Online Appendix to "Dynamic Efficiencies of the 1997 Boeing-McDonnell Douglas Merger"

Yonghong An* Wei Zhao†

October 22, 2018

A. Details of the Medium-Wide-Ratio R_w

To construct the ratio R_w, we use monthly data on the total number of flights for each aircraft sub-model (e.g., B777-200) on any U.S. domestic and international route during the 1990-2011 period from Department of Transportation (www.transportation.gov). For each month-aircraft-route observation, we also observe number of passengers, pound of freights, distance of routes, and total flying time.

To compute R_w, we focus on those routes where wide-body aircraft are flying in a nontrivial frequency. For this purpose, we select the routes with distances being longer than 1000 miles and there is at least one flight of medium-sized, wide-body aircraft; merge all of the post-merger years' data (1997-2011) and then keep only those routes that have, on average, at least 50 flights of any wide-body aircraft per year; and drop all non-jet observations because they are not expected to fly on a route where wide-body aircraft are also flying anyway.

We end up with 908 routes. Among these routes, a substantial portion of international ones are exclusively served by medium-sized, wide-body aircraft. For example, 55% of the flights are A340 and 45% are B777 for the route from New York to Shanghai, China. By contrast, typical routes with a medium-wide-ratio close to 0.5 are hub-to-hub domestic routes, e.g., Los Angeles to Chicago. A histogram of R_w is presented in panel (b) of Figure 1 in the main text. The figure illustrates that the ratio concentrates at 0 and 1: only 38.5% of the routes are in the interval $[0.2, 0.8]$ and 26% in $[0.3, 0.7]$.

B. A two-level nested logit model of demand

The nested logit model in Section 4 of An and Zhao (2018) by nature does not allow greater substitution between submodels of the same model. To relax this restriction, we

*Department of Economics, Texas A&M University, College Station, TX 77843; email: y.an@tamu.edu.
†Competition Economics LLC, 2000 Powell Street, Suite 510, Emeryville, CA 94608; email: wzhao@c-econ.com.

consider an alternative approach to estimating demand function using a two-level nested logit model (Verboven, 1996). Specifically, consumer i's utility function from aircraft j at time t is expressed as:

$$\begin{aligned} v_{ijt} &= \delta_{jt} + \omega_{ijt}, \\ \delta_{jt} &= X_{jt}\beta - \alpha p_{jt} + \xi_{jt}, \\ \omega_{ijt} &= \zeta_{igt} + (1-\sigma_1)\varepsilon_{ijt} + (1-\sigma_2)\eta_{ihgt}, \end{aligned} \quad (1)$$

where g refers to group and h to model. The error terms ζ_{igt}, ε_{ijt} and η_{ihgt} are respectively the random tastes at the group, model and submodel levels. They have the unique distribution such that ζ_{igt}, $(1-\sigma_1)\varepsilon_{ijt} + (1-\sigma_2)\eta_{ihgt}$ and $\zeta_{igt} + (1-\sigma_1)\varepsilon_{ijt} + (1-\sigma_2)\eta_{ihgt}$ have the extreme value distribution. The two nesting parameters σ_1 and σ_2 capture the consumers' preference correlation across models and submodels. The random utility theory requires that $0 \leq \sigma_2 \leq \sigma_1 \leq 1$.

Verboven (1996) provides the well-known formula of the market share:

$$\ln(s_{jt}/s_{0t}) = X_{jt}\beta - \alpha p_{jt} + \sigma_1 \ln(s_{j|hg,t}) + \sigma_2 \ln(s_{h|g,t}) + \xi_{jt}, \quad (2)$$

where s_{jt} and s_{0t} are defined the same as in the one-level model. $s_{j|hg,t}$ and $s_{h|g,t}$ are market shares of submodel j in model h of group g, and of model h in group g.

$$s_{j|hg,t} = \frac{q_{jt}}{\sum_{k=1}^{J_{hg}} q_{kt}}, \; s_{h|g,t} = \frac{\sum_{k=1}^{J_{hg}} q_{kt}}{\sum_{h=1}^{J_g} \sum_{k=1}^{J_{hg}} q_{kt}},$$

where J_{hg} is the number of submodels in model h of group g, and J_g is the number of models in group g. The total number of submodels is $J = \sum_{h=1}^{J_g} J_{hg}$.

