Online Appendix to "Price Competition and Endogenous Product Choice in Networks: Evidence from the US Airline Industry"

Christian Bontemps* Cristina Gualdani† Kevin Remmy‡

February 21, 2023

A Proofs

Proof of Theorem 1 (i) comes from the convergence of $\widehat{\Gamma}_I$ to Γ_I with respect to the Hausdorff distance. (ii) comes from Shapiro et al. (2014), Theorem 5.11, p.193. The uniformity in q for both (i) and (ii) comes from the compactness of the unit ball and the identified set. ∎

Proof of Theorem 2 (i) comes from the convergence of $\widehat{\Gamma}_I^\alpha$ to Γ_I^α with respect to the Hausdorff distance. (ii) comes from Shapiro et al. (2014), Theorem 5.11, p.193 combined with the Delta method for the asymptotic variance of the estimated constraints. As a matter of fact, denoting $b_r = \mathbb{E}(Z_{r,m}B_m)$ and $a_r = \mathbb{E}(Z_{r,m}A_m)$, for $r = 1, \ldots, \mathrm{R}$, we have

$$\frac{\partial g_\alpha}{\partial b_r}(\gamma) = \frac{\gamma \exp\left(\alpha\left[b_r^\top \gamma - a_r\right]\right)}{1+\exp\left(\alpha\left[b_r^\top \gamma - a_r\right]\right)} = \frac{\gamma}{1+\exp\left(-\alpha\left[b_r^\top \gamma - a_r\right]\right)},$$
$$\frac{\partial g_\alpha}{\partial a_r}(\gamma) = \frac{-1}{1+\exp\left(-\alpha\left[b_r^\top \gamma - a_r\right]\right)}.$$

*Email: christian.bontemps@tse-fr.eu, ENAC & Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France.

†Email: c.gualdani@qmul.ac.uk, Queen Mary University of London, London, United Kingdom.

‡Email: kevin.remmy.001@gmail.com, University of Mannheim, Mannheim, Germany.

Therefore:

$$\begin{aligned}\sqrt{\mathrm{M}}(\hat{g}_\alpha(\gamma) - g_\alpha(\gamma)) &= \sum_{r=1}^{\mathrm{R}} \frac{\sqrt{\mathrm{M}}\left((\hat{b}_r - b_r)^\top \gamma - (\hat{a}_r - a_r)\right)}{1 + \exp\left(-\alpha\left[b_r^\top \gamma - a_r\right]\right)} + o_P(1), \\ &= \sum_{r=1}^{\mathrm{R}} \frac{W_r(\gamma)}{1 + \exp\left(-\alpha\left[b_r^\top \gamma - a_r\right]\right)} + o_P(1).\end{aligned}$$

The uniformity in q comes from the compactness of the unit ball. ■

B Existence of Nash equilibrium networks

As discussed in Section 3.3 of the main paper, proving the existence of a pure strategy Nash equilibrium (PSNE) $G \coloneqq (G_f : f \in \mathcal{N})$ is difficult due to the presence of spillovers from entry across markets on the demand, marginal cost and fixed cost sides.

Berry (1992)establishes the existence of a PSNE in one of the first empirical models of entry that incorporates strategic interactions between firms in the second-stage pricing game. His proof relies on the assumption that the entry decisions are independent across markets. It is therefore not applicable to our framework. Another approach used in the network formation literature to show the existence of a PSNE is to represent the model as a potential game (Monderer and Shapley, 1996). This is possible if the payoff function is additive separable in the linking decisions and linear in the spillovers (as for example in Mele, 2017), which is not the case here. Alternatively, it is possible to show the existence of a PSNE under the assumption that the game is supermodular, in order to exploit the fixed point theorem for isotone mappings (Topkis, 1979). However, supermodularity does not hold in our setting due to the second-stage competition between airlines. Finally, one could try to decompose the original game into "local" games such that the original game is in equilibrium if and only if each local game is in equilibrium (Gualdani, 2021). In turn, the existence of a PSNE in each local game - which is typically easier to establish - is sufficient for the existence of a PSNE in the original game. However, the classes of spillovers considered in our model do not allow us to implement such a decomposition.

One might also ask whether allowing for private fixed cost shocks could simplify the existence proof. Espín-Sánchez et al. (2021) prove equilibrium existence in an entry model where firms have some private information at the entry stage. However, they do not allow for multi-product firms and they do not allow for spillovers from entry across markets. Moreover, in our setting it is more reasonable to assume that the fixed cost shocks are common knowledge among airlines, as discussed in Section 3.2 of the main paper.

Note that the moment inequalities in Section 4.2 of the main paper are based on necessary conditions for PSNE. Therefore, one could consider a first-stage equilibrium notion that is weaker than PSNE. In particular, given our focus on one-link deviations, inequalities (10) and (11) resemble the notion of pairwise stability used in network theory, according to which no player has profitable deviations by adding or removing a link (Jackson and Wolinsky, 1996). Definition 1 introduces a notion of first-stage equilibrium along the lines of pairwise stability.

Definition 1. (*Pairwise Stability*) The networks $G_1, \ldots, G_{\mathrm{N}}$ represent a pairwise stable outcome if, for each market $\{a, b\} \in \mathcal{M}$ and airline $f \in \mathcal{N}$, it holds that

$$G_{ab,f} = 0 \Rightarrow \Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(+ab),f}, G_{-f}; \theta) + \gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(+ab),f} + \gamma_{1,f} + \eta_{ab,f} \geq 0,$$
$$G_{ab,f} = 1 \Rightarrow \Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(-ab),f}, G_{-f}; \theta) - \gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(-ab),f} - \gamma_{1,f} - \eta_{ab,f} \geq 0.$$

◇

In the absence of ties (it is sufficient that the fixed cost shocks have a continuous distribution), Definition 1 can be rewritten as a simultaneous equation model.

Lemma 1. (*Equivalent representation of pairwise stability*) In the absence of ties, the networks $G_1, \ldots, G_{\mathrm{N}}$ represent a pairwise stable outcome if and only if:

$$G_{ab,f} = \mathbb{1}\{\Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(-ab),f}, G_{-f}; \theta) - \gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(-ab),f} - \gamma_{1,f} - \eta_{ab,f} \geq 0\},$$
$$\forall\{a, b\} \in \mathcal{M}, \forall f \in \mathcal{N}. \tag{B.1}$$

◇

See Menzel (2017) or Sheng (2020) for a proof of Lemma B.1. Note that although pairwise stability is a weaker equilibrium notion than PSNE, establishing the existence of a pairwise stable outcome does not appear to be easier in our setting. In particular, according to Jackson and Watts (2002), for any payoff function there is either a pairwise stable outcome or a closed cycle.[1] A typical way used in the literature to exclude the presence of closed cycles is to show that the model can be represented as a potential game, as discussed by Jackson and Watts (2001) and Hellmann (2013). As before, however, this is possible if the payoff function is additive separable in the link decisions and linear in the spillovers (as in Sheng, 2020), which is not our case.

C How to deal with incoherence

In Section 4.2 of the main paper, we have constructed the identified set for the first-stage parameters under the assumption that PSNE networks exist for each parameter value and variable realisation. As discussed above, proving the existence of PSNE networks is difficult. Therefore, it is legitimate to wonder whether one should modify the definition

[1] A closed cycle represents a situation in which individuals never reach a stable state and constantly alternate between forming and severing links.

of the identified set when non-existence is possible, i.e., when our model is incoherent in the terminology of Tamer (2003) and Lewbel (2007).

To explain how we deal with incoherence, we first report here the moment inequalities predicted by our model as derived in Section 4.2 of the main paper:

$$\begin{aligned}
&\mathbb{E}_{\Pr}\Big[Z_{r,(-ab),f} \times G_{ab,f} \times \Big(\Pi^e_f(G_f, G_{-f};\theta) - \Pi^e_f(G_{-ab,f}, G_{-f};\theta) - \Big(\gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(-ab),f} + \gamma_{1,f}\Big)\Big)\Big] \geq 0,\\
&r = 1, \ldots, \mathrm{R}_-,\\
&\mathbb{E}_{\Pr}\Big[Z_{r,(+ab),f} \times (1 - G_{ab,f}) \times \Big(\gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(+ab),f} + \gamma_{1,f} - \big(\Pi^e_f(G_{(+ab),f}, G_{-f};\theta) - \Pi^e_f(G_f, G_{-f};\theta)\big)\Big)\Big] \geq 0,\\
&r = 1, \ldots, \mathrm{R}_+,
\end{aligned} \tag{C.1}$$

where $\mathbb{E}_{\Pr}$ is the expectation operator based on the probability function Pr associated with the probability space where the random variables of the model are defined. Second, to simplify the exposition, we focus on one moment inequality from (C.1):

$$\mathbb{E}_{\Pr}\Big[Z_{r,(-ab),f} \times G_{ab,f} \times \Big(\Pi^e_f(G_f, G_{-f};\theta) - \Pi^e_f(G_{-ab,f}, G_{-f};\theta) - \Big(\gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(-ab),f} + \gamma_{1,f}\Big)\Big)\Big] \geq 0. \tag{C.2}$$

Third, we streamline the notation of (C.2) as:

$$\mathbb{E}_{\Pr}(G_m\tilde{A}_m) - \mathbb{E}_{\Pr}(G_m\tilde{B}_m^\top)\gamma \geq 0, \tag{C.3}$$

where the subscripts f and r are omitted, m is a market $\{a, b\}$, $\tilde{A}_m$ is $Z_{r,(-ab),f}(\Pi^e_f(G_f, G_{-f};\theta) - \Pi^e_f(G_{-ab,f}, G_{-f};\theta))$, $\tilde{B}_m$ is such that $\tilde{B}_m^\top\gamma$ is equal to $Z_{r,(-ab),f}(\Delta\overline{\mathrm{Q}}_{(-ab),f}\gamma_{2,f} + \gamma_{1,f})$.

