APPENDIX TO "Dog Eat Dog: Measuring Network Effects Using a Digital Platform Merger"

By Chiara Farronato, Jessica Fong, Andrey Fradkin

A Extensions to the Theory Model

In this Appendix, we discuss some extensions of the theory model from Section 4, which motivate some heterogeneous effects that we test in Section 6. In the theory model, we have assumed that the proportion of buyers relative to sellers is held constant. But in a two-sided platform where buyers choose from the same set of service providers, we would expect that more sellers, holding constant the number of buyers, will be beneficial for each individual buyer by increasing the expected match value v and the match probability q, and by decreasing the transacted price p. We let N denote the number of sellers in a market, as in Section 4. Here we add γ to denote the number of buyers for every seller, so that the total number of buyers is γN. So now we have $v(N, \gamma)$, $p(N, \gamma)$, and $q(N, \gamma)$.

We make the following assumptions.

1. *$p(N, \gamma)$, price, is independent of N and is increasing in γ.* This assumption means that average prices increase if buyers increase relative to sellers – because for example, sellers have heterogeneous costs and the marginal seller sets the market clearing price – but that doubling both the number of buyers and sellers does not affect average prices – because for example, the new sellers and buyers are drawn from the same distributions as the existing participants.

2. *$v(N, \gamma)$, the buyer's match value, is increasing in N and independent of γ.* Doubling both buyers and sellers allows buyers to find a better match – because for example, they find somebody living closer to their home, but an increase in buyers relative to sellers does not change the match value, conditional on matching.

3. *$q(N, \gamma)$, the match probability, is increasing in N and decreasing in γ.* Doubling both buyers and sellers allows buyers to find a match with higher probability – because for

example, it is more likely that they find somebody available for the required dates – but increasing only the number of buyers reduces the value of the match – because for example, the available sitter is now transacting with another buyer.

4. *Buyers and sellers do not multi-home.*

Assumptions 1-3 lead to the following: (1) network effects imply that u is increasing in N because of an increase in v or q or both, while leaving the price unchanged, and (2) u is decreasing in γ because of a decrease in q and in increase in p. Assumption 4 is made for simplicity: participants on the merged platform are the sum of the participants in each of the two separate platforms.

Intuitively, if two merging platforms have a different number of buyers relative to sellers, the merger will shock buyers from the two platforms in opposite directions. Both sets of buyers will see an increase in the seller pool, but if one platform had 2 buyers for every seller and the other platform had 3 buyers for every seller, the merged platform would have a number of buyers for every seller between 2 and 3, depending on their market shares. This means that after controlling for the increase in market size, buyers in the first platform will be relatively worse off – because they now need to compete with relatively more buyers for the sellers available compared to before – while buyers in the second platform would be better off.

As long as the utility function has decreasing marginal returns to each additional user then the benefits from the merger are bigger in markets where the two platforms have very different relative proportions of buyers and sellers. This is because the average utility can be written as a function of the total number of sellers N and of the number of buyers relative to sellers γ. Assuming that $\gamma N u(N, \gamma)$ is twice continuously differentiable, has increasing returns to scale, and has decreasing returns to each individual buyer or seller implies that its first derivatives with respect to both N and γ are positive, but while the second derivative with respect to N is positive, the second derivative with respect to γ is negative. $u(N, \gamma)$ is thus concave in γ. Let's now assume that both competing platforms have the same number of sellers, $0.5N$, but the larger platform has a share $\pi \geq 0.5$ of buyers. This means that the larger platform has a number of buyers relative to sellers equal to $\pi\gamma$ where γ is the market-level number of buyers per seller. Analogously the smaller platform, which has the remaining buyers, has a number of buyers relative to sellers equal to $(1-\pi)\gamma$. So the aggregate utility in the market, which is the sum of the utilities of both platforms, will be equal to $\pi\gamma 0.5Nu(0.5N, \pi\gamma) + (1-\pi)\gamma 0.5Nu(0.5N, (1-\pi)\gamma)$. After the platforms merge, the combined market has N sellers, and γN buyers, and thus, aggregate utility when the platforms merge is $\gamma Nu(N, \gamma)$ regardless of the pre-merger share of buyers π. However, due to Jensen's inequality, the closer π is to 1, the lower the aggregate pre-merger utility from the two separate platforms: $\pi\gamma 0.5Nu(0.5N, \pi\gamma) + (1-\pi)\gamma 0.5Nu(0.5N, (1-\pi)\gamma)$. Therefore the increase in benefits from the merger is greater in markets with higher π.