The unobserved characteristic ξ_{jt} is likely correlated with price p_{jt}, and consequently the two market shares $s_{j|hg,t}$, $s_{h|g,t}$ are also endogenous. We use a standard instrumenting approach to estimate the parameters through two-stage least squares (2SLS). Instruments used include hourly wage in manufacturing and its lagged terms, price of aluminum and its lagged terms, average characteristics for products of the same model and group as product j, respectively, the number of products in the same model and group as product j. The correlation parameters σ_1 is identified by covariation between the within-subgroup market share of the plane $s_{j|hg,t}$ and its total market share s_{jt}. Similarly, σ_2 is identified by the correlation between $s_{h|g,t}$ and s_{jt}.

To estimate the two-level nested logit model, we classify the twelve products into two groups as in the one-level model: new medium-sized aircraft and the outside aircraft, which is defined to be all new non-medium-sized wide-body aircraft and all used medium-sized wide-body aircraft. The group of new medium-sized aircraft contains four models: A330, A340, B777, and MD-11. Their submodels are A330-200, A330-300; A340-200, A340-300, A340-500, A340-600; B777-200, B777-200ER, B777-200LR, B777-300, B777-300ER, and MD-11 has no submodel.

We present the estimation results in Table 1. The results demonstrate that the estimates are similar to those in the one-level nested logit model. The submodel and model nesting parameters in the two-level mode are estimated to be $\sigma_1 = 0.86$ and $\sigma_2 = 0.82$ in the full specification. These estimates satisfy the requirements $1 \geq \sigma_1 \geq \sigma_2 \geq 0$ for the model to be consistent with the random utility theory. The strict inequality $\sigma_1 > \sigma_2$ would imply that products of the same model–e.g., A330-200 and A330-300–are the closest substitutes for consumers. Products of medium-sized but not the same model–e.g., A330-200 and B777-200–are weaker substitutes, and products of medium-sized and the outside ones are the weakest substitutes. Nevertheless, a formal test indicates that we cannot reject the hypothesis $\sigma_1 = \sigma_2$ at 1%, 5%, and 10% levels of significance (p-value is 0.50). With $\sigma_1 = \sigma_2$, the two-level model is reduced to one level, where the nests are groups and consumers' preferences are only correlated across products of the same group; there is no additional correlation across products of different subgroups within a group.

C. An Alternative Experience Accumulation Function

In this section, we propose a more general experience accumulation function than we present in the main text.

Specifically, we allow the contribution of product k's quantity to product j's experience level to depend on the "distance" between these two products, where the distance is measured by resemblance in aircraft characteristics.[1] The generalized accumulation function of experience is:

$$E_{j,t+1} = \delta E_{jt} + \sum_{k=1}^{J} \theta_{jk} f(X_j, X_k; \upsilon) q_{kt};\ E_{j1} = 1,\ \forall j. \tag{3}$$

where

$$\theta_{jk} = \begin{cases} 1 & \text{if } j = k, \\ \theta_1 & \text{if } k \text{ is a different submodel from } j, \\ \theta_2 & \text{if } k \text{ is a different model in the same firm,} \\ \theta_3 & \text{if } k \text{ is a model from another firm.} \end{cases}$$

The parameters θ_1, θ_2 and θ_3 capture spillover across submodels (e.g., A330-200 and A330-300), across-products in the same firm (e.g., A330 and A340), and across-firms (e.g., MD-11 and B777), respectively. The function $f(X_j, X_k; \upsilon)$ measures products' distance in two dimensions of characteristics: number of seats and maximum ranges.[2]

$$f(X_j, X_k; \upsilon) \equiv \upsilon_1^{\frac{|X_{j1} - X_{k1}|}{D_1}} \upsilon_2^{\frac{|X_{j2} - X_{k2}|}{D_2}}, \tag{4}$$

where the indices 1 and 2 stand for "number of seats" and "maximum range," respectively. $D_h = \max_j |X_{jh} - X_{kh}|$, $h = 1, 2$ is the maximum difference, and υ_1 and υ_2 are both

[1] The spillover effects measured by the parameters here are net effects in the sense that increases in quantities of other aircraft also may spur competition for experienced workers in the labor market. Thus, the parameters represent net effects of experience spillover and labor market competition.