Let $\mathbb{P}$ be the distribution of $(G_m\tilde{A}_m, G_m\tilde{B}_m)$ identified by the sampling process. If the set of PSNE networks is non-empty for each parameter value and variable realisation, then we can replace $\mathbb{E}_{\Pr}$ with $\mathbb{E}_{\mathbb{P}}$ in (C.3) and obtain the identified set for γ associated with $\mathbb{P}$:

$$\Gamma_I \coloneqq \Big\{\gamma \in \Gamma : \ \mathbb{E}_{\mathbb{P}}(G_m\tilde{A}_m) - \mathbb{E}_{\mathbb{P}}(G_m\tilde{B}_m^\top)\gamma \geq 0\Big\}. \tag{C.4}$$

If the set of PSNE networks is empty for some parameter values and variable realisations, then the relationship between $\mathbb{P}$ and Pr is not completely defined because our model is silent about the realisations of $(G_m\tilde{A}_m, G_m\tilde{B}_m)$ when the set of PSNE networks is empty. Since non-existence outcomes are never observed in our data, we approach the incoherence problem by assuming that the data are drawn from the subset of the sample space in which the set of PSNE networks is non-empty. That is, $\mathbb{P}$ comes from a truncated version of Pr, as discussed in Section 4.2 of Chesher and Rosen (2020). In what follows, we show that the identified set for γ associated with $\mathbb{P}$ is still defined by (C.4).

For ease of explanation, let us assume that $\tilde{A}_m$ and $\tilde{B}_m$ are discrete random variables. Given $\gamma \in \gamma_I$, our model predicts that

$$\sum_{a\in\mathcal{A}} a \times \Pr(\tilde{A}_m = a, G_m = 1) - \sum_{b\in\mathcal{B}} b^\top \times \Pr(\tilde{B}_m = b, G_m = 1) \times \gamma \geq 0, \tag{C.5}$$

where $\mathcal{A}$ and $\mathcal{B}$ are the supports of $\tilde{A}_m$ and $\tilde{B}_m$, respectively. Let $\mathcal{S}_{\theta,\gamma}(X^{\oplus}, W^{\oplus}, \text{MS}, \eta)$ be the random closed set of PSNE networks.[2] If our model is correctly specified, then the observed realisation of G is associated with realisations of $X^{\oplus}, W^{\oplus}, \text{MS}, \eta$ from the truncated support $\{(x^{\oplus}, w^{\oplus}, ms, \bar{\eta}) \in \text{Supp}_{X^{\oplus},W^{\oplus},\text{MS},\eta} : \mathcal{S}_{\theta,\gamma}(x^{\oplus}, w^{\oplus}, ms, \bar{\eta}) \neq \emptyset\}$. Therefore, it holds that:

$$\begin{aligned}&\mathbb{P}(\tilde{A}_m = a, G_m = 1) = \Pr(\tilde{A}_m = a, G_m = 1 | \mathcal{S}_{\theta,\gamma}(X^{\oplus}, W^{\oplus}, \text{MS}, \eta) \neq \emptyset)\\&= \frac{\Pr(\tilde{A}_m = a, G_m = 1, \mathcal{S}_{\theta,\gamma}(X^{\oplus}, W^{\oplus}, \text{MS}, \eta) \neq \emptyset)}{\Pr(\mathcal{S}_{\theta,\gamma}(X^{\oplus}, W^{\oplus}, \text{MS}, \eta) \neq \emptyset)} = \frac{\Pr(\tilde{A}_m = a, G_m = 1)}{\Pr(\mathcal{S}_{\theta,\gamma}(X^{\oplus}, W^{\oplus}, \text{MS}, \eta) \neq \emptyset)}.\end{aligned} \tag{C.6}$$

In turn, we can write:

$$\begin{aligned}\Pr(\tilde{A}_m = a, G_m = 1) &= \mathbb{P}(\tilde{A}_m = a, G_m = 1) \times \Pr(\mathcal{S}_{\theta,\gamma}(X^{\oplus}, W^{\oplus}, \text{MS}, \eta) \neq \emptyset),\\ \Pr(\tilde{B}_m = b, G_m = 1) &= \mathbb{P}(\tilde{B}_m = b, G_m = 1) \times \Pr(\mathcal{S}_{\theta,\gamma}(X^{\oplus}, W^{\oplus}, \text{MS}, \eta) \neq \emptyset).\end{aligned} \tag{C.7}$$

We plug (C.7) in (C.5) and obtain:

$$\Pr(\mathcal{S}_{\theta,\gamma}(X^{\oplus}, W^{\oplus}, \text{MS}, \eta) \neq \emptyset) \times [\mathbb{E}_{\mathbb{P}}(G_m \tilde{A}_m) - \mathbb{E}_{\mathbb{P}}(G_m \tilde{B}_m^{\top})\gamma] \geq 0, \tag{C.8}$$

which is equivalent to

$$\mathbb{E}_{\mathbb{P}}(G_m \tilde{A}_m) - \mathbb{E}_{\mathbb{P}}(G_m \tilde{B}_m^{\top})\gamma \geq 0. \tag{C.9}$$

Hence, the identified set associated with $\mathbb{P}$ is:

$$\Gamma_I \coloneqq \Big\{\gamma \in \Gamma : \ \mathbb{E}_{\mathbb{P}}(G_m \tilde{A}_m) - \mathbb{E}_{\mathbb{P}}(G_m \tilde{B}_m^{\top})\gamma \geq 0\Big\}, \tag{C.10}$$

as in (C.4).

D Computing the first-stage moment inequalities

We provide some directions on how to compute $\Pi_f^e(G_{(+ab),f}, G_{-f}; \theta) - \Pi_f^e(G_f, G_{-f}; \theta)$ and $\text{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma)$ entering (10). A similar procedure can be followed to compute (11).

First, we compute $\text{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma)$. If none of cities a and b are firm f's hubs, then $\text{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma) = \gamma_{1,f} + \eta_{ab,f}$. If only city a (resp. b) is one of firm f's hubs, then $\text{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma) = \gamma_{1,f} + \gamma_{2,f} \times ((D_{a,f}+1)^2 - D_{a,f}^2) + \eta_{ab,f}$ (resp. $\text{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma = \gamma_{1,f} + \gamma_{2,f} \times ((D_{b,f}+1)^2 - D_{b,f}^2) + \eta_{ab,f})$, where $D_{a,f}$ (resp. $D_{b,f}$) is the number of spokes of hub a (resp. b). If both cities a and b are firm f's hubs, then $\text{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma) =$

[2] For the formal definition of a random closed set, see Molchanov and Molinari (2018) and Molinari (2020).

$\gamma_{1,f} + \gamma_{2,f} \times ((D_{a,f}+1)^2 - D^2_{a,f}) + \gamma_{2,f} \times ((D_{b,f}+1)^2 - D^2_{b,f})$.

Second, we determine the realisations of the second-stage shocks used to evaluate the airlines' expected variable profits. In particular, from the vector of second-stage estimates, $\hat{\theta}$, we compute the second-stage shocks for each product offered using the BLP inversion. For each airline f, we compute the mean and variance of the second-stage shocks just obtained and denote them by μ_f and Σ_f respectively. For each potential product of each airline f, we take 100 random draws from a normal distribution with mean μ_f and variance Σ_f. We store all such draws in a matrix Ξ.