Relaxing Assumption 4, to the extent that a share of sellers multi-home, we would expect the benefits from the merger to be lower. So we can test whether network effects are larger in markets where a lower share of sellers multi-home.

Finally, we have compared markets with the same number of participants but different proportions of participants across the two competing platforms. To the extent that increasing returns to scale are higher for markets with a smaller number of participants (smaller N), we would expect the benefits from the merger to be concentrated in geographies with fewer participants.

B Extensions to the Empirical Results

In this appendix, we provide additional results to Section 6.

First, we provide results on additional outcomes in Figure B.1.

Second, we provide additional results separately for new users, Rover users, DogVacay users (Figure B.2), and multi-homing users (Figure B.3).

Third, we provide additional results for large versus small markets, for market with little versus substantial multihoming, and for markets with little versus large differences in the relative proportions of buyers and sellers across the two competing platforms. the results are in Figure B.4, Figure B.5, and Figure B.6.

Fourth, we present the coefficients from the matching regressions in tables for better readability. Instead of estimating a coefficient for each month, as in Equation (3), we estimate a coefficient for the transition period (March to June 2017) and post-acquisition (July to December 2017). Instead of normalizing February 2017 to 0, we normalize all 3 months before the acquisition (December 2016 - February 2017) to 0. We refer to this period as the baseline. We also estimate a pre-trend coefficient for the 3 months before the baseline. The interpretation of each coefficient is the average difference between the treated market and a matched control unit in the respective time period, relative to the baseline period. Note that if all matched markets had identical pre-trends, we would expect the coefficient for the 3 months before the baseline to be not statistically different from 0. The below regression is estimated separately for each Rover market share group.

$$y_{zt} - y_{z't} = \alpha + \beta_1 \mathbf{1}\{t \in 3\ Months\ PreBaseline\} + \beta_2 \mathbf{1}\{t \in\ Transition\} + \beta_3 \mathbf{1}\{t \in PostMerger\} + \epsilon_{z,z',t} \quad (4)$$

Results are presented in Tables B.1 through B.6.

Fifth, we provide results using a simple difference-in-differences estimation with no matching, which accounts for differential pre-trends across market share groups. We replace Equation (3) with the following, un-matched, equation:

$$y_{zt} = \beta_{s(z)t} + \gamma_{s(z)} t + \delta_{s(z)} \mathbb{1}\{t \geq Dec2016\} + \mu_t + \mu_z + \epsilon_{zt}. \quad (5)$$

By adding $\gamma_{s(z)}t + \delta_{s(z)}\mathbb{1}\{t \geq Dec2016\}$, we allow for the observations in the treatment and control groups to have a different linear pre-trend. Results are presented in Figure B.7.

Sixth and finally, in the paper we have defined markets at the zip code level. The problem with this definition is that zip codes are not independent of each other. There are over 20 zip codes in Seattle, and dog owners may search for sitters across many zip codes within their city. It is possible that in zip code A, Rover had 50% of the market before the acquisition, and in neighboring zip code B it had 75% of the market. After the acquisition, the bigger increase in options in zip code A may cause some dog owners to substitute away from sitters in B towards sitters in A. This would amplify the post-acquisition outcome differences between A and B. The above example demonstrates how the stable unit treatment value assumption (SUTVA) of causal inference does not hold. This bias has been studied in the context of online marketplaces for inferences from A/B experiments (Holtz and Aral, 2018).

To reduce bias from violations of SUTVA, we form clusters of zip codes separately for each CBSA. The construction of clusters must balance two competing objectives. On one hand, larger clusters reduce interactions between units of observation. On the other hand, larger clusters mean fewer observations and less statistical power. For this reason, we choose a clustering procedure that allows us to explore this trade-off.

We use a geographically constrained hierarchical clustering algorithm,[37] which allows us to impose that a cluster be formed by a spatially contiguous set of zip codes. A key advantage of this algorithm is that more aggregated clustering nests less aggregated clustering — i.e. all zip codes belonging to one cluster when the clustering is less aggregated map to the same (larger) cluster when the clustering is more aggregated. Therefore, it is easy to vary the desired size of clusters to evaluate the bias-precision trade-off.