[2] We also tried fuselage and some other characteristics, and we found that the results did not change significantly.

between 0 and 1. Note that from (3) and (4), the larger the difference is for a given υ, the smaller the spillover effect; υ_h is close to 0 implies that characteristic h has strong impacts, whereas υ_h is close to 1 suggests that characteristic h has little impact on the spillover rate. In the special case where $\upsilon_1 = \upsilon_2 = 1$, the transition of experience (3) degenerates to the simplified version in our main text:

$$E_{j,t+1} = \delta E_{jt} + \sum_{k=1}^{J} \theta_{jk} q_{kt},$$

where there are no effects of aircraft characteristics differences on the experience accumulation.

We estimate the learning parameters using the specification (3) and the labor input function in Section 5.2 of our main text. The estimating approach and instrumental variables are the same as in An and Zhao (2018). We present the estimation result of the learning curve in Table 2, the estimates are similar to Table 5. This is because the effects of characteristics are estimated to be marginal, i.e., both the estimates of υ_1 and υ_2 are both close to 1. The results imply that when we model learning-by-doing, it is sufficient to consider spillover across submodels, models, and firms, but not similarity in characteristics.

D. Modeling Generation Upgrade

In this section, we extend the dynamic game in the main text by adding firms' endogenous decisions of upgrading their products, and by including products' generation level as a state variable.

D.1. Definition of generation upgrade

As in Goettler and Gordon (2011), we measure a product's generation as the difference between this product and an outside good, where the outside good stochastically improves over time, and the difference is bounded. The outside good is a composition of any products that can be viewed by consumers as substitutes of the product we consider. In our paper, the outside good consists of old wide-body aircraft for sale and new non-medium-sized, wide-body aircraft, and narrow-body aircraft. Formally, such a relative generation of product j at time t, denoted as G_{jt}, is defined as

$$G_{jt} = g_{jt} - g_{0t}, \tag{5}$$

where g_{jt} and g_{0t} are the generation levels of product j and the outside good at period t, respectively.

There is no unique way to define a generation g_{jt} of an aircraft. A criterion is that a new generation should have substantial demand-side advantages that can be attributed to more desirable characteristics over the old generation. The wide-body aircraft industry had evolved for 28 years before the merger in 1997, and there was hardly any room left for firms to introduce new aircraft with characteristics (e.g., range, plane size, number of seats, etc.) that meet market demand but had not yet covered by existing aircraft. In actuality, aircraft models introduced after the 1997 merger were generally driven by concern over operating cost[3]. Therefore, we define a new generation of aircraft as one that provides at least 5% lower operating cost for airlines. In the medium-sized, wide-body aircraft market, we treat introduction of Boeing 777, 787, and Airbus A350 as new generations, i.e., B787 and A350 are new generations of B777 and A330, respectively. Based on this definition, there are no more than two generation lags among all new medium-sized, wide-body aircraft, hence the support of the state variable G_{jt} is $\{-1, 0, 1\}$. $G_{jt} = -1$ denotes that product j is of one generation behind the outside aircraft; $G_{jt} = 1$ denotes one generation ahead of it.

