamma + \delta_j + \xi_{jt} = x'_{jt}0 + 0.75 + \xi_{jt},$$

for any product j . So $\alpha = 0.5$, $\gamma = 0$ and $\delta_j = 0.75$ for any product j . Products in the simulation are differentiated by the observed price, p_{jt} , and unobserved characteristics, ξ_{jt} . The discount factor β is set to 0.80. We maintain the independence and logit specification about ε_{ijt} , i.e. Assumption 3.

We next describe the data generation process of price and the unobserved product characteristics. We specifically account for correlation between ξ_{jt} and p_{jt} . Such a formulation is motivated by the price endogeneity problem researchers face when employing aggregate data, where firms can observe ξ_{jt} and then set prices optimally. We use a reduced form price model to characterize this dependence. Specifically,

$$p_{jt} = c_j + MC_{jt} + \omega_{jt} \quad \text{and} \quad \xi_{jt} = \phi_j \xi_{j,t-1} + \nu_{jt},$$

where $(\omega_{jt}, \nu_{jt})'$ is *iid* across products and time periods, and follows a normal distribution,

$$\begin{pmatrix} \omega_{jt} \\ \nu_{jt} \end{pmatrix} \sim \mathcal{N} \left(0, \begin{pmatrix} \sigma_p^2 & \rho \sigma_\nu \sigma_p \\ & \sigma_\nu^2 \end{pmatrix} \right).$$

Here MC_{jt} denotes the marginal cost of product j at time t . MC_{jt} is independent of $(\omega_{j\tilde{t}}, \nu_{j\tilde{t}})'$ for any period t and \tilde{t} . Specifically, MC_{jt} takes the form

$$MC_{jt} = \psi_j MC_{j,t-1}$$

We will use MC_{jt} as the instrumental variable in both estimation steps 1 and 2 outlined in §4.1.

In our simulations the maximum number of products is 5, and we assign the following parameter values. We let $(c_1, \dots, c_5) = (1, 2.5, 3.5, 4.5, 5.5)$ and $(\psi_1, \dots, \psi_5) = (0.98, 0.92, 0.88, 0.84, 0.80)$. For the initial state of MC_{j0} , we let $(MC_{10}, \dots, MC_{50}) = (15, 14.5, 14, 13.5, 13)$. Such specification ensures that product marginal cost, MC_{jt} , has a declining trajectory, which is consistent with durable goods models.

In addition, we let $\phi_j = 0.25$ for any product j .¹ Let $\sigma_p = 0.5$, $\rho = 0.25$, and $\sigma_\nu = 0.1$. Since ξ_{jt} is a stationary AR(1) process, it is easy to see that $\sigma^2 = \text{Var}(\xi_{jt}) = 0.1^2 / (1 - 0.25^2)$, that is $\sigma \approx 0.1033$. Moreover, $\text{corr}(\xi_{jt}, p_{jt}) = \rho$ by serial independence of both ω_{jt} and ν_{jt} .

In Fig. 1, we present prices and the outside option's market share in order to illustrate that the data generation process (DGP) is consistent with a durable goods setting. Note the declining prices and decreasing market share of the outside option in Fig. 1.

Suppose for J products and one market we have simulated panel data (s_t, p_t, MC_t, ξ_t) for T periods. We first estimate α with an instrumental variable regression. We use the marginal cost variable above as a price instrument. Given the estimates of α we have estimates of y_t

¹We also performed simulations when ξ_{jt} has no serial correlation, i.e. $\phi_j = 0$. Results are available upon request.

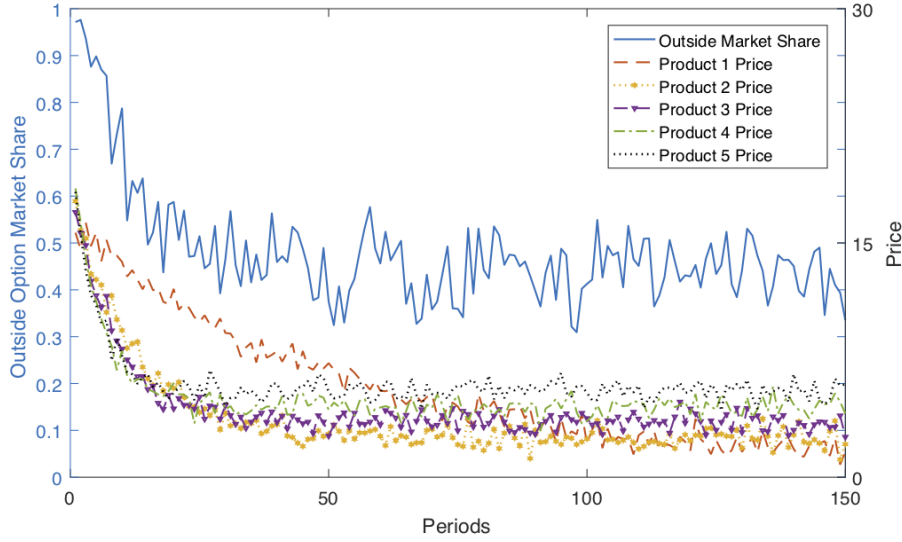


Figure 1: Monte Carlo Prices and Outside Market Share ($J = 5$)

and w_t . We then estimate β using two stage least squares as discussed in §5.1.2, using the demeaned price instrument as the instrument. Once β is estimated, we can estimate γ by multiplying the estimate of $\tilde{\gamma}$ with $1 - \hat{\beta}$, if we included other observed product characteristics to estimate. Yet, since the DGP only consists of a constant term, we estimate the constant using step 3 in section 5.1.3. The estimation of $\text{Var}(\xi_{jt})$ follows the steps in the previous section. We also estimate $E(\xi_{jt} | p_{jt})$ using step 4 in §5.2.1 to recover ρ and σ .

Each set of simulations we analyze was based on 250 replications. We also analyze sets with varying number of markets (1 and 10), time periods (150, or 300) and the number of J .