Third, we compute the expected variable profits of airline f under $(G_{(+ab),f}, G_{-f})$. To do so, we update the list of products offered by firm f, by adding direct flights between cities a and b. Further, note that setting $G_{ab,f} = 1$ creates a "domino effect" in neighbouring markets, due to the possibility for airline f to offer one-stop flights and the presence of spillovers in entry across markets. Specifically, if a is one of firm f's hubs, then we add one-stop flights, via a, between b and all cities d such that $G_{da,f} = 1$. Similarly, if b is one of firm f's hubs, then we add one-stop flights, via b, between a and all cities d such that $G_{db,f} = 1$. We update the matrices of product covariates by including the observed demand and marginal cost shifters of the new products. We also update the covariates (namely, "*Nonstop Origin*" and "*Connections*") of the pre-existing products that are affected by the new products. Let $\mathcal{M}_{ab,f}$ be the list of markets containing either new products or products with modified covariates. For each market $m \in \mathcal{M}_{ab,f}$, we let the firms reoptimise their prices by iterating on the F.O.C.s in (4), for every draw of the second-stage shocks stored in the matrix Ξ.[3] We compute the variable profits of airline f, average across draws, and get the simulated expected variable profits of airline f, which we denote by $\sum_{m \in \mathcal{M}_{ab,f}} \Pi^e_{f,m}(G_{(+ab),f}, G_{-f}; \theta)$. We implement a similar procedure to compute the expected variable profits of airline f in each markets $m \in \mathcal{M}_{ab,f}$ under G, which we denote by $\sum_{m \in \mathcal{M}_{ab,f}} \Pi^e_{f,m}(G_f, G_{-f}; \theta)$. Lastly, we calculate $\Pi^e_f(G_{(+ab),f}, G_{-f}; \theta) - \Pi^e_f(G_f, G_{-f}; \theta) = \sum_{m \in \mathcal{M}_{ab,f}} \Pi^e_{f,m}(G_{(+ab),f}, G_{-f}; \theta) - \sum_{m \in \mathcal{M}_{ab,f}} \Pi^e_{f,m}(G_f, G_{-f}; \theta)$.

E Bounds under many-link deviations

In this section, we show that many-link deviations do not provide a substantial improvement in the bounds. In particular, we show that two-link deviations generate many redundant inequalities compared to those generated by the one-link deviations.

[3] We have decided to use the F.O.C.s in (4) as a contraction mapping. While we do not formally prove that (4) is indeed a contraction mapping, we have found that the resulting price vector does not change when starting from different values and that the mapping converges in all the cases considered.

Adding links to the factual network

Consider markets $\{a, b\}$ and $\{c, d\}$ that are not served by airline f with direct flights (i.e., $G_{ab,f} = G_{cd,f} = 0$). From the revealed preference principle, it holds that

$$\Pi_f^e(G_{(+ab),f}, G_{-f}; \theta) - \Pi_f^e(G_f, G_{-f}; \theta) \leq \text{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma), \tag{E.1}$$

$$\Pi_f^e(G_{(+cd),f}, G_{-f}; \theta) - \Pi_f^e(G_f, G_{-f}; \theta) \leq \text{FC}_f(G_{(+cd),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma), \tag{E.2}$$

$$\Pi_f^e(G_{(+ab,+cd),f}, G_{-f}; \theta) - \Pi_f^e(G_f, G_{-f}; \theta) \leq \text{FC}_f(G_{(+ab,+cd),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma). \tag{E.3}$$

(E.1) and (E.2) are taken into account by our identification methodology, as they refer to one-link deviations. (E.3) is ignored by our identification methodology, as it refers to a two-link deviation. In what follows, we show that if markets $\{a, b\}$ and $\{c, d\}$ are non-hub markets for airline f and have no cities in common, or they share a hub endpoint, then (E.1) and (E.2) imply (E.3). Hence, (E.3) is redundant.

First, consider the case where markets $\{a, b\}$ and $\{c, d\}$ are non-hub markets for airline f and have no cities in common. Given our fixed cost equation, it holds that

$$\text{FC}_f(G_{(+ab,+cd),f}, \eta_f; \gamma) - \text{FC}_f(G_{(+cd),f}, \eta_f; \gamma) = \text{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma).$$

Therefore, the right-hand-side of (E.3) is equal to

$$\begin{aligned} \text{FC}_f(G_{(+ab,+cd),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma) =& \text{FC}_f(G_{(+cd),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma) \\ &+ \text{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \text{FC}_f(G_f, \eta_f; \gamma). \end{aligned} \tag{E.4}$$

Observe that the left-hand-side of (E.3) can be rewritten as

$$\Pi_f^e(G_{(+ab,+cd),f}, G_{-f}; \theta) - \Pi_f^e(G_{(+cd),f}, G_{-f}; \theta) + \Pi_f^e(G_{(+cd),f}, G_{-f}; \theta) - \Pi_f^e(G_f, G_{-f}; \theta).$$

Furthermore, from our second-stage estimates, it generally holds that

$$\Pi_f^e(G_{(+ab,+cd),f}, G_{-f}; \theta) - \Pi_f^e(G_{(+cd),f}, G_{-f}; \theta) \leq \Pi_f^e(G_{(+ab),f}, G_{-f}; \theta) - \Pi_f^e(G_f, G_{-f}; \theta). \tag{E.5}$$

In other words, adding an independent edge $\{a, b\}$ to the counterfactual network $G_{(+cd),f}$ does not tend to generate more expected variable profits than adding it to the actual network G_f. In fact, adding $\{a, b\}$ to $G_{(+cd),f}$ increases expected variable profits due to two effects. First, the demand in market $\{a, b\}$ increases because the passengers of market $\{a, b\}$ can now fly directly between a and b instead of flying through a hub of f which is neither c nor d (recall the variable "*Indirect*" entering the demand function). Second, the demand in markets having a or b as endpoints is increased by adding the direct service between a and b (recall the variable "*Nonstop Origin*" entering the demand function). From Table 2 (demand panel) we can see that the first effect dominates the second: flying direct increases utility by 1.794; adding *one* direct connection increases utility by 0.00868. In turn, through (E.1), (E.2) and (E.5), we see that

$$\begin{aligned}&\Pi^e_f(G_{(+ab,+cd),f}, G_{-f}; \theta) - \Pi^e_f(G_{(+cd),f}, G_{-f}; \theta) + \Pi^e_f(G_{(+cd),f}, G_{-f}; \theta) - \Pi^e_f(G_f, G_{-f}; \theta) \\ &\leq \mathrm{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \mathrm{FC}_f(G_f, \eta_f; \gamma) + \mathrm{FC}_f(G_{(+cd),f}, \eta_f; \gamma) - \mathrm{FC}_f(G_f, \eta_f; \gamma).\end{aligned} \tag{E.6}$$

Hence, by combining (E.4) and (E.6), (E.3) is verified.

Second, consider the case where markets $\{a, b\}$ and $\{c, d\}$ share a hub endpoint. For instance suppose $a = c$ and a is a hub. Then,

$$\begin{aligned}\mathrm{FC}_f(G_{(+ab,+cd),f}, \eta_f; \gamma) - &\mathrm{FC}_f(G_f, \eta_f; \gamma) \\ &= \mathrm{FC}_f(G_{(+ab),f}, \eta_f; \gamma) - \mathrm{FC}_f(G_f, \eta_f; \gamma) \\ &+ \mathrm{FC}_f(G_{(+cd),f}, \eta_f; \gamma) - \mathrm{FC}_f(G_f, \eta_f; \gamma) \\ &+ \gamma_2(2D_{a,f} + 3),\end{aligned}$$

where $D_{a,f}$ is the number of hub a's spokes in the factual network G_f (mean 20 in the dataset). Again, given our second-stage estimates, it generally holds that

$$\big(\Pi^e_f(G_{(+ab,+cd),f}, G_{-f}; \theta) - \Pi^e_f(G_{(+cd),f}, G_{-f}; \theta)\big) - \big(\Pi^e_f(G_{(+ab),f}, G_{-f}; \theta) - \Pi^e_f(G_f, G_{-f}; \theta)\big)$$

is small, compared to $\gamma_2(2D_{a,f} + 3)$. (E.5) is not always satisfied because adding $\{a, b\}$ and $\{a, d\}$ creates opportunities to fly from b to d via a. However, in our data, it is always possible to fly from b to d via other hubs in the factual network for the same airline f. As a result, it is reasonable to believe that (E.5) holds for most, if not all, two-link deviations. Therefore, using the same steps as above, we conclude that (E.3) holds.