D.2. The dynamic game with generation upgrade

Once the relative generation of product G_{jt} is incorporated into the dynamic model as a state variable, the state of the industry is then characterized by a profile $\omega = (E, G, \xi, M)$. A firm n makes joint decisions on upgrading and quantity choices in each period. Let $d_{jt} \in \{0, 1\}$ denote the decision on upgrading product j at t. Then firm n's choices are $D_{nt} \equiv (d_{1t}, d_{2t}, \cdots, d_{J_n t}) \in \{0, 1\}^{J_n}$, and $Q_{nt} \equiv (q_{n1t}, q_{n2t}, \cdots, q_{nJ_n t})$. Upgrading a product j ($d_j = 1$) incurs direct cost c^G_{jt}, and it is also associated with an indirect cost through a setback in experience. Specifically, if $d_{jt} = 1$, then the experience of product j decreases from E_{jt} to $\psi(E_{jt})$, where $\psi(x)$ is a given function and $\psi(x) < x$ for any x. The dynamic game can be divided into three stages in each period:

- Stage (i): Nature draws shocks on demand (M and ξ), and innovation of the outside good g_{0t}. All draws are immediately observed by all firms.
- Stage (ii): Firms learn their upgrading cost, which is private information. Then firms simultaneously make adoption decisions (D_n). The resulting new generation levels of all products are immediately observed by all firms.
- Stage (iii): Firms compete in a simultaneous quantity competition game. Experience level for each product is realized based on quantity choices and is revealed to all firms.

Note that quantity and upgrading decisions are made in different stages. Thus, expected future values need to be constructed differently from the main text when solving for optimal quantity and upgrading policies. To deal with these complexity, it is useful to be specific about stages for the state profile: we denote the state profile at the beginning of stage (ii) and (iii) as ω and $\widetilde{\omega}$, respectively. At the beginning of stage (ii), the realized upgrading costs and upgrading choices are private information. Given the state profile ω,

[3] This is confirmed by discussions with Edmund S. Greenslet, an aircraft industry expert, and publisher of *The Airline Monitor*.

a firm believes that the upgrading decision $d_k \in \{0,1\}$ for its competitor's product k will be realized with probability p_k. Let P_{-n} denote the vector of probabilities firm n forms for its competitors' decisions.

Given the state profile ω, firm n makes a joint upgrading decision D_n. Choosing D_n incurs total upgrading cost $C_n^G \cdot D_n$, where C_n^G is the vector of upgrading costs $C_n^G = (c_1^G, c_2^G, \cdots, c_{J_n}^G)$. Denote firm n's expected value, excluding upgrading cost, of choosing D_n as EV^{D_n}, which is the summation of expected values across all products firm n owns. The expectation is over other firms' upgrading probabilities P_{-n} as well as the associated upgrading costs. Let D_n and D'_n be two different vectors of upgrading choices from the set $\{0,1\}^{J_n}$.

Firm n will choose D_n if it gives the firm the largest net continuation value (expected future value less upgrade cost). Thus, the probability of choosing the vector D_n is

$$P^{D_n} = \Pr\left(EV^{D_n} - C_n^G \cdot D_n \geq EV^{D'_n} - C_n^G \cdot D'_n\right), \forall\, D'_n \neq D_n, \forall n = 1, 2, \cdots, N, \quad (6)$$

where the vector of probabilities P_{-n} is implicitly contained in EV^{D_n} and $EV^{D'_n}$.

With equilibrium (P^{D_n}, P_{-n}) being solved from (6), we now turn to equilibrium quantity choices. Since production affects future variable cost through its direct impact on experience accumulation, production decisions for each period could no longer be modeled as static. Quantities enter both the current profit function and the next period value function in the Bellman equation. Aside from this quantity effect on future costs, the per period game is a quantity competition with heterogeneous goods and multi-product firms. The flow profit of a firm from product j is

$$\pi_j(\boldsymbol{Q}, \tilde{\boldsymbol{\omega}}) = p_j(\boldsymbol{Q}; X_j, \xi_j, M, G_j)q_j - C_j(q_j; E_j, G_j), \quad (7)$$

where we suppress the index t whenever there is no ambiguity. The cost function $C_j(q_j; E_j, G_j)$ is the sum of a fixed cost, total variable cost, and potential upgrading costs.