The first set of simulations in Table 1 and 2 illustrate how well and precise our methodology is able to identify the data generating process—including the discount factor. Furthermore, if the discount factor is known (or assumed), the results exhibit less small sample bias and more precision, particularly for the parameters that include the discount factor in estimation, γ , σ , ρ , and ϕ . Specifically, we determine that estimation of ρ is quite challenging in practice and requires a sizeable amount of data and products to precisely estimate when the discount factor is estimated. This again is from the fact that

$$g_{3jt}(\theta) = z_{\rho,j,t}(p_{jt})r_{jt}$$

$$r_{jt} = (1 - \beta)(y_{j,t} + \beta w_{j,t+1}) - (1 - \beta)\delta_j - \tilde{\rho}_j(\tilde{p}_{j,t} - \beta \tilde{p}_{j,t+1})$$

Table 1: Monte Carlo Simulation Results: 10 Markets and 150 Periods

DGP: 10 Markets, $T = 150$						
	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7323 (0.0083)	-0.5003 (0.0075)	0.0963 (0.0063)	0.1405 (0.0637)	0.2609 (0.0686)	0.8153 (0.0116)
$J = 3$	0.7381 (0.0082)	-0.5003 (0.0075)	0.0941 (0.0088)	0.1199 (0.0655)	0.2538 (0.0480)	0.8192 (0.0161)
$J = 4$	0.7431 (0.0088)	-0.5002 (0.0077)	0.0911 (0.0102)	0.1403 (0.0600)	0.2473 (0.0435)	0.8253 (0.0190)
$J = 5$	0.7463 (0.0109)	-0.5001 (0.0074)	0.0913 (0.0121)	0.1736 (0.0582)	0.2422 (0.0436)	0.8257 (0.0222)

	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7391 (0.0093)	-0.5003 (0.0075)	0.1053 (0.0025)	0.2086 (0.0692)	0.2555 (0.0684)	–
$J = 3$	0.7399 (0.0089)	-0.5003 (0.0075)	0.1050 (0.0017)	0.1902 (0.0560)	0.2496 (0.0472)	–
$J = 4$	0.7411 (0.0089)	-0.5002 (0.0077)	0.1054 (0.0017)	0.2126 (0.0490)	0.2426 (0.0424)	–
$J = 5$	0.7417 (0.0098)	-0.5001 (0.0084)	0.1057 (0.0015)	0.2318 (0.0428)	0.2381 (0.0425)	–

Mean and standard deviation for 250 simulations.

Table 2: Monte Carlo Simulation Results: 10 Markets and 300 Periods

DGP: 10 Markets, $T = 300$						
	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7494 (0.0058)	-0.5003 (0.0074)	0.1065 (0.0062)	0.2171 (0.0408)	0.2360 (0.0313)	0.7972 (0.0116)
$J = 3$	0.7457 (0.0086)	-0.5003 (0.0076)	0.1041 (0.0080)	0.2382 (0.0371)	0.2388 (0.0289)	0.8028 (0.0145)
$J = 4$	0.7473 (0.0114)	-0.5002 (0.0079)	0.1018 (0.0090)	0.2546 (0.0321)	0.2363 (0.0359)	0.8078 (0.0168)
$J = 5$	0.7483 (0.0134)	-0.5001 (0.0087)	0.1015 (0.0096)	0.2664 (0.0333)	0.2345 (0.0440)	0.8087 (0.0176)

	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7496 (0.0057)	-0.5003 (0.0074)	0.1050 (0.0018)	0.2091 (0.0514)	0.2365 (0.0318)	–
$J = 3$	0.7451 (0.0062)	-0.5003 (0.0076)	0.1057 (0.0014)	0.2438 (0.0392)	0.2383 (0.0285)	–
$J = 4$	0.7443 (0.0070)	-0.5002 (0.0079)	0.1061 (0.0013)	0.2653 (0.0340)	0.2354 (0.0354)	–
$J = 5$	0.7443 (0.0083)	-0.5001 (0.0087)	0.1064 (0.0013)	0.2661 (0.0318)	0.2337 (0.0436)	–

Mean and standard deviation for 250 simulations.

is impacted by the discount factor. Thus, any bias associated with the discount factor will propagate through and into the estimation of the correlation parameter. Lastly, as is the case in much of the static choice literature where the variance covariance matrix is estimated, it is known that sizeable amounts of data are required to precisely estimate the parameter. This is made more clear with our second set of simulations which increases the time duration to 300 periods from 150. This increase doubles the amount of data and provides improvement in the estimation of ρ and the discount factor.

Finally, Table 3 and 4 present the analysis where only 1 market is employed and T equals 150 or 300 periods. The results are similar to the set of simulations which employ 10 markets, but with less precision—most notably for ρ and ϕ .

Table 3: Monte Carlo Simulation Results: 1 Market and 150 Periods

DGP: 1 Market, $T = 150$						
	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7339 (0.0265)	-0.5041 (0.0249)	0.0990 (0.0402)	0.1462 (0.2087)	0.2471 (0.1975)	0.8167 (0.0386)
$J = 3$	0.7419 (0.0247)	-0.5048 (0.0234)	0.0971 (0.0314)	0.1301 (0.2272)	0.2380 (0.1293)	0.8190 (0.0503)
$J = 4$	0.7447 (0.0264)	-0.5040 (0.0233)	0.0922 (0.0607)	0.1374 (0.1906)	0.2273 (0.1181)	0.8260 (0.0607)
$J = 5$	0.7487 (0.0331)	-0.5034 (0.0250)	0.0893 (0.0369)	0.1678 (0.1825)	0.2220 (0.1246)	0.8324 (0.0684)

	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7439 (0.0300)	-0.5041 (0.0249)	0.1075 (0.0091)	0.2211 (0.2205)	0.2407 (0.1951)	–
$J = 3$	0.7462 (0.0273)	-0.5084 (0.0234)	0.1077 (0.0143)	0.2088 (0.1964)	0.2364 (0.1304)	–
$J = 4$	0.7456 (0.0268)	-0.5040 (0.0233)	0.1069 (0.0068)	0.2178 (0.1774)	0.2227 (0.1154)	–
$J = 5$	0.7462 (0.0290)	-0.5034 (0.0250)	0.1078 (0.0147)	0.2464 (0.1597)	0.2168 (0.1205)	–

Mean and standard deviation for 250 simulations.

Table 4: Monte Carlo Simulation Results: 1 Market and 300 Periods

DGP: 1 Market, $T = 300$						
	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7510 (0.0197)	-0.5041 (0.0248)	0.1083 (0.0293)	0.2265 (0.1653)	0.2256 (0.0966)	0.7994 (0.0344)
$J = 3$	0.7486 (0.0267)	-0.5048 (0.0238)	0.1047 (0.0246)	0.2424 (0.1462)	0.2354 (0.0960)	0.8040 (0.0449)
$J = 4$	0.7503 (0.0337)	-0.5041 (0.0239)	0.1008 (0.0276)	0.2543 (0.1084)	0.2381 (0.1094)	0.8109 (0.0512)
$J = 5$	0.7515 (0.0416)	-0.5035 (0.0259)	0.1011 (0.0317)	0.2697 (0.0984)	0.2412 (0.1322)	0.8111 (0.0578)

	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7527 (0.0191)	-0.5041 (0.0248)	0.1069 (0.0098)	0.2209 (0.1841)	0.2250 (0.0964)	–
$J = 3$	0.7490 (0.0206)	-0.5048 (0.0238)	0.1069 (0.0050)	0.2566 (0.1335)	0.2350 (0.0951)	–
$J = 4$	0.7477 (0.0223)	-0.5041 (0.0239)	0.1070 (0.0040)	0.2705 (0.1097)	0.2364 (0.1068)	–
$J = 5$	0.7483 (0.0253)	-0.5035 (0.0259)	0.1071 (0.0036)	0.2828 (0.0927)	0.2400 (0.1303)	–

Mean and standard deviation for 250 simulations.