Removing links from the factual network

Consider the mirror case where markets $\{a, b\}$ and $\{c, d\}$ are served by airline f with direct flights (i.e. $G_{ab,f} = G_{cd,f} = 1$). From the revealed preference principle we can see that

$$\Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(-ab),f}, G_{-f}; \theta) \geq \mathrm{FC}_f(G_f, \eta_f; \gamma) - \mathrm{FC}_f(G_{(-ab),f}, \eta_f; \gamma), \tag{E.7}$$

$$\Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(-cd),f}, G_{-f}; \theta) \geq \mathrm{FC}_f(G_f, \eta_f; \gamma) - \mathrm{FC}_f(G_{(-cd),f}, \eta_f; \gamma), \tag{E.8}$$

$$\Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(-ab,-cd),f}, G_{-f}; \theta) \geq \mathrm{FC}_f(G_f, \eta_f; \gamma) - \mathrm{FC}_f(G_{(-ab,-cd),f}, \eta_f; \gamma). \tag{E.9}$$

(E.7) and (E.8) are taken into account by our identification methodology, as they refer to one-link deviations. (E.9) is ignored by our identification methodology, as it refers to a two-link deviation. By following the steps above, it is possible to show that, in most of the cases, (E.9) is redundant.

F Inference on the demand and supply parameters

We conduct inference on θ via GMM under the assumption that the number of markets goes to infinity. Formally, we consider the moment conditions of Section 4.1 and use their sample analogues to construct a GMM objective function which should be minimised with respect to $\theta \in \Theta$:

$$Q(\theta) = D(\theta)' A D(\theta), \tag{F.1}$$

where

$$D(\theta) \coloneqq \begin{pmatrix} \frac{1}{|\mathcal{J}|} \sum_{m \in \mathcal{M}} \sum_{j \in \mathcal{J}_m} [\tau_{j,m}(X_m^{\oplus}, W_m^{\oplus}, \mathrm{MS}_m, s_m^{\oplus}, P_m^{\oplus}, G; \theta) \times z_{j,m,1}(X_m^{\oplus}, W_m^{\oplus})] \\ \frac{1}{|\mathcal{J}|} \sum_{m \in \mathcal{M}} \sum_{j \in \mathcal{J}_m} [\tau_{j,m}(X_m^{\oplus}, W_m^{\oplus}, \mathrm{MS}_m, s_m^{\oplus}, P_m^{\oplus}, G; \theta) \times z_{j,m,2}(X_m^{\oplus}, W_m^{\oplus})] \\ \vdots \\ \frac{1}{|\mathcal{J}|} \sum_{m \in \mathcal{M}} \sum_{j \in \mathcal{J}_m} [\tau_{j,m}(X_m^{\oplus}, W_m^{\oplus}, \mathrm{MS}_m, s_m^{\oplus}, P_m^{\oplus}, G; \theta) \times z_{j,m,\mathrm{L}}(X_m^{\oplus}, W_m^{\oplus})] \end{pmatrix},$$

$\mathcal{J} \coloneqq \cup_{m \in \mathcal{M}} \mathcal{J}_m$ is the set of all offered products, and A is an appropriate $2\mathrm{L} \times 2\mathrm{L}$ weighting matrix. In particular, A is computed via the usual two-step procedure: first, we estimate the parameters using the optimal weighting matrix under conditional homoskedasticity; second, we use the obtained estimates to construct the optimal weighting matrix under conditional heteroskedasticity and re-estimate the parameters.

Note that we estimate the demand and supply sides jointly. We could also estimate the demand and supply sides separately by following a two-step procedure: first estimating the demand parameters; then using these estimates to calculate the mark-ups; finally regressing the resulting marginal costs on the observed marginal cost shifts to obtain the supply parameters. We have chosen to estimate the demand and supply sides together because it allows us to take into account the potential correlation between the demand and supply moments and thus obtain more precise estimates, as discussed in Berry et al. (1995). Moreover, since we have a computationally "light" demand specification, the additional cost of estimating the demand and supply sides jointly is negligible.

G Inference on the fixed cost parameters

G.1 Writing (21) as a linear optimisation problem with exponential cone constraints

In what follows, we show that (21) is a linear optimisation problem with exponential cone constraints. First, we simplify the notation of (21) and write it as

$$\begin{aligned} \delta(q, \Gamma_I^\alpha) &:= \sup_{\gamma \in \Gamma} q^\top \gamma, \\ &\text{s.t.} \ \sum_{r=1}^{\mathrm{R}} f_\alpha(b_r \gamma - a_r) - \mathrm{R} \log(2)/\alpha \le 0, \end{aligned} \tag{G.1}$$

where b_r stands for $\mathbb{E}(Z_{r,m} B_m)$ and a_r for $\mathbb{E}(Z_{r,m} A_m)$. Both quantities can be estimated consistently from their empirical analogue. Second, observe that

$$\sum_{r=1}^{\mathrm{R}} f_\alpha(b_r \gamma - a_r) - \mathrm{R} \log(2)/\alpha \le 0 \tag{G.2}$$

$$\Leftrightarrow \log(1 + \exp(\alpha(b_r \gamma - a_r))) \le t_r \text{ for } r = 1, \dots, \mathrm{R} \text{ and } \sum_{r=1}^{\mathrm{R}} t_r \le \mathrm{R} \log 2 \tag{G.3}$$

$$\Leftrightarrow \exp(-t_r) + \exp(-t_r + \alpha(b_r \gamma - a_r)) \le 1 \text{ for } r = 1, \dots, \mathrm{R} \text{ and } \sum_{r=1}^{\mathrm{R}} t_r \le \mathrm{R} \log 2 \tag{G.4}$$

Therefore, (G.1) is equivalent to

$$\max q^\top \gamma + \sum_{r=1}^{R} 0.t_r + 0.u_r + 0.v_r,$$

under the constraints

$$\begin{aligned} \sum_{r=1}^{\mathrm{R}} t_r &\le R \log 2, \\ u_r + v_r &\le 1, \ r = 1, \dots, \mathrm{R}, \\ (v_r, 1, -t_r) &\in K_{\exp} \ r = 1, \dots, \mathrm{R}, \\ (u_r, 1, \alpha(b_r \gamma - a_r) - t_r) &\in K_{\exp} \ r = 1, \dots, \mathrm{R}. \end{aligned}$$

The exponential cone $K_{\exp}$ is a convex subset of $\mathbb{R}^3$ such that

$$K_{\exp} = \{(x_1, x_2, x_3) : \ x_1 \ge x_2 \exp(x_3/x_2); x_2 > 0\} \cup \{(x_1, 0, x_3), \ x_1 \ge 0, x_3 \ge 0\}.$$

The constraints above ensure, in particular, that for any r, $v_r \geq \exp(-t_r)$ and $u_r \geq \exp(-t_r + \alpha(b_r\gamma - a_r))$, and, therefore, ensure (G.4).

See `https://docs.mosek.com/modeling-cookbook/expo.html#softplus-function` for further details.

G.2 Constructing a confidence interval for a component of γ

Suppose we want to construct a confidence interval for a specific linear combination of components of γ, $c^\top\gamma$. Let $q = c/\|c\|$. By Theorem 2,

$$\sqrt{\mathrm{M}}\left(\hat{\delta}(q;\Gamma_I^\alpha) - \delta(q;\Gamma_I^\alpha)\right) \xrightarrow[\mathrm{M}\to\infty]{d} Z_\alpha(q).$$

The optimisation routine detailed in Section G.1 gives us the unique point, γ_q, which achieves the maximum of $c^\top\gamma$ on $\widehat{\Gamma}_I^\alpha$. Let λ_q be the Lagrange multiplier solving

$$\lambda_q \nabla g_\alpha(\gamma_q) = q,$$

where

$$\nabla g_\alpha(\gamma_q) = \sum_{r=1}^{\mathrm{R}} b_r \frac{\exp\left(\alpha\left[b_r^\top\gamma - a_r\right]\right)}{1 + \exp\left(\alpha\left[b_r^\top\gamma - a_r\right]\right)}.$$

Let $W_r(\gamma_q)$ be a random normal variable with variance equal to the asymptotic variance of $\frac{1}{\mathrm{M}}\sum_{m=1}^{\mathrm{M}}(Z_{r,m}B_m^\top\gamma_q - Z_{r,m}A_m)$. In turn, we can compute the variance of $Z_\alpha(q)$, which is the variance of a centered normal random variable. We denote it $v_\alpha(q)$. The quantity

$$c^\top\gamma_q + \|c\| n_{1-\beta}\sqrt{v_\alpha(q)},$$

is the upper bound of the $1-\beta$ confidence interval for $c^\top\gamma$, where $n_{1-\beta}$ is the $1-\beta$ quantile of the standard normal distribution.

Similarly, let $-q = -c/\|c\|$ and γ_{-q} be the point which achieves the maximum of $-c^\top\gamma$ on $\widehat{\Gamma}_I^\alpha$. Let λ_{-q} be the Lagrange multiplier solving

$$\lambda_{-q}\nabla g_\alpha(\gamma_{-q}) = -q.$$

As above, we can compute the variance of $Z_\alpha(-q)$ and denote it $v_\alpha(-q)$. The quantity

$$c^\top\gamma_{-q} - \|c\| n_{1-\beta}\sqrt{v_\alpha(-q)},$$

is the lower bound of the $1-\beta$ confidence interval for $c^\top\gamma$.