Let ρ denote the discount factor of firms, then the model described above can be characterized by the following Bellman equations for firm n,

$$V_n(\tilde{\boldsymbol{\omega}}) = \max_{Q_n} \left\{ \sum\nolimits_{j \in \mathcal{J}_n} \pi_j(Q_n, Q_{-n}, \tilde{\boldsymbol{\omega}}) + \rho \int V_n(\tilde{\boldsymbol{\omega}}')\mathcal{F}(\tilde{\boldsymbol{\omega}}'|\tilde{\boldsymbol{\omega}}, \boldsymbol{Q})d\tilde{\boldsymbol{\omega}}' \right\}, \quad (8)$$

In solving the model numerically, we track the probability vector (P^{D_n}, P_{-n}) for each state profile ω, and quantity and value function for each state profile $\widetilde{\omega}$. We find that tracking $V_n(\widetilde{\omega})$ instead of $V_n(\omega)$ makes computation much easier. As in the base model, we use the symmetric Markov perfect Nash equilibrium (MPE) as the equilibrium concept.

D.3. State transition and calibration of parameters

D.3.1. State transition

We discuss state transition in this section. Since M_t and ξ_{jt} evolve exogenously, we only need to specify transitions of E_{jt} and G_{jt}, which are both affected by upgrading decisions. Note that G_{jt} is only updated in stage (i) and (ii) in the dynamic game, then transition of G_{jt} does not depend on transition of E_{jt} but G_{jt}'s transition affects that of E_{jt}. Specifically,

$$\Pr(\tilde{\omega}'|\tilde{\omega}) = \Pr(\omega'|\tilde{\omega}) \times \Pr(\tilde{\omega}'|\omega'), \tag{9}$$

where the transition of G_{jt} and E_{jt} are in the first and second terms on the right-hand-side, respectively.

Let p^G be the probability that the generation of outside good g_{0t} advances in each period. Then the transition of g_{0t} is

$$g_{0,t+1} = \begin{cases} 1 + g_{0,t}, & \text{with probability } p^G, \\ g_{0,t}, & \text{with probability } 1 - p^G. \end{cases} \tag{10}$$

Note that the probability p^G can be understood as the long-run industry innovation rate in the equilibrium. Evolution of g_{jt} is controlled by joint upgrading decisions. When product j is upgraded in period t, its generation will increase by 1. Otherwise, g_{jt} stays the same. Combining the evolution of g_{0t} and g_{jt}, we obtain the transition of G_{jt} and exemplify it in the following for the case of $G_{jt} = 0$.

$$G_{j,t+1} = \begin{cases} 1, & \text{with prob. } p_j^u, \\ 0, & \text{with prob. } p_j^s, \\ -1, & \text{with prob. } p_j^d, \end{cases} \tag{11}$$

where p_j^u, p_j^s and p_j^d denote probability of $G_{j,t}$ after stage (ii) in the next period increases, remains the same, and decreases, respectively.

$$\begin{aligned} p_j^u &= (1 - p^G) \Pr(d_{jt} = 1 | g_{0t} = g_{0,t-1}), \\ p_j^s &= p^G \Pr(d_{jt} = 1 | g_{0t} = g_{0,t-1} + 1) + (1 - p^G) \Pr(d_{jt} = 0 | g_{0t} = g_{0,t-1}), \\ p_j^d &= p^G \Pr(d_{jt} = 0 | g_{0t} = g_{0,t-1} + 1). \end{aligned}$$

Note that if $G_{jt} = 1$ and $G_{jt} = -1$, then $G_{j,t+1} \in \{0, 1\}$ and $G_{j,t+1} \in \{-1, 0\}$, respectively. The transition of G_{jt} can be derived analogously to (11).

In our dynamic model, whenever $d_{jt} = 1$ the experience level will be set back. The experience setback function $\psi(\cdot)$ is difficult to estimate because it requires observations of a products's unit labor requirement and generation upgrade choices. Our strategy is to assume generation upgrades setback experience by m grid points for the discrete experience state introduced in the dynamic model. That is,

$$\psi(E_{jt}) = \max\{E^{k-m}, E^1\}, \tag{12}$$

where E^k, $k \in \{1, 2, ..., 7\}$ is the discretized grid that E_{jt} takes. Based on this approach and the transition of generation G_{jt}, it is readily to derive the transition of E_{jt}.