Table 5: Nested Logit Monte Carlo Simulation Results: 10 Markets and 150 Periods

DGP: 10 Markets, $T = 150$							
	$\delta = 0.75$	$\alpha = -0.5$	$\zeta = 0.8$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 3$	0.7339 (0.0352)	-0.5018 (0.0274)	0.7920 (0.1021)	0.0954 (0.0215)	0.1358 (0.0967)	0.2326 (0.0567)	0.8176 (0.0319)
$J = 4$	0.7367 (0.0381)	-0.5015 (0.0266)	0.7967 (0.0674)	0.0917 (0.0183)	0.1503 (0.1086)	0.2447 (0.0516)	0.8245 (0.0288)
$J = 5$	0.7405 (0.0243)	-0.5005 (0.0156)	0.8005 (0.0432)	0.0902 (0.0135)	0.1730 (0.0899)	0.2411 (0.0439)	0.8261 (0.0262)

	$\delta = 0.75$	$\alpha = -0.5$	$\zeta = 0.8$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 3$	0.7416 (0.0387)	-0.5018 (0.0274)	0.7920 (0.1021)	0.1039 (0.0057)	0.2029 (0.1625)	0.2261 (0.0527)	-
$J = 4$	0.7415 (0.0403)	-0.5015 (0.0266)	0.7967 (0.0674)	0.1041 (0.0051)	0.2186 (0.1476)	0.2384 (0.0477)	-
$J = 5$	0.7401 (0.0263)	-0.5005 (0.0156)	0.8005 (0.0431)	0.1038 (0.0036)	0.2327 (0.0810)	0.2412 (0.0423)	-

Mean and standard deviation for 250 simulations.

Table 6: Nested Logit Monte Carlo Simulation Results: 10 Markets and 300 Periods

DGP: 10 Markets, $T = 300$							
	$\delta = 0.75$	$\alpha = -0.5$	$\zeta = 0.8$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 3$	0.7449 (0.0192)	-0.5008 (0.0156)	0.7974 (0.0426)	0.1024 (0.0115)	0.2390 (0.0515)	0.2300 (0.0415)	0.8019 (0.0204)
$J = 4$	0.7435 (0.0351)	-0.5003 (0.0241)	0.7991 (0.0477)	0.1011 (0.0151)	0.2507 (0.0769)	0.2386 (0.0322)	0.8058 (0.0247)
$J = 5$	0.7460 (0.0284)	-0.5003 (0.0197)	0.7998 (0.0442)	0.1002 (0.0112)	0.2609 (0.0773)	0.2420 (0.0298)	0.8071 (0.0223)

	$\delta = 0.75$	$\alpha = -0.5$	$\zeta = 0.8$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 3$	0.7455 (0.0157)	-0.5008 (0.0156)	0.7974 (0.0426)	0.1032 (0.0024)	0.2461 (0.0886)	0.2288 (0.0441)	-
$J = 4$	0.7440 (0.0238)	-0.5003 (0.0241)	0.7991 (0.0477)	0.1039 (0.0037)	0.2584 (0.0978)	0.2378 (0.0327)	-
$J = 5$	0.7437 (0.0273)	-0.5003 (0.0197)	0.7998 (0.0442)	0.1040 (0.0039)	0.2697 (0.0764)	0.2413 (0.0292)	-

Mean and standard deviation for 250 simulations.

1.2 Nested Logit Model

Next, we present the result of several Monte Carlo simulations with a nested logit data generating process. Particularly, we analyze the case where there of 3-5 products with product one relegated to one nest and all other products to a second nest. The within nested correlation for product one is normalized to 1 with the second nest taking the value of 0.80. The remaining data generating processes follows exactly as above in the simply MNL case.

We present the same variation of simulations as in the Logit case. The tables below illustrate that our estimator is able to precisely estimate the model primitives associated with the nested logit model. Finally, the presence of multimarkets aids in the recovery of model parameters.

1.3 Heterogeneous Logit Model

Lastly, we present the result of several Monte Carlo simulations where the DGP includes consumer heterogeneity in price, but we estimate a multinomial logit model. Doing so

Table 7: Nested Logit Monte Carlo Simulation Results: 1 Markets and 150 Periods

DGP: 1 Market, $T = 150$							
	$\delta = 0.75$	$\alpha = -0.5$	$\zeta = 0.8$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 3$	1.0176 (4.7648)	-0.5113 (0.0926)	0.7635 (0.3300)	0.3646 (3.4888)	0.2148 (0.3081)	0.2253 (0.1907)	1.0357 (3.7963)
$J = 4$	0.6528 (0.5058)	-0.5057 (0.0838)	0.7943 (0.2161)	0.1424 (0.2823)	0.1982 (0.2974)	0.2718 (0.1648)	0.7696 (0.3559)
$J = 5$	0.7338 (0.1032)	-0.5020 (0.0533)	0.8038 (0.1406)	0.1011 (0.0577)	0.1907 (0.2524)	0.2700 (0.1343)	0.8133 (0.1044)

	$\delta = 0.75$	$\alpha = -0.5$	$\zeta = 0.8$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 3$	0.7561 (0.1276)	-0.5113 (0.0926)	0.7635 (0.3300)	0.1113 (0.0247)	0.2533 (0.4121)	0.2134 (0.1427)	-
$J = 4$	0.7469 (0.1241)	-0.5057 (0.0838)	0.7943 (0.2161)	0.1117 (0.0219)	0.2374 (0.3609)	0.2576 (0.1211)	-
$J = 5$	0.7420 (0.0852)	-0.5020 (0.0533)	0.8038 (0.1406)	0.1089 (0.0150)	0.2416 (0.2417)	0.2694 (0.1220)	-

Mean and standard deviation for 250 simulations.