Note that, following Stoye (2009), we can adapt the choice of the quantile to handle near to point-identified cases.

G.3 Drawing points from the confidence region for γ

In this section, we outline the steps to draw points from the confidence region for the true value γ_0 in order to run our counterfactual analysis.

1. We look for an interior point γ_c in $\widehat{\Gamma}_I^\alpha$. This is known in the convex optimization literature as the Chebyshev center of a polyhedron (Boyd and Vandenberghe, 2004, page 148). Interestingly, it can be solved by linear programming:

$$\begin{aligned}
&\max_{r\geq 0} r, \\
&\text{s.t. } \frac{1}{\mathrm{M}}\sum_{m=1}^{\mathrm{M}}(-Z_{r,m}B_m^\top\gamma + Z_{r,m}A_m) + r\|\frac{1}{\mathrm{M}}\sum_{m=1}^{\mathrm{M}} Z_{r,m}B_m\|_2 \leq 0, \\
&r = 1, \dots, \mathrm{R}.
\end{aligned}$$

2. Draw a random direction q on the unit sphere and find the frontier point $\gamma_q = \gamma_c + r_q q$ of $\widehat{\Gamma}_I^\alpha$ ($r_q \geq 0$). Again, this is a linear program.

3. Calculate the outer normal vector of $\widehat{\Gamma}_I^\alpha$ at γ_q. This is the direction q' such that $\delta(q', \widehat{\Gamma}_I^\alpha) = q'^\top\gamma_q$. It can be done analytically by calculating the gradient of $g_\alpha(\cdot)$ at γ_q.

4. Calculate the variance $V_\alpha(q')$ of $Z_\alpha(q')$ using Theorem 2.

5. The point $f_q = \gamma_q + \sqrt{V_\alpha(q')}n_{1-\beta}q'$ is a frontier point of the (conservative) confidence region $CR_{1-\beta}(\gamma_0)$ (drawn from γ_q in direction q').

6. Draw a norm l uniformly on $[0, 1]$.

7. Pick the point $\gamma_c + lf_q$ which belongs to $CR_{1-\beta}(\gamma_0)$.

Figure G.1 illustrates the sequence.

H Empirical application

H.1 Data

Table H.1 lists the airlines' hubs. Table H.2 reports the airlines belonging to the groups LCC and Other.

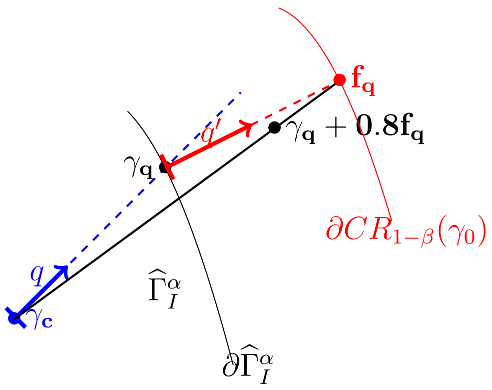

Figure G.1: Drawing from the confidence region

AA	**DL**	**UA**	**US**	**WN**
Dallas	Atlanta	Washington DC	Charlotte	Washington DC
New York	Cincinnati	Denver	Washington DC	Denver
Los Angeles	Detroit	Houston	Philadelphia	Houston
Miami	New York	New York	Phoenix	Las Vegas
Chicago	Memphis	Los Angeles		Chicago
	Minneapolis-Saint Paul	Chicago		Phoenix
	Salt Lake City	San Francisco		

Table H.1: Hubs of the legacy carriers and focus cities of Southwest Airlines in 2011.

LCC	**Other**
Frontier Airlines	AirTran Airways
Alaska Airlines	USA3000 Airlines
Spirit Airlines	
Jetblue Airlines	
Virgin America	
Sun County Airlines	
Allegiant Air	

Table H.2: Airlines in the categories LLC and Other.

H.2 Instruments

Table H.3 lists the instruments we use in the estimation of the fixed cost parameters. Table H.4 lists the instruments we use in the estimation of the demand and supply parameters.

Table H.3: First-stage instruments.

	$Z_{r,(-ab),f} = 1$ if
All firms	$\{a, b\}$ is not a hub market and has been continuously served since 1979 Q1
AA	$\{a, b\}$ is a hub market with size above 6 million
DL	$\{a, b\}$ is a hub market with size above 6 million
UA	$\{a, b\}$ is a hub market with size above 6 million
US	$\{a, b\}$ is a hub market with size above 5 million
WL	$\{a, b\}$ is a hub market with size above 6 million
	$Z_{r,(+ab),f} = 1$ if
All firms	$\{a, b\}$ is not a hub market and a competitor has a hub at a or b
AA	$\{a\}$ is a hub and $\{b\}$ is closer to at least 2 other AA hubs
DL	$\{a\}$ is a capacity constrained hub and $\{b\}$ is closer to all other DL hubs
UA	$\{a\}$ is a hub and $\{b\}$ is closer to at least 2 other UA hubs
US	$\{a\}$ is a capacity constrained hub and $\{b\}$ is closer to all other US hubs
WL	$\{a\}$ is a capacity constrained hub and $\{b\}$ is closer to all other WN hubs

Note: Capacity constrained airports are defined to be airports in need of capacity improvements according to the Federal Aviation Administration's FACT3 report. Note that this definition is not equivalent to an airport being slot constrained.

Table H.4: Second-stage instruments.

Number of firms present in the market
Number of itineraries offered in the market
Number of products offered in the market
Indicator for destination being a hub
Indicator for the market being a monopoly
Number of rival firms offering direct flights in the market
Square of the number of rival firms offering direct flights in the market

H.3 Results from demand and supply

Table H.5 shows the estimated variable profits, prices, marginal costs, and markups at the firm level. For each airline, the first, second, and third rows contain quantities averaged over all products, direct flights and one-stop flights respectively. The fourth and fifth rows contain quantities averaged over direct flights where at least one of the endpoints is a hub, and direct flights where no endpoint is a hub. We can see that airlines charge higher markups on direct flights compared to one-stop flights, which is in line with the fact that consumers prefer to take direct flights (see "*Indirect*" in Table 2, demand panel). The legacy carriers charge higher markups on direct flights where at least one of the endpoints is a hub than on direct flights where no endpoint is a hub, suggesting the existence of a hub premium. This hub premium may be due to the fact that consumers value flying from dense hubs (see "*Nonstop Origin*" in Table 2, demand panel) or to fixed costs due to congestion effects at hubs (see Table 4). While American Airlines, US Airways and Southwest Airlines have lower marginal costs for direct flights, the opposite is true for Delta and United Airlines.[4] The marginal cost of Southwest Airlines is lower than the marginal cost of the legacy carriers. For direct flights, the difference is quite substantial. For one-stop flights, Southwest Airlines' advantage is small, consistent with the fact that Southwest Airlines uses focus cities rather than hubs. Therefore, the marginal cost savings of offering one-stop flights (see "*Connections*" and "*Indirect*" in Table 2, Supply panel) may be less pronounced as not all the features of traditional hubs are used.

H.4 Estimated shares of the variable costs over the operating costs

We compute the estimated share of the variable costs over the operating costs. The former are obtained as marginal costs times number of passengers. The latter are defined as the sum of the variable costs and fixed costs, without considering the congestion costs. Table H.6 reports this share for each airline based on our results. We compare such shares with estimates from the FAA for 2018 based on administrative data (Table H.7) and observe similar orders of magnitude.

I Counterfactuals

I.1 Descriptions of the counterfactual algorithm

The possibility of multiple PSNE networks raises the question of how to obtain counterfactuals when airlines are allowed to reoptimise their networks and prices. Although the data tell us which equilibrium was played in the past, they do not tell us which equilibrium will be chosen by the players once we change the environment. Previous literature has suggested several ways of solving this problem. For example, the analyst could enumerate all possible equilibria and report some summary measures of the resulting range of counterfactuals (Eizenberg, 2014). Alternatively, the analyst could implement a learning algorithm and use it to select a probability distribution of possible equilibria (Lee and

[4]Note that the fact that American Airlines, US Airways and Southwest Airlines have lower marginal costs on direct flights does not contradict the negative sign of the coefficient on "*Connections*" in Table 2. In fact, recall that the results in Table 2 should be interpreted *ceteris paribus*. Instead, the results in Table H.5 are obtained by averaging over all itineraries, including those with different characteristics.

Table H.5: Profits by firms.