$$E_{j,t+1} = \begin{cases} E_{jt}^u \text{ downgrade by } m, & \text{with prob. } p_j^* \cdot \eta_j(1), \\ E_{jt}^d \text{ downgrade by } m, & \text{with prob. } p_j^* \cdot \eta_j(0), \\ E_{jt}^u, & \text{with prob. } (1 - p_j^*) \cdot \eta_j(1), \\ E_{jt}^d, & \text{with prob. } (1 - p_j^*) \cdot \eta_j(0), \end{cases} \tag{13}$$

where $p_j^* = p^G \Pr(d_{jt} = 1 | g_{0t} = g_{0,t-1} + 1) + (1 - p^G) \Pr(d_{jt} = 1 | g_{0t} = g_{0,t-1})$.

D.3.2. Calibration of model parameters

We discussed the computation of expected value function without the generation upgrade state variable G_{jt} in Appendix B of the main paper. The computation can be readily extended given the transitions of G_{jt} and E_{jt}.

Next, we discuss all the parameters related to generation upgrading that are used in the dynamic game. We first assume that the upgrading cost c_j^G for any product j is a random draw from a uniform distribution on $[C^d, C^u]$. The lower and upper bounds, C^d and C^u, are set as the minimum and maximum of reported upgrading costs collected from various sources for the wide-body aircraft. In our merger analysis presented in the main text, C^d and C^u are \$330 and \$614 in 1994 million dollars.

Since there are not enough instances of generation upgrade observed to fully estimate the experience setback in the merger evaluation, we set $m = 0, 1, 2$ and find that the results do not change qualitatively. We present the results for $m = 1$ in the main text.

Finally, p^G, which represents the industry long-run innovation rate, is obtained as the inverse of the average years across products before generation advances. For our merger analysis we choose $p^G = 0.09$.

References

An, Y. and W. Zhao, "Dynamic Efficiencies of the 1997 Boeing-McDonnell Douglas Merger," Working Paper, 2018. (document)

Goettler, R. L. and B. R. Gordon, "Does AMD Spur Intel to Innovate More?," *Journal of Political Economy* 119(6) (2011), 1141–1200. (document)

Pakes, A. and P. McGuire, "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model," *RAND Journal of Economics* 25(4) (1994), 555–589.

Verboven, F., "International price discrimination in the European car market," *RAND Journal of Economics* (1996), 240–268. (document)

Table 1: Estimate of Demand Estimation

	Estimates	
	(a)	(b)
Constant	-2.84***	-3.17***
	(0.43)	(0.48)
Price/100	1.89***	1.90***
	(0.27)	(0.27)
Fuel eff.	-0.36***	-0.30**
	(0.13)	(0.13)
Seat/100	0.29***	0.28***
	(0.094)	(0.092)
Range/10000	1.05***	0.97***
	(0.165)	(0.165)
Engine	-0.16***	-0.15***
	(0.031)	(0.030)
911 (dummy)	-0.62***	-0.62***
	(0.087)	(0.086)
Generation	—	0.15
		(0.10)
Subgroup corr. (σ_1)	0.86***	0.86***
	(0.030)	(0.030)
Group corr. (σ_2)	0.82***	0.82***
	(0.068)	(0.066)
Own-price elasticity	[-13.06, -4.41]	[-13.12,-4.51]

Note: The number of observations is 115. The dependent variable is $\ln(s_{jt}/s_{0t})$. Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Estimates of Learning Curve Parameters

Parameters	Explanation	Estimate
$\ln A$	Labor Cost Intercept	9.26***
		(3.29)
γ_2	Return to Scale	0.32
		(0.59)
γ_1	Learning Parameter	-1.15***
		(0.14)
δ	Depreciation of Experience	0.96***
		(0.001)
θ_1	In-family Spillover	0.99***
		(0.024)
θ_2	In-firm Spillover	0.24***
		(0.003)
θ_3	Across-firm Spillover	0.014***
		(0.002)
υ_1	Seats Diff.	0.99***
		(0.0037)
υ_2	Maximum Range Diff.	0.99***
		(0.0032)
$1 - 2^{\gamma_1}$	Implied Learning Rate†	55%

Note: Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. †Implicit learning rate measures percent of labor saving when experience doubles.