Table 8: Nested Logit Monte Carlo Simulation Results: 1 Markets and 300 Periods

DGP: 1 Market, $T = 300$							
	$\delta = 0.75$	$\alpha = -0.5$	$\zeta = 0.8$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 3$	0.7393 (0.0703)	-0.5073 (0.0523)	0.7828 (0.1284)	0.1092 (0.0437)	0.2667 (0.2042)	0.2199 (0.1165)	0.7931 (0.0751)
$J = 4$	0.7200 (0.1497)	-0.5038 (0.0803)	0.7933 (0.1558)	0.1170 (0.0700)	0.2545 (0.2325)	0.2460 (0.1100)	0.7917 (0.1058)
$J = 5$	0.7410 (0.1155)	-0.5023 (0.0666)	0.8030 (0.1465)	0.1129 (0.0495)	0.2542 (0.2112)	0.3188 (0.1143)	0.7976 (0.0827)

	$\delta = 0.75$	$\alpha = -0.5$	$\zeta = 0.8$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 3$	0.7526 (0.0510)	-0.5073 (0.0523)	0.7828 (0.1284)	0.1050 (0.0083)	0.2475 (0.2563)	0.2227 (0.1171)	-
$J = 4$	0.7481 (0.0927)	-0.5038 (0.0803)	0.7933 (0.1558)	0.1108 (0.0168)	0.2481 (0.2835)	0.2757 (0.1040)	-
$J = 5$	0.7456 (0.0913)	-0.5023 (0.0666)	0.8030 (0.1456)	0.1122 (0.0177)	0.2612 (0.2126)	0.3180 (0.1091)	-

Mean and standard deviation for 250 simulations.

allows us to determine how model primitives are impacted from this model misspecification. These sets of Monte Carlo studies differ from the above in that the number of simulations run is 100 vs 250 and the number of markets is equal to 1. This change is due to the computational complexity and the time it takes to generate the data. That said, the process does follow the above multinomial logit data generating process with the exception that there are three different consumer types rather than one. The three consumers have price preference parameters equal to $\alpha_1 = -0.4$, $\alpha_2 = -0.5$, $\alpha_3 = -0.6$. The initial weights for each of these consumers in period 0 takes four different parameterizations in order to capture varying degrees of consumer heterogeneity, with case (1) the most heterogeneous and (4) being no heterogeneity.

(1) $\omega_{1,0} = 0.33$, $\omega_{2,0} = 0.34$, $\omega_{3,0} = 0.33$

(2) $\omega_{1,0} = 0.20$, $\omega_{2,0} = 0.60$, $\omega_{3,0} = 0.20$

(3) $\omega_{1,0} = 0.10$, $\omega_{2,0} = 0.80$, $\omega_{3,0} = 0.10$

(4) $\omega_{1,0} = 0.00$, $\omega_{2,0} = 1.00$, $\omega_{3,0} = 0.00$

Below we present four different tables, one for each of the above cases along with varying the number of product from 2 to 5. Within each table, we present the results for all the

Table 9: Heterogeneous Monte Carlo Simulation Results Case (1): 1 Market and 300 Periods

DGP: 1 Market, $T = 300$						
	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7375 (0.0176)	-0.5573 (0.0278)	0.0850 (0.0149)	0.2482 (0.2398)	0.1741 (0.0838)	0.8344 (0.0291)
$J = 3$	0.7402 (0.0200)	-0.5604 (0.0255)	0.0794 (0.0205)	0.2890 (0.1878)	0.1425 (0.0907)	0.8459 (0.0389)
$J = 4$	0.7449 (0.0233)	-0.5602 (0.0245)	0.0727 (0.0214)	0.2277 (0.2052)	0.0936 (0.0971)	0.8592 (0.0411)
$J = 5$	0.7462 (0.0290)	-0.5583 (0.0273)	0.0715 (0.0232)	0.2710 (0.1674)	0.0496 (0.1135)	0.8619 (0.0439)

	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7423 (0.0213)	-0.5573 (0.0278)	0.1028 (0.0050)	0.3534 (0.1845)	0.1681 (0.0822)	–
$J = 3$	0.7342 (0.0211)	-0.5604 (0.0255)	0.1029 (0.0033)	0.3670 (0.1976)	0.1327 (0.0870)	–
$J = 4$	0.7290 (0.0223)	-0.5602 (0.0245)	0.1033 (0.0025)	0.3715 (0.2036)	0.0828 (0.0905)	–
$J = 5$	0.7239 (0.0268)	-0.5583 (0.0273)	0.1034 (0.0024)	0.3678 (0.0931)	0.0435 (0.1083)	–

Mean and standard deviation for 100 simulations.

Table 10: Heterogeneous Monte Carlo Simulation Results Case (2): 1 Market and 300 Periods

DGP: 1 Market, $T = 300$						
	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7337 (0.0182)	-0.5464 (0.0276)	0.0872 (0.0152)	0.1913 (0.2566)	0.1547 (0.0881)	0.8298 (0.0300)
$J = 3$	0.7349 (0.0211)	-0.5490 (0.0255)	0.0810 (0.0208)	0.2278 (0.2159)	0.1161 (0.0911)	0.8425 (0.0395)
$J = 4$	0.7394 (0.0246)	-0.5486 (0.0242)	0.0737 (0.0217)	0.2045 (0.1532)	0.0583 (0.0964)	0.8573 (0.0417)
$J = 5$	0.7401 (0.0305)	-0.5464 (0.0271)	0.0721 (0.0234)	0.2364 (0.1349)	0.0042 (0.1066)	0.8610 (0.0441)

	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7373 (0.0211)	-0.5464 (0.0276)	0.1026 (0.0049)	0.3218 (0.1837)	0.1493 (0.0829)	–
$J = 3$	0.7280 (0.0210)	-0.5490 (0.0255)	0.1028 (0.0033)	0.3262 (0.1736)	0.1082 (0.0889)	–
$J = 4$	0.7218 (0.0221)	-0.5486 (0.0242)	0.1033 (0.0025)	0.3278 (0.1908)	0.0507 (0.0916)	–
$J = 5$	0.7154 (0.0267)	-0.5464 (0.0271)	0.1036 (0.0025)	0.3331 (0.0945)	0.0023 (0.1043)	–

Mean and standard deviation for 100 simulations.

model parameters. First, our method does a fair job at recovering consumer preferences for varying degrees of heterogeneity. Naturally, as the degree of heterogeneity decreases the precision and lack of bias increases. However, recovering parameters associated with the unobservables is quite difficult, particularly when J increases and even with modest amounts of consumer heterogeneity.