The clustering procedure takes in two dissimilarity matrices. The first matrix gives dissimilarities in the "feature space" and it is computed from data on co-occurrence of searches,[38] i.e. cases when a dog owner sees listings from two zip codes in the same set of search results. The more frequently the two zip codes co-occur, the more similar they are. The second matrix gives the dissimilarities in the "constrained space", and each element

[37] We use the R package *ClustGeo* (Chavent et al., 2018).

[38] We use 2017 search results from Rover to construct the matrix of dissimilarity in the feature space.

is 0 or 1 depending on whether two zip codes are geographically contiguous. There is a final parameter, α, which controls the importance of each dissimilarity matrix — higher α increases the importance of the geographic distances. We also have the freedom to choose the number of clusters in a given CBSA. We choose α and the number of clusters to maximize the number of observations — clusters — subject to a threshold on the level of interactions among distinct clusters.

Specifically, we implement the Ward-like hierarchical clustering method with spatial constraints proposed by Chavent et al. (2018). The algorithm takes in the following inputs:

- A dissimilarity matrix D_0 composed of distances $(d_{0,ij})$ between zip codes i and j. The distances are based on how frequently two zip codes occur together in search results.[39] We measure co-occurrences in the following way. For each search s, we take the corresponding search results and create all unique zip code pairings. For the pair of zip codes i and j we compute the probability of obtaining the pair i, j out of a draw of two search results from search s.[40] The probability $p_{s,ij}$ takes values between 0—if i or j do not appear in the search results from search s—and .5 —if search s has only two results, one from zip code i and the other from zip code j. We aggregate at the zip code-pair level by summing over searches, and we normalize by the minimum number of searches with results from zip code i or zip code j. We call this the co-occurrence share. The distance $d_{0,ij}$ is equal to the reciprocal of the co-occurrence share:

$$d_{0,ij} = \frac{min(\sum_s \mathbb{1}\{\text{search } s \text{ contains zip code } i\}, \sum_s \mathbb{1}\{\text{search } s \text{ contains zip code } j\})}{\sum_s \mathbb{1}\{\text{search } s \text{ contains zip codes } i \text{ and } j\} p_{s,ij}}$$

  Infinite values are set to $2 \underset{d_{0,ij}<\infty}{max}\, d_{0,ij}$. This guarantees that after normalizing the dissimilarity matrix $\frac{D_0}{max(D_0)}$, the distance values are either 1 (for zip codes with no co-occurrences) or between 0 and .5. The diagonal values are set to 0.

- A matrix D_1 of geographic distances $(d_{1,ij})$ between zip codes i and j. The distance $d_{1,ij}$ is equal to 1 if zip codes i and j are not geographic neighbors, and it is equal to 0 otherwise. Every zip code has a distance 0 from itself so the diagonal is once again

[39] We have search results data from 2017 for Rover.

[40] For computational ease, we sample search results with replacement to compute $p_{s,ij}$.

set to 0.

- A set of weights (w_i), one for each zip code. We set $w_i = 1$ for all zip codes.
- A parameter, α, which determines the importance of the geographic distance matrix D_1 relative to the co-occurrence distance matrix D_0.

The values in the normalized matrix $\frac{D_0}{max(D_0)}$ and in D_1 are all between 0 and 1 so the matrices have the same order of magnitude. The algorithm then proceeds in steps starting from a partition $\mathcal{P}_n^\alpha$ where each of the n zip codes is a separate cluster. At each following step k, for each cluster $\mathcal{C}_k^\alpha$ we compute the mixed pseudo inertia as

$$I_\alpha\left(\mathcal{C}_k^\alpha\right) = (1-\alpha)\sum_{i\in\mathcal{C}_k^\alpha}\sum_{j\in\mathcal{C}_k^\alpha}\frac{w_i w_j}{2\mu_k^\alpha}d_{0,ij}^2 + \alpha\sum_{i\in\mathcal{C}_k^\alpha}\sum_{j\in\mathcal{C}_k^\alpha}\frac{w_i w_j}{2\mu_k^\alpha}d_{1,ij}^2,$$