	Profits (100k)	Price	Marginal cost	Markup	Lerner Index
AA					
All	1.78	453.36	335.20	118.16	0.28
Direct	13.77	402.37	277.42	124.94	0.32
One-stop	0.39	459.26	341.89	117.38	0.27
Direct, hub endpoint	15.06	402.75	276.66	126.09	0.33
Direct, non-hub endpoints	2.00	398.87	284.48	114.40	0.30
DL					
All	1.41	436.45	310.40	126.05	0.31
Direct	12.31	463.26	321.03	142.23	0.33
One-stop	0.33	433.80	309.35	124.45	0.31
Direct, hub endpoint	13.49	482.67	336.83	145.84	0.32
Direct, non-hub endpoints	4.47	334.75	216.44	118.31	0.38
UA					
All	1.25	445.56	328.43	117.13	0.28
Direct	9.17	458.50	334.97	123.53	0.29
One-stop	0.20	443.85	327.56	116.28	0.28
Direct, hub endpoint	11.03	456.82	332.24	124.58	0.29
Direct, non-hub endpoints	2.17	464.88	345.33	119.55	0.29
US					
All	1.30	453.43	336.77	116.67	0.27
Direct	8.99	407.34	275.17	132.17	0.35
One-stop	0.35	459.10	344.34	114.76	0.26
Direct, hub endpoint	10.42	418.96	282.96	136.00	0.35
Direct, non-hub endpoints	3.95	366.22	247.58	118.64	0.36
WN					
All	2.79	419.43	299.51	119.92	0.31
Direct	12.09	365.14	237.09	128.05	0.38
One-stop	0.23	434.40	316.73	117.67	0.29
Direct, hub endpoint	16.49	362.34	233.95	128.39	0.38
Direct, non-hub endpoints	8.88	367.19	239.39	127.80	0.38

Note: Quantities are in USD.

Table H.6: Estimated shares of the variable costs over the operating costs.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
AA	80.5%	81.5%	82.4%	82.5%	83.7%	84.8%
DL	78.2%	79.3%	80.2%	80.4%	81.7%	82.9%
UA	75.0%	76.2%	77.2%	77.4%	78.8%	80.1%
US	73.9%	75.1%	76.2%	76.4%	77.9%	79.2%
WN	66.3%	67.7%	69.0%	69.3%	71.0%	72.6%

Table H.7: Passenger Air Carriers Filing Schedule P-5.2 Operating and Fixed Costs per Block Hours.

Aircraft Category	Cost per Block Hour								
	Fuel and Oil	Mainte-nance	Crew	Total Variable	Deprec.	Rentals	Other	Total Fixed	Share Variable
Wide-body more than 300 seats	\$5,411	\$1,331	\$2,356	\$9,097	\$845	\$406	\$5	\$1,254	87.9%
Wide-body 300 seats and below	\$4,080	\$1,289	\$1,857	\$7,227	\$685	\$366	\$8	\$1,058	87.2%
Narrow-body more than 160 seats	\$2,054	\$718	\$1,152	\$3,925	\$355	\$217	\$10	\$582	87.1%
Narrow-body 160 seats and below	\$1,741	\$737	\$1,034	\$3,512	\$306	\$215	\$12	\$533	86.8%
RJ more than 60 seats	\$115	\$431	\$444	\$991	\$131	\$252	\$14	\$397	71.4%
RJ 60 seats and below	\$92	\$479	\$470	\$1,041	\$58	\$227	\$8	\$293	78.0%
Turboprop more than 60 seats	\$0	\$880	\$360	\$1,241	\$439	\$103	\$2	\$544	69.5%
All Aircraft	\$1,681	\$727	\$1,012	\$3,420	\$314	\$239	\$11	\$564	85.8%

Source: FAA, `https://www.faa.gov/regulations_policies/policy_guidance/benefit_cost`, Section 4 of the Benefit-Cost analysis, Table 4-6.

Pakes, 2009; Wollmann, 2018). The first approach is not computationally feasible in our setting, due to the large number of markets and the presence of entry spillovers. Therefore, we follow the second approach. We fix an order of markets and firms. For a given value of the parameters, the first firm in the first market best responds to its competitors in terms of entry and pricing decisions. The second firm similarly best responds, taking into account the best response of the first firm. The third company also best responds, taking into account the best responses of the first and second companies. The algorithm cycles through the firms and markets until no airline wishes to deviate. The procedure is repeated for 50 draws of parameter values from the estimated identified set of first-stage parameters. For each parameter value, we consider four market orderings. In the first ordering, we rank the markets according to which hub is involved, whether the market is served by the merged firm, the size of the merged firm's operations at the endpoints, and the market size(ordering A). In the second ordering, we reverse this ranking (ordering B). In the third and fourth orderings, werank markets randomly (orderings C and D). For each of the four market orderings, we consider two firm orderings: AA-DL-UA-WL (ordering 1) and the reverse (ordering 2). This procedure generates a distribution of possible equilibria over 400 (i.e., $50 \times 4 \times 2$) counterfactual runs. In the tables of Section 8.2, we report the minimum, maximum, and median changes in the relevant outcomes under such distribution.

The remainder of the section illustrates the details of the counterfactual algorithm. In particular, we explain the algorithm implemented to simulate the merger under the *Networks vary - w/o remedies* scenario, given an order of markets and firms and a value of the parameters. The algorithm is structured in the following steps:

1. Latent variables. We determine the realisations of the latent variables that are needed to evaluate the airlines' profits. In particular, from the vector of second-stage estimates, $\hat{\theta}$, we compute the second-stage shocks for each product offered by the airlines before the merger, via BLP inversion. For each airline f, we compute the mean and variance of the second-stage shocks and denote them by μ_f and Σ_f, respectively. When computing μ_f and Σ_f for the merged airline, we consider the second-stage shocks associated with all the products offered by the merging firms before the merger. If both American Airlines and US Airways offer a given itinerary before the merger, then we take the mean value of the second-stage shocks of the two pre-merger products. For each potential product of every airline f, we take 100 random draws from a normal distribution with mean μ_f and variance Σ_f. We store all such draws in a matrix Ξ. Further, for each market $\{a, b\}$ and airline f, we impute the fixed cost shock $\eta_{ab,f}$ as explained in Section I.2.

2. Initial state. At the start, all firms except the merged entity are assigned their pre-merger networks and products. The merged entity is assigned the network resulting from combining the pre-merger networks of American Airlines and US Airways. The products initially offered by the merged entity and their observed characteristics are constructed from such merged network. We denote by $G \coloneqq (G_1, \ldots, G_{\text{N-1}})$ the initial networks of the carriers. We let the firms play the simultaneous pricing game described in Section 3.1, for each draw of the second-stage shocks stored in the matrix Ξ. We save the initial equilibrium prices in a matrix P.

3. Iterations. We take the first firm f in the first market $\{a, b\}$ and let it play its best response as follows. Suppose, for instance, that the initial network G_f is characterised by $G_{ab,f} = 0$. First, we compute airline f's expected variable profits under $(G_{(+ab),f}, G_{-f})$. To do so, we update the list of products offered by firm f, by adding direct flights between cities a and b. Further, note that setting $G_{ab,f} = 1$ creates a "domino effect" in neighbouring markets, due to the possibility for airline f to offer one-stop flights and the presence of spillovers in entry across markets. Hence, if a is one of firm f's hubs, then we add one-stop flights, via a, between b and all cities d such that $G_{da,f} = 1$. Similarly, if b is one of firm f's hubs, then we add one-stop flights, via b, between a and all cities d such that $G_{db,f} = 1$. We update the matrices of product covariates by including the observed demand and marginal cost shifters of the new products. We also update the product covariates (namely, "*Nonstop Origin*" and "*Connections*") of the pre-existing products that are affected by the new products. Let $\mathcal{M}_{ab,f}$ be the list of markets containing either new products or products with modified covariates. For each of these products in every market $m \in \mathcal{M}_{ab,f}$, we let airline f find the best-response price, while holding the other prices in P fixed, for every draw of the second-stage shocks stored in the matrix Ξ. We compute airline f's variable profits, average across draws, and get the simulated airline f's expected variable profits, which we denote by $\sum_{m \in \mathcal{M}_{ab,f}} \Pi^e_{f,m}(G_{(+ab),f}, G_{-f}; \hat{\theta})$. Next, we implement a similar procedure to compute airline f's expected variable profits in each markets $m \in \mathcal{M}_{ab,f}$ under G, which we denote by $\sum_{m \in \mathcal{M}_{ab,f}} \Pi^e_{f,m}(G_f, G_{-f}; \hat{\theta})$. We take the difference between airline f's fixed costs under $(G_{(+ab),f}, G_{-f})$ and G, which is $\text{FC}_f(G_{(+ab),f}, \eta_f; \hat{\gamma}) - \text{FC}_f(G_f, \eta_f; \hat{\gamma}) = \hat{\gamma}_{2,f} \Delta\overline{\text{Q}}_{(+ab),f} + \hat{\gamma}_{1,f} + \eta_{ab,f}$, where $\hat{\gamma}$ is the value of the fixed costs parameters drawn from the estimated identified set and $\eta_{ab,f}$ is the imputed value of the fixed cost shock. Lastly, we compute:

$$\sum_{m \in \mathcal{M}_{ab,f}} \Pi^e_{f,m}(G_{(+ab),f}, G_{-f}; \hat{\theta}) - \sum_{m \in \mathcal{M}_{ab,f}} \Pi^e_{f,m}(G_f, G_{-f}; \hat{\theta}) - (\hat{\gamma}_{2,f} \Delta\overline{\text{Q}}_{(+ab),f} + \hat{\gamma}_{1,f} + \eta_{ab,f}). \tag{I.1}$$

If (I.1) is positive (negative), then the best-response entry of airline f is $G_{ab,f} = 1$ ($G_{ab,f} = 0$). We update G and P and move to the second firm in the first market. We let this firm best respond, while taking into account the first firm's best response. The third firm similarly best responds, while taking into account the first and second firms' best responses, and so on.