Table 11: Heterogeneous Monte Carlo Simulation Results Case (3): 1 Market and 300 Periods

DGP: 1 Market, $T = 300$						
	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7294 (0.1900)	-0.5333 (0.0272)	0.0897 (0.0157)	0.1342 (0.2514)	0.1264 (0.0838)	0.8245 (0.0311)
$J = 3$	0.7289 (0.0224)	-0.5352 (0.0253)	0.0828 (0.0213)	0.2085 (0.1975)	0.0790 (0.0916)	0.8459 (0.0403)
$J = 4$	0.7332 (0.0262)	-0.5345 (0.0240)	0.0747 (0.0221)	0.1588 (0.1936)	0.0150 (0.0941)	0.8556 (0.0424)
$J = 5$	0.7336 (0.0323)	-0.5319 (0.0269)	0.0726 (0.0236)	0.1909 (0.1207)	-0.0429 (0.1031)	0.8607 (0.0444)

	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7318 (0.0209)	-0.5333 (0.0272)	0.1024 (0.0049)	0.2569 (0.2332)	0.1219 (0.0839)	–
$J = 3$	0.7211 (0.0210)	-0.5353 (0.0253)	0.1027 (0.0032)	0.2776 (0.1979)	0.0748 (0.0893)	–
$J = 4$	0.7136 (0.0219)	-0.5345 (0.0240)	0.1034 (0.0025)	0.2691 (0.1708)	0.0126 (0.0900)	–
$J = 5$	0.7058 (0.0265)	-0.5319 (0.0269)	0.1040 (0.0025)	0.2931 (0.0977)	-0.0438 (0.0988)	–

Mean and standard deviation for 100 simulations.

Table 12: Heterogeneous Monte Carlo Simulation Results Case (4): 1 Market and 300 Periods

DGP: 1 Market, $T = 300$						
	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7509 (0.0216)	-0.5034 (0.0260)	0.1055 (0.0179)	0.2241 (0.1685)	0.2262 (0.0888)	0.7943 (0.0361)
$J = 3$	0.7491 (0.0271)	-0.5057 (0.0246)	0.1025 (0.0240)	0.2542 (0.1299)	0.2449 (0.0975)	0.8012 (0.0459)
$J = 4$	0.7543 (0.0322)	-0.5065 (0.0233)	0.0964 (0.0253)	0.2547 (0.1141)	0.2429 (0.1069)	0.8134 (0.0494)
$J = 5$	0.7562 (0.0402)	-0.5068 (0.0263)	0.0969 (0.0288)	0.2762 (0.1009)	0.2519 (0.1312)	0.8128 (0.0548)

	$\delta = 0.75$	$\alpha = -0.5$	$\sigma = 0.1033$	$\rho = 0.25$	$\phi = 0.25$	$\beta = 0.8$
$J = 2$	0.7528 (0.0203)	-0.5034 (0.0260)	0.1028 (0.0051)	0.2138 (0.1941)	0.2261 (0.0889)	–
$J = 3$	0.7503 (0.0206)	-0.5057 (0.0246)	0.1031 (0.0034)	0.2543 (0.1377)	0.2441 (0.0957)	–
$J = 4$	0.7508 (0.0216)	-0.5065 (0.0233)	0.1034 (0.0026)	0.2719 (0.1217)	0.2411 (0.1045)	–
$J = 5$	0.7520 (0.0261)	-0.5068 (0.0263)	0.1034 (0.0024)	0.2896 (0.0964)	0.2512 (0.1302)	–

Mean and standard deviation for 100 simulations.

2 Assumption Table

Here we detail the nature of the assumptions we have made in the paper, noting situations that are consistent with our assumptions as well as those that are inconsistent with the assumption. We expect this might help the reader understand and evaluate the suitability of the method to their application.

Table 13: Summary of Assumptions

A#	Interpretation	Consistent	Inconsistent
A5(i)	Time invariance of marginal and conditional (on x, p) unobservable state distribution	Unobserved quality control process is constant over time, or changes in distribution of unobservable quality are accompanied with changes in observed product characteristics or prices. If we interpret ξ_{jt} as advertising (or quality control) of the product (either hardware or software in our empirical application), and if the observables (price and characteristics) for the products don't change, then the conditional distribution of unobservable quality remains the same.	If we interpret ξ as advertising, then the advertising expenditures becomes less (or more) volatile over time, while product characteristics and price remain constant. Similarly, quality control process improves while x and p remain the same. <i>Note: Since this is a conditional expectation (on p and x), it does not restrict advertising from increasing in volatility when prices decrease, for example. Thus, in practice it is fairly flexible.</i>
A5(ii)	Future unobservable state $\xi_{j,t+1}$ is independent of current observable state (x_{jt}, p_{jt}) , conditional on future observed state $(x_{j,t+1}, p_{j,t+1})$.	If $\xi_{j,t+1}$ is set based on $x_{j,t+1}$ and $p_{j,t+1}$, then we are ok. Similarly, advertising expenditures are made after the product is manufactured and price is set. Also, if product quality control process is independent of past period features and prices.	In period $t + 1$, firm observes only the prior period's x_{jt} and p_{jt} and sets advertising (or quality control) levels $\xi_{j,t+1}$ based on that, and before current period's $x_{j,t+1}$ and $p_{j,t+1}$.
A6(i)	(Conditional on x_t, p_t) independence of contemporaneous unobservable states (ξ_{jt}) across products j	Each firm makes its quality control or advertising choices independently based on x and p . Note that these choices can depend on the observable characteristics of the prices and characteristics of products made by competitors. <i>Note: Even in the strategic case, if the strategy only depends on observable characteristics of all products, this assumption will be valid.</i>	Firms set advertising expenditures based on expected strategic responses of competitors and they have some information about competitor's advertising choice.
A6(ii)	Two or more products with same variance of unobservable (ξ) conditional on (x_t, p_t)	We have some subset of products that have same variance, conditional on observables (i.e. when their observable characteristics and prices are the same). If we have at least one firm with multiple products, and we expect that the (conditional) variance for these multiple products is identical, then the condition is satisfied. This might happen when the firm has a single quality control process across all products.	Each product has different conditional variance. This maybe possible if each firm has only one product, and each firm has very different advertising policy or quality control policy even when the observable product characteristics and prices of these products are same.