4. Stop. We cycle through the firms and markets. When no firm wants to deviate in none of the markets, we stop the procedure. In practice, we have obtained convergence in all the cases considered.

Due to computational costs, the above algorithm does not consider all possible entry deviations by each firm. In fact, it imposes that each firm considers adding/deleting direct flights in one market at a time. Nevertheless, at the rest point of the procedure, the necessary conditions for PSNE that are used in the estimation of the fixed cost parameters hold. Hence, the algorithm provides an equilibrium that is internally consistent with our model. Similar restrictions on the set of admissible deviations are assumed by Eizenberg (2014) and Wollmann (2018).[5]

We also adopt the above algorithm in the merger simulation for the *Networks vary - w/ remedies* and *6. Networks vary - PHX dehubbed* scenarios. However, in scenario *Networks vary - w/ remedies*, we do not allow the merged entity to exit the markets out of Charlotte, New York, Los Angeles, Miami, Chicago, Philadelphia, and Phoenix that were served before the merger by American Airlines or US Airways. In scenario *Networks vary - PHX dehubbed* we delete all flights of the merged entity between Phoenix and non-hub cities and do not allow the merged entity to re-enter those markets.

I.2 Imputation of the fixed cost shocks in the counterfactuals

To perform the counterfactuals, we need a measure of the fixed cost shocks. Different approaches have been taken in the literature. For example, Wollmann (2018) draws the fixed cost shocks from a normal distribution with zero mean and variance equal to a fraction of the variance of the systematic fixed costs. Kuehn (2018) finds, for each market, the range of realisations of the fixed cost shocks generating the observed entry/exit patterns and takes the midpoint. We use a procedure that is similar to Kuehn (2018). We repeat the steps below for each value of γ drawn from the estimated identified set at which we run the counterfactual algorithm. When we observe airline f serving market $\{a, b\}$ with direct flights (i.e., $G_{ab,f} = 1$), we infer that this choice must be profitable, giving us an upper bound for $\eta_{ab,f}$. In fact, let $\Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(-ab),f}, G_{-f}; \theta) - \gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(-ab),f} - \gamma_{1,f} - \eta_{ab,f}$ be the difference between the factual profits of airline f and the profits that airline f would get if deviating to $G_{ab,f} = 0$. By best-response arguments, it must be that $\Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(-ab),f}, G_{-f}; \theta) - \gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(-ab),f} - \gamma_{1,f} - \eta_{ab,f} \geq 0$, i.e., $\eta_{ab,f} \leq \Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(-ab),f}, G_{-f}; \theta) - \gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(-ab),f} - \gamma_{1,f}$. Thus, $\Pi^e_f(G_f, G_{-f}; \theta) - \Pi^e_f(G_{(-ab),f}, G_{-f}; \theta) - \gamma_{2,f}\Delta\overline{\mathrm{Q}}_{(-ab),f} - \gamma_{1,f}$ represents an upper bound for $\eta_{ab,f}$. Next, we collect all the markets where airline f does not enter, that are hub markets (non-hub markets) if market $\{a, b\}$ is a hub market (is not a hub market) for airline f, and that

[5]The networks at the rest point of our algorithm constitute a pairwise stable outcome, in the sense illustrated by Section B. In fact, our algorithm resembles the tâtonnement dynamics discussed by Jackson and Watts (2002), in which agents form or destroy individual connections, taking the remaining network as given and not anticipating future adjustments. Jackson and Watts (2002) show that pairwise stable networks can be achieved by tâtonnement dynamics.

Table I.1: Comparison of merger prediction with data from 2015-2019.

	2011	**Prediction**	**2015**	**2016**	**2017**	**2018**	**2019**	**Mean 15-19**
Segments: AA/US	430	498	445	465	464	478	491	468.6
		[435, 546]						
Segments: Other major airlines	736	689	669	678	698	681	678	680.8
		[606, 710]						

face similar congestion costs. These markets give us a vector of lower bounds for $\eta_{ab,f}$. We take the 2.5th percentile of these lower bounds and use it as a lower bound for $\eta_{ab,f}$. Lastly, we set $\eta_{ab,f}$ equal to the mid-point between the lower and upper bounds. We implement a similar procedure to determine the fixed cost shocks for the markets that are not served by airline f in the data. However, instead of the 2.5th percentile, in that case we take the 97.5th percentile to obtain an upper bound. When simulating the merger, the merged entity gets the mean value of the fixed cost shocks imputed to the merging firms by following the above procedure.

I.3 Comparison with post-merger data

Table I.1 shows a comparison of our *Networks vary - w/ remedies* scenario with post-merger data on the markets served with direct flights by American Airlines and its competitors before and after the merger out of American Airlines and US Airways' hubs. Note that such a comparison is always fraught with difficulties because other changes occurred at the same time the merger was consummated, such as changes in preferences, costs (e.g., a significant drop in the price of kerosene in the 2010s), and other changes in market structure. See, for instance, Bontemps et al. (2022). Nevertheless, our model predicts relatively well the actual entry-exit dynamics. In particular, we correctly predict the post-merger expansion of American Airlines' network and reduction of competitors' networks. In particular, towards 2019, the observed number of markets served with direct flights closely matches the median prediction of our scenarios. Further, the observed number of markets served with direct flights lies within the lower-and upper bounds of our predictions in every year considered.

I.4 Additional tables

Table I.2 shows the hub-level changes in the number of direct flights offered by American Airlines and the other major airlines. The column Av. presence reports the average number of main carriers present across all possible markets out of a given hub. Table I.3 reports the percentage change in prices, marginal costs, and markups of American Airlines and the other major airlines. It distinguishes between direct flights and one-stop flights.

Table I.2: Changes in direct flights offered in the hub markets of AA and US.

	Before			**Merger**								
				w/o remedies			w/ remedies			PHX dehubbed		
	AA/US	Others	Av. presence	AA/US	Others	Av. presence	AA/US	Others	Av. presence	AA/US	Others	Av. presence
AA hubs												
DFW	68	55	1.6	69	57	1.52	68	57	1.52	68	57	1.52
				[66, 73]	[55, 57]	[1.5, 1.59]	[66, 73]	[55, 57]	[1.5, 1.59]	[67, 74]	[54, 57]	[1.49, 1.6]
LAX	28	90	1.51	30	90	1.47	30	90	1.48	34	90	1.52
				[21, 33]	[87, 91]	[1.36, 1.52]	[28, 33]	[87, 91]	[1.44, 1.52]	[22, 35]	[88, 91]	[1.38, 1.56]
ORD	59	129	2.35	63	105	2.05	63	110	2.1	62	108	2.09
				[56, 70]	[73, 116]	[1.68, 2.21]	[60, 71]	[73, 115]	[1.73, 2.23]	[53, 69]	[75, 118]	[1.71, 2.2]
MIA	40	51	1.17	25	52	0.94	42	52	1.13	25	52	0.94
				[15, 44]	[50, 52]	[0.8, 1.16]	[40, 47]	[50, 52]	[1.11, 1.21]	[14, 46]	[50, 52]	[0.79, 1.18]
JFK	41	113	2	58	95	1.85	56	96	1.88	61	97	1.85
				[29, 81]	[49, 113]	[1.55, 2.17]	[43, 81]	[49, 106]	[1.59, 2.16]	[29, 81]	[51, 112]	[1.57, 2.16]
US hubs												
CLT	61	41	1.29	64	41	1.28	65	41	1.29	64	41	1.28
				[43, 69]	[39, 42]	[1.02, 1.35]	[61, 69]	[39, 42]	[1.23, 1.35]	[43, 68]	[39, 42]	[1.04, 1.33]
PHX	41	74	1.49	40	66	1.29	42	65	1.3	8	68	0.93
				[23, 43]	[61, 69]	[1.1, 1.35]	[41, 43]	[59, 68]	[1.22, 1.35]	[8, 8]	[63, 70]	[0.87, 0.95]
DCA	40	130	2.16	81	127	2.52	75	128	2.46	81	127	2.52
				[34, 82]	[123, 133]	[2, 2.6]	[28, 82]	[124, 133]	[1.91, 2.61]	[32, 82]	[124, 133]	[1.98, 2.59]
PHL	52	53	1.33	54	56	1.36	56	55	1.37	55	56	1.37
				[25, 67]	[54, 60]	[1.04, 1.52]	[52, 67]	[54, 56]	[1.32, 1.52]	[28, 66]	[54, 61]	[1.07, 1.49]
Total												
Total	430	736	1.66	491	686	1.6	498	689	1.61	457	693	1.56
				[348, 531]	[607, 720]	[1.43, 1.65]	[435, 546]	[606, 710]	[1.54, 1.67]	[335, 497]	[612, 721]	[1.41, 1.62]

Note: Median outcomes are reported, with minimum and maximum outcome in brackets.

I.5 Inference on counterfactuals

In this section, we report the confidence intervals for the counterfactuals presented in Section 8.2 of the main paper. To construct these confidence intervals, we run the counterfactual algorithm discussed in Section I.1 at 50 draws of parameter values from the 95% confidence region for γ. Section G.3 explains how we take such draws. In particular, Table I.4 reports the confidence intervals for Table 9, Table I.5 reports the confidence intervals for Table 10.

Table I.3: Percentage change in prices, marginal cost, and markups.

	Before	Merger		
		w/o remedies	w/ remedies	PHX dehubbed
AA/US: Direct				
Price	406.24	-4.71	-4.68	-4.72
		[-6.73, -3.67]	[-5.36, -3.53]	[-6.71, -3.60]
Marginal cost	276.70	-10.23	-10.05	-10.06
		[-12.62, -9.66]	[-10.87, -9.43]	[-12.52, -9.28]
Markup	129.54	+7.39	+7.04	+7.30
		[+4.6, +9.41]	[+5.64, +9.39]	[+4.22, +9.23]
Others: Direct				
Price	413.19	+0.56	+0.55	+0.78
		[-0.25, +2.50]	[-0.18, +1.43]	[-0.14, +2.65]
Marginal cost	291.60	+1.39	+1.35	+1.38
		[+0.30, +3.40]	[+0.51, +2.26]	[+0.20, +3.23]
Markup	121.59	-1.44	-1.43	-0.64
		[-1.80, +1.16]	[-1.97, -0.21]	[-1.12, +1.50]
AA/US: One-stop				
Price	466.39	-5.69	-5.42	-6.19
		[-8.15, -5.19]	[-5.90, -4.74]	[-7.62, -5.60]
Marginal cost	351.28	-12.67	-12.43	-12.85
		[-14.08, -11.75]	[-12.91, -10.49]	[-13.31, -10.33]
Markup	115.11	+15.27	+15.44	+13.63
		[+8.22, +17.82]	[+12.01, +18.52]	[+8.07, +16.62]
Others: One-stop				
Price	416.12	+4.05	+3.97	+4.13
		[+3.42, +4.97]	[+3.43, +4.52]	[+3.55, +4.98]
Marginal cost	301.18	+6.00	+5.94	+6.03
		[+5.35, +6.63]	[+5.41, +6.49]	[+5.40, +6.58]
Markup	114.94	-1.25	-1.31	-0.84
		[-1.81, +0.99]	[-2.00, -0.44]	[-1.49, +1.07]

Note: Percentage changes with respect to the pre-merger scenario are reported.

Table I.4: Percentage change in consumer surplus across different scenarios.

	Networks fixed	Networks vary		
		w/o remedies	w/ remedies	PHX dehubbed
Total consumer surplus	+0.08	+0.77	+0.91	-0.67
	[-0.47, +3.40]	[-8.92, +3.47]	[-3.92, +3.84]	[-10.01, +1.79]
New markets	0	45.15	45.02	42.87
		[30.77, 52.29]	[29.47, 53.37]	[23.58, 53.02]
Old markets	+0.08	-5.28	-5.12	-4.67
	[-0.47, +3.40]	[-10.67, -3.97]	[-8.18, -3.94]	[-11.23, -3.32]

Note: Consumer surplus is computed using the log-sum formula and it is in USD 1 million up to constant of integration. Percentage changes with respect to the pre-merger scenario are reported.

Table I.5: Outcomes across different scenarios

	Before	Merger			
		Networks fixed	Networks vary		
			w/o remedies	w/ remedies	PHX dehubbed
Total	2807.06	+0.08	+0.77	+0.91	-0.67
		[-0.47, +3.40]	[-8.92, +3.47]	[-3.92, +3.84]	[-10.01, +1.79]
Mean	4.09	+0.08	-0.73	-0.44	-1.96
		[-0.47, +3.40]	[-9.58, +1.83]	[-4.76, +2.20]	[-10.67, +0.34]
Markups: AA/US	119.20	+7.34	+12.86	+12.96	+12.36
		[+5.98, +8.64]	[+7.44, +16.30]	[+10.05, +16.37]	[+6.41, +15.50]
Markups: Other major airlines	116.22	-0.45	-1.30	-1.37	-0.93
		[-0.68, +0.07]	[-2.11, +1.10]	[-2.22, -0.40]	[-1.58, +1.17]
Segments: AA/US	430	430	493.5	500	467
			[346, 551]	[434, 559]	[330, 514]
Segments: Other major airlines	736	736	686	688.5	691
			[594, 717]	[596, 710]	[613, 719]

Note: Consumer surplus is computed using the log-sum formula and it is in USD 1 million up to constant of integration. Percentage changes with respect to the pre-merger scenario are reported for total consumer surplus, mean consumer surplus, and markups.

References

Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica*, 60(4):889–917.

Berry, S. T., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Bontemps, C., Remmy, K., and Wei, J. (2022). Ex-post evaluation of the american airlines–us airways merger a structural approach. *Journal of Transport Economics and Policy (JTEP)*, 56(2):129–155.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization.* Cambridge University Press.

Chesher, A. and Rosen, A. M. (2020). Econometric modeling of interdependent discrete choice with applications to market structure. CeMMAP Woking Paper 25/20.

Eizenberg, A. (2014). Upstream innovation and product variety in the us home pc market. *The Review of Economic Studies*, 81(3):1003–1045.

Espín-Sánchez, J.-A., Parra, A., and Wang, Y. (2021). Equilibrium uniqueness in entry games with private information.

Gualdani, C. (2021). An econometric model of network formation with an application to board interlocks between firms. *Journal of Econometrics*, 224(2):345–370.

Hellmann, T. (2013). On the existence and uniqueness of pairwise stable networks. *International Journal of Game Theory*, 42:211–237.

Jackson, M. O. and Watts, A. (2001). The existence of pairwise stable networks. *Seul Journal of Economics*, 14(3):299–321.

Jackson, M. O. and Watts, A. (2002). The evolution of social and economic networks. *Journal of Economic Theory*, 106(2):265–295.

Jackson, M. O. and Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1):44–74.

Kuehn, J. (2018). Spillovers from entry: the impact of bank branch network expansion. *The RAND Journal of Economics*, 49(4):964–994.

Lee, R. S. and Pakes, A. (2009). Multiple equilibria and selection by learning in an applied setting. *Economics Letters*, 104(1):13–16.

Lewbel, A. (2007). Coherency and completeness of structural models containing a dummy endogenous variable. *International Economic Review*, 48(4):1379–1392.

Mele, A. (2017). A structural model of dense network formation. *Econometrica*, 85(3):825–850.

Menzel, K. (2017). Strategic network formation with many agents. Working Papers.

Molchanov, I. and Molinari, F. (2018). *Random Sets in Econometrics*. Cambridge University Press.

Molinari, F. (2020). Econometrics with partial identification. *The Handbook of Econometrics*, 7a:355–486.

Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, 14(1):124–143.

Shapiro, A., Dentcheva, D., and Ruszczyński, A. (2014). *Lectures on Stochastic Programming: Modelling and Theory*. SIAM.

Sheng, S. (2020). A structural econometric analysis of network formation games through subnetworks. *Econometrica*, 88(5):1829–1858.

Stoye, J. (2009). More on confidence intervals for partially identified parameters. *Econometrica*, 77(4):1299–1315.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies*, 70(1):147–165.

Topkis, D. M. (1979). Equilibrium points in nonzero-sum n-person submodular games. *SIAM Journal of Control and Optimization*, 17(6):773–787.

Wollmann, T. G. (2018). Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review*, 108(6):1364–1406.