A#	Interpretation	Consistent	Inconsistent
A7	(Conditional on x_t, p_t) unobservable states (ξ_{jt}) across products j have same distribution (except mean)	At least two products made by similar manufacturer with same quality control process, or similar advertising policies. Note that only the (conditional) distribution is required to be identical, not the actual realizations. Also the conditional mean can be different, so only higher-order moments, that is the shape, not the location, of their probability density functions, need to be same.	There are no two products with similar advertising or quality control policies, implying all products have materially different processes that vary in higher-order moments, conditional on observable state.
A8	State Evolution Dependence Structure	Observable characteristics evolve based on previous period observables (characteristics and prices). Observables characteristics do not depend on current or prior unobservables, except price. Unobservables can be quality control process that impact fit and finish of product which do not impact the features developed in future. <i>Note: price and quality control may be contemporaneously related, as might be expected, since firm can set price based on realization of unobservable quality.</i>	Firm invests more in observable product characteristics (e.g. better camera) because its unobservable quality control (or finish) was poor.
A8'	State Evolution Dependence Structure	Current advertising or quality control process does not impact future prices or product characteristics. Current advertising only depends on current prices and product characteristics, but not on past observable characteristics or prices	Firm sets higher advertising level to compensate because its past observable product characteristics did not drive demand.
A9	Unobservable State Evolution	Unobservable characteristics do not depend on competitor's price levels. In the quality control interpretation of ξ_{jt} , this is very likely since the quality control processes are long-term and are unlikely to be changed in response to a competitor's contemporaneous price level. In the advertising interpretation, it implies that advertising or promotional budgets are set independent of current competitor prices. They can depend on own prices.	A firm (Apple) sets advertising budget to be higher to respond to a competitor (Samsung) slashing its price levels.

3 Alternative Counterfactual Procedure

Here we present an alternative to solving the ex-ante value function that does not require value function iteration, the discretization of state variables, nor the use of interpolation. Our counterfactual method is implemented in two steps. The first step recovers the counterfactual impact on within market shares, relative to a given product. This step thus captures the competitive substitution effects between products and does not depend on consumer beliefs in our model specification. The second step moves beyond the competitive effects and determines the impact on the outside market share. The second step allows the researcher to quantify the impact on overall demand, and evaluates whether the counterfactual change leads to expansion or contraction of overall demand.

We consider the counterfactual change of a current product characteristic x_{jt} to counterfactual x_{jt}^c without changing product fixed effect, δ_j , or unobserved product characteristic ξ_{jt} . Other counterfactuals, such as changes to the distribution of state variables, can be addressed similarly. As is standard in structural models, we assume the counterfactual does not affect consumers' preference, product fixed effects and unobserved characteristics. Hence we use the estimated coefficients and unobservable residuals $(\alpha, \beta, \gamma, \delta_j, \xi_{jt})$. In the sequel, we use superscript "c" to denote counterfactual objects, e.g. s_{jt}^c denotes counterfactual market share of product j . We also assume the counterfactual price p_{jt}^c is held constant.

The first step is to generate the counterfactual within or relative market share. By eq. (6), we have counterfactual relative market share as a function of counterfactual (x_{jt}^c, p_{jt}^c) ,

$$\ln \left(\frac{s_{jt}^c}{s_{1t}^c} \right) = (x_{jt}^c - x_{1t}^c)' \tilde{\gamma} - \alpha(p_{jt}^c - p_{1t}^c) + \frac{\delta_j - \delta_1}{1 - \beta} + \frac{\xi_{jt} - \xi_{1t}}{1 - \beta}.$$

After estimation of $(\alpha, \beta, \tilde{\gamma}, \delta_j, (\xi_{jt} - \xi_{1t}))$, we are able to express s_{jt}^c/s_{1t}^c as a known function of $(x_{jt}^c, p_{jt}^c, \delta_j, \xi_{jt})$. For simplicity of exposition, let

$$\tilde{s}_{jt}^c = s_{jt}^c/s_{1t}^c,$$

and let m_t^c denote the vector of all counterfactual state variables.

We can express the counterfactual market share s_{1t}^c as a function of the counterfactual relative market shares and the counterfactual outside market share s_{0t}^c :

$$s_{1t}^c = \frac{1 - s_{0t}^c(m_t^c)}{\sum_{j=1}^J \tilde{s}_{jt}^c(m_t^c)}. \quad (1)$$

We write $s_{0t}^c(m_t^c)$ to emphasize that the counterfactual outside market share s_{0t}^c is a function of counterfactual market state variables.

The second step finds the counterfactual outside market share $s_{0t}^c(m_t^c)$ from the following equation:

$$\ln \left(\frac{1 - s_{0t}^c(m_t^c)}{s_{0t}^c(m_t^c)} \right) = \lambda(m_t^c) + \beta \mathbb{E} \left[\ln(1 - s_{0,t+1}^c(m_{t+1}^c)) \mid m_t^c \right], \quad (2)$$

where

$$\lambda(m_t^c) = \ln \left(\sum_{j=1}^J \tilde{s}_{j,t}^c \right) - \beta \mathbb{E} \left[\ln \left(\sum_{j=1}^J \tilde{s}_{j,t+1}^c \right) \mid m_t^c \right] + v_{1t}^c(m_t^c) - \beta \mathbb{E}(v_{1,t+1}^c(m_{t+1}^c) \mid m_t^c),$$

$$v_{1t}^c(m_t^c) = x_{1t}^c \tilde{\gamma} - \alpha p_{1t}^c + \frac{\delta_1}{1 - \beta} + \frac{\xi_{1t}}{1 - \beta}. \quad (3)$$

From the first step, we have determined $\ln \left(\sum_{j=1}^J \tilde{s}_{j,t}^c \right)$. If ξ_{1t} was known, $v_{1t}^c(m_t^c)$ and hence $\lambda(m_t^c)$, are known as well. We discuss how to determine ξ_{1t} below.

Eq. (2) follows from eq. (11), from which we have

$$\ln \left(\frac{s_{1t}^c}{s_{0t}^c} \right) - v_{1t}^c(m_t^c) = -\beta \mathbb{E}(v_{1,t+1}^c(m_{t+1}^c) - \ln s_{1,t+1}^c \mid m_t^c).$$

Substituting s_{1t}^c above with its formula from eq. (1), we get eq. (2). For a stationary dynamic programming problem, $s_{0t}^c(m_t^c)$ is a time invariant function. Eq. (2) is then an integral equation of s_{0t}^c , from which one solve s_{0t}^c .

3.1 Dimension reduction and other details

In many applications, the dimension of the market state variables m_t^c is proportional to the number of states per product with the number of products as an exponential, and could be computationally infeasible to solve. The curse of dimensionality could arise if either the number of products or observed characteristics is large. For example, in our mobile phone application, there are 7 brands and 9 product characteristics (including 7 product features, price and 1 unobservable characteristic), leading to a $9 \times 7 = 63$ -dimensional continuous state space. Thus, if we discretize the continuous variables and represent them each with n points, the dimension of the state space m_t^c is n^{63} . Thus, if we choose $n = 10$, we have 10^{63} points in the state space. Solving this problem with value function iteration, for example, becomes computationally infeasible.

Thus, we consider using alternative approaches to computing the value function. Tradi-

tionally, researchers assume consumers track all state variables, but as noted above this leads to a curse of dimensionality. One widely known approach that eliminates this problem is to assume consumers track the inclusive value as the relevant state variable (Melnikov, 2013; Gowrisankaran and Rysman, 2012) so that consumers make choices based on the evolution of the inclusive value. An alternative and less restrictive option as it does not rely on the inclusive value sufficiency assumption (which implies that if two different states have the same option value, then they also have the same value function) is to assume consumers track the conditional value function v_{jt} of all products. Thus, the state space in this latter example is of dimension J . This is more general than the inclusive value assumption, since the inclusive value is a deterministic function of the conditional values of all products. Broadly speaking, our counterfactual approach could accommodate any conceivable set of assumptions that can be used to generate the consumer choice data. Depending on the application context, different methods might be more or less suitable.

Below we reduce the dimension by replacing m_t^c with $(v_{1t}^c, \dots, v_{Jt}^c)$, which is defined by eq. (3). Then eq. (2) reads

$$\ln \left(\frac{1 - s_{0t}^c(v_{1t}^c, \dots, v_{Jt}^c)}{s_{0t}^c(v_{1t}^c, \dots, v_{Jt}^c)} \right) = \lambda(v_{1t}^c, \dots, v_{Jt}^c) + \beta \mathbb{E} \left[\ln(1 - s_{0,t+1}^c(v_{1,t+1}^c, \dots, v_{J,t+1}^c)) \mid v_{1t}^c, \dots, v_{Jt}^c \right], \quad (4)$$

In practice the conditional expectation terms in the above display and $\lambda(v_{1t}^c, \dots, v_{Jt}^c)$ can be estimated by a nonparametric regression. Because $s_{0t}^c(v_{1t}^c, \dots, v_{Jt}^c)$ is a conditional probability of choice, one can use the series logit method in the treatment effects literature (Hirano, Imbens, and Ridder, 2003) to approximate it:

$$s_0(v_{1t}^c, \dots, v_{Jt}^c; \rho) = \frac{\exp(\psi(v_{1t}^c, \dots, v_{Jt}^c)' \rho)}{1 + \exp(\psi(v_{1t}^c, \dots, v_{Jt}^c)' \rho)}, \quad (5)$$

where $\psi(v_{1t}^c, \dots, v_{Jt}^c)$ is a vector of known approximating functions, e.g. polynomials, of $v_{1t}^c, \dots, v_{Jt}^c$. We use this functional form for convenience since the market share is bounded, i.e. $s_0 \in [0, 1]$, and other functions that constrained it in such a manner would be applicable as well. We then use eq. (4) to find ρ , e.g. by least squares, to recover the counterfactual outside market share. Once we calculate the counterfactual outside market share, we can determine s_{jt}^c from eq. (1).

As we discussed in §5 of the main text, we also need to know ξ_{1t} , which appears in $\lambda(v_{1t}^c, \dots, v_{Jt}^c)$ above, but more generally in order to solve for the ex-ante value function. There are two different ways to implement this, which trades off an additional assumption

for computational simplicity. The first is discussed in the text and consists of drawing from the estimated distribution of ξ_{jt} . The second is to take an alternative approach that uses the following formula for ξ_{1t} , which follows from eq. (11),

$$\left(\frac{1 - \beta\phi_1}{1 - \beta}\right) \xi_{1t} = y_{1t} - \delta_1 + \beta \mathbf{E}(w_{1,t+1} \mid x_t, p_t, \xi_t). \quad (6)$$

If we assume that

$$\mathbf{E}(w_{1,t+1} \mid x_t, p_t, \xi_t) = \mathbf{E}(w_{1,t+1} \mid x_t, p_t, \xi_{2t} - \xi_{1t}, \dots, \xi_{Jt} - \xi_{1t}), \quad (7)$$

we can identify and estimate ξ_{jt} , because $\xi_{jt} - \xi_{1t}$ is identified (c.f. eq. (18)). When ξ_{1t} and p_{1t} is highly correlated, the bias (difference between the left-hand side and the right-hand side in the above display) is expected to be small. The extreme case is when (p_{1t}, ξ_{1t}) follow a bivariate normal distribution (as we assumed in estimation), and their correlation coefficient is one. In this extreme case, knowing p_{1t} is equivalent to knowing ξ_{1t} , hence eq. (7) holds.

4 Derivatives for Calculating Asymptotic Variance

We derive the formulas for $\partial g_{1,(j,k),t}(\theta_1)/\partial\theta$, $\partial g_{2,(j,0),t}(\theta_1)/\partial\theta$, $\partial g_{3,j,t}(\theta)/\partial\theta$, $\partial g_{4,(j,k),t}(\theta)/\partial\theta$, and $\partial g_{5,(j,k),t}(\theta)/\partial\theta$. It is easier to calculate the derivatives for $\theta_1 = (\alpha, \beta, \tilde{\gamma}', \delta)'$ and $\theta_2 = (\tilde{\rho}', \sigma^2, \phi)'$. For $\partial g_{1,(j,k),t}(\theta_1)/\partial\theta$, we have

$$\begin{aligned} g_{1,(j,k),t,\alpha}(\theta) &= z_{(j,k),t}(p_{jt} - p_{kt}) \\ g_{1,(j,k),t,\beta}(\theta) &= -z_{(j,k),t}(\delta_j - \delta_k)/(1 - \beta)^2 \\ g_{1,(j,k),t,\tilde{\gamma}}(\theta) &= -z_{(j,k),t}(x_{jt} - x_{kt})' \\ g_{1,(j,k),t,\delta_i}(\theta) &= \begin{cases} 0 & \text{if } i \neq j, i \neq k \\ -z_{(j,k),t}/(1 - \beta) & \text{if } i = j \\ z_{(j,k),t}/(1 - \beta) & \text{if } i = k \end{cases} \\ g_{1,(j,k),t,\theta_2}(\theta) &= 0'. \end{aligned}$$

For $\partial g_{2,(j,0),t}(\theta_1)/\partial\theta$, we have

$$\begin{aligned}
g_{2,(j,0),t,\alpha}(\theta) &= x_{j,t,IV,t}(p_{jt} - \beta p_{j,t+1}) \\
g_{2,(j,0),t,\beta}(\theta) &= x_{j,t,IV,t}w_{j,t+1} \\
g_{2,(j,0),t,\tilde{\gamma}}(\theta) &= x_{j,t,IV,t}(-x'_{jt} + \beta x'_{j,t+1}) \\
g_{2,(j,0),t,\delta_i}(\theta) &= \begin{cases} 0 & \text{if } i \neq j \\ -x_{j,t,IV,t} & \text{if } i = j \end{cases} \\
g_{2,(j,0),t,\theta_2}(\theta) &= 0'.
\end{aligned}$$

For $\partial g_{3,j,t}(\theta)/\partial\theta$, we have

$$\begin{aligned}
g_{3,j,t,\alpha}(\theta) &= z_{\rho,jt}(1 - \beta)(p_{jt} - \beta p_{j,t+1}) \\
g_{3,j,t,\beta}(\theta) &= z_{\rho,jt}[-(y_{jt} + \beta w_{j,t+1}) + (1 - \beta)w_{j,t+1} + \tilde{\rho}_j \tilde{p}_{j,t+1}] \\
g_{3,j,t,\tilde{\gamma}}(\theta) &= z_{\rho,jt}(1 - \beta)(-x'_{jt} + \beta x'_{j,t+1}) \\
g_{3,j,t,\delta}(\theta) &= 0' \\
g_{3,j,t,\tilde{\rho}_i}(\theta) &= \begin{cases} 0 & \text{if } i \neq j \\ -z_{\rho,jt}(\tilde{p}_{jt} - \beta \tilde{p}_{j,t+1}) & \text{if } i = j \end{cases} \\
g_{3,j,t,\sigma^2}(\theta) &= 0' \\
g_{3,j,t,\phi}(\theta) &= 0'.
\end{aligned}$$

For $\partial g_{4,(j,k),t}(\theta)/\partial\theta$, we will need $\partial d_{(j,k),t}(\theta)/\partial\theta$:

$$\begin{aligned}
d_{(j,k),t,\alpha}(\theta) &= (1 - \beta)(p_{jt} - p_{kt}) \\
d_{(j,k),t,\beta}(\theta) &= - \left[\ln \left(\frac{s_{jt}}{s_{kt}} \right) - (x_{jt} - x_{kt})' \tilde{\gamma} + \alpha(p_{jt} - p_{kt}) \right] \\
d_{(j,k),t,\tilde{\gamma}}(\theta) &= -(1 - \beta)(x_{jt} - x_{kt})' \\
d_{(j,k),t,\delta_i}(\theta) &= \begin{cases} 0 & \text{if } i \neq j, i \neq k \\ -1 & \text{if } i = j \\ 1 & \text{if } i = k \end{cases} \\
d_{(j,k),t,\theta_2}(\theta) &= 0'.
\end{aligned}$$

We have

$$\begin{aligned}
g_{4,(j,k),t,\theta_1}(\theta) &= d_{(j,k),t}d_{(j,k),t,\theta_1}(\theta) \\
g_{4,(j,k),t,\tilde{\rho}_i}(\theta) &= \begin{cases} 0 & \text{if } i \neq j, i \neq k \\ \tilde{\rho}_k\tilde{p}_{jt}\tilde{p}_{kt} & \text{if } i = j \\ \tilde{\rho}_j\tilde{p}_{jt}\tilde{p}_{kt} & \text{if } i = k \end{cases} \\
g_{4,(j,k),t,\sigma^2}(\theta) &= -1 \\
g_{4,(j,k),t,\phi}(\theta) &= 0'.
\end{aligned}$$

For $\partial g_{5,(j,k),t}(\theta)/\partial\theta$, we have

$$\begin{aligned}
g_{5,(j,k),t,\alpha}(\theta) &= \frac{d_{(j,k),t}}{\beta\sigma^2}d_{(j,k),t,\alpha}(\theta) - \left(\frac{1-\beta}{\beta}\right)(p_{jt} - \beta p_{j,t+1})\frac{d_{(j,k),t}}{\sigma^2} - \\
&\quad \left(\frac{1-\beta}{\beta}\right)(y_{jt} + \beta w_{j,t+1})\frac{d_{(j,k),t,\alpha}(\theta)}{\sigma^2} \\
g_{5,(j,k),t,\beta}(\theta) &= \left(\frac{d_{(j,k),t}d_{(j,k),t,\beta}(\theta)}{\beta\sigma^2} - \frac{d_{(j,k),t}^2}{2\beta^2\sigma^2}\right) + \\
&\quad \frac{1}{\beta^2}(y_{jt} + \beta w_{j,t+1})\frac{d_{(j,k),t}}{\sigma^2} - \\
&\quad \left(\frac{1-\beta}{\beta}\right)\left[w_{j,t+1}\frac{d_{(j,k),t}}{\sigma^2} + (y_{jt} + \beta w_{j,t+1})\frac{d_{(j,k),t,\beta}(\theta)}{\sigma^2}\right] \\
g_{5,(j,k),t,\tilde{\gamma}}(\theta) &= \frac{d_{(j,k),t}}{\beta\sigma^2}d_{(j,k),t,\tilde{\gamma}}(\theta) - \\
&\quad \left(\frac{1-\beta}{\beta}\right)\left[(-x'_{jt} + \beta x'_{j,t+1})\frac{d_{(j,k),t}}{\sigma^2} + (y_{jt} + \beta w_{j,t+1})\frac{d_{(j,k),t,\tilde{\gamma}}(\theta)}{\sigma^2}\right] \\
g_{5,(j,k),t,\delta}(\theta) &= \left[\frac{d_{(j,k),t}}{\beta\sigma^2} - \left(\frac{1-\beta}{\beta}\right)(y_{jt} + \beta w_{j,t+1})\frac{1}{\sigma^2}\right]d_{(j,k),t,\delta}(\theta) \\
g_{5,(j,k),t,\tilde{\rho}}(\theta) &= 0' \\
g_{5,(j,k),t,\sigma^2}(\theta) &= -\left[\frac{d_{(j,k),t}^2}{2\beta} - \left(\frac{1-\beta}{\beta}\right)(y_{jt} + \beta w_{j,t+1})d_{(j,k),t}\right]\frac{1}{\sigma^4} \\
g_{5,(j,k),t,\phi_i}(\theta) &= \begin{cases} 0 & \text{if } i \neq j \\ -1 & \text{if } i = j \end{cases}.
\end{aligned}$$

References

- Gowrisankaran, G. and M. Rysman (2012). Dynamics of consumer demand for new durable goods. *Journal of political Economy* 120(6), 1173–1219.
- Hirano, K., G. W. Imbens, and G. Ridder (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* 71(4), 1161–1189.
- Melnikov, O. (2013). Demand for differentiated durable products: The case of the us computer printer market. *Economic Inquiry* 51(2), 1277–1298.