where $\mu_k^\alpha = \sum_{i\in\mathcal{C}_k^\alpha} w_i$ is the aggregate weight of cluster $\mathcal{C}_k^\alpha$. The mixed pseudo inertia is a measure of homogeneity within a cluster, which is a function of the dissimilarity values in characteristics and geography. In order to obtain a new partition $\mathcal{P}_k^\alpha$ in k clusters from a given partition $\mathcal{P}_{k+1}^\alpha$ in $k+1$ clusters, we choose to combine clusters $\mathcal{A}$ and $\mathcal{B}$ belonging to $\mathcal{P}_{k+1}^\alpha$ to minimize mixed within cluster inertia:

$$\arg \min_{\mathcal{A},\mathcal{B}\in\mathcal{P}_{k+1}^\alpha} I_\alpha\left(\mathcal{A}\cup\mathcal{B}\right) - I_\alpha\left(\mathcal{A}\right) - I_\alpha\left(\mathcal{B}\right).$$

We can graphically represent the hierarchically-nested set of partitions $\{\mathcal{P}_n^\alpha, ..., \mathcal{P}_k^\alpha, ...\mathcal{P}_1^\alpha\}$ with a tree. We are free to choose where to 'cut' the tree, i.e. the number k of clusters to include in our partition. We are also free to choose α. To select α and k we implement the following algorithm:

1. We divide zip codes into Core-Based Statistical Areas (CBSAs). We perform steps 2-4 separately for each CBSA, which means that we choose α, k separately for each CBSA.[41]

[41] A handful of CBSAs have zip codes with no neighbors. For example, Odessa, TX, has a zip code that only borders an airport. These zip codes pose a problem for the Ward-based algorithm. In this case we cluster zip codes ignoring the geographic dissimilarity matrix. So for these CBSAs, we set $\alpha = 0$.

2. We implement the hierarchical clustering with spatial constraints for a grid of values for $\alpha \in \{.25, .5, .75, 1\}$ and for k between 1 and $min(100, n)$, where n is the number of zip codes in the CBSA.[42]

3. Our measure of cluster quality Q_k^α is derived from the search data in a similar manner to the dissimilarity matrix. For each cluster in partition $\mathcal{P}_k^\alpha$ we compute the weighted number of search co-occurrences within each cluster and divide it by the weighted total co-occurrences in the CBSA. We then sum across clusters within CBSA to get the cluster quality.

$$Q_{k,CBSA}^\alpha = \frac{\sum_{c \in \mathcal{C}_k^\alpha} \sum_{i,j \in c} \sum_s \mathbb{1}\{\text{search } s \text{ contains zip codes } i \text{ and } j\} p_{k,ij}}{\sum_{i,j \in CBSA} \sum_s \mathbb{1}\{\text{search } s \text{ contains zip codes } i \text{ and } j\} p_{k,ij}}.$$

   If all co-occurrences are within cluster, then $Q_k^\alpha = 1$, representing a perfect clustering. In practice, some co-occurrences inevitably occur across clusters. These are driven by the dispersion of search results shown by Rover's ranking algorithm and by the willingness of owners to consider many zip codes.

4. We pick the partition $\mathcal{P}_k^\alpha$ with the highest k subject to $Q_k^\alpha > .65$.

Intuitively we find the partition with the most distinct clusters subject to a minimum quality threshold that controls the potential interdependencies across clusters. Setting the threshold at 65% means that on average 65% of requests have booking inquiries only within the cluster. Note that this threshold is far from 100%. 100% means that all booking inquiries for the same request happen within the same cluster.

Figure B.8 plots the clusters that our procedure finds in four of the largest cities in our data. The clusters are reasonably contiguous in space, and in general much larger than individual zip codes. On average each cluster has 6.26 zip codes. There are also a few separate clusters in each city, implying that not all zip codes in a CBSA are equally substitutable between one another.

We then estimate Equation (3) with cluster-month as unit of observation. Results are presented in Figure B.9.

[42] For CBSAs with more than 200 zip codes the 25 limit can be binding in practice, so we use k between 1 and $min(50, n)$, where n is the number of zip codes in the CBSA.

Table B.1: Estimates of Merger Effects - Market Level

Period	Rover Share	Buyers (log)	Sitters (log)	Stays (log)	Match Rate	Price	Pr(Request Again)	Pr(Repeat Stay)	Pr(5 star)
3 Mons Before Baseline	[0.0,0.2)	-0.002	0.025	0.043	0.023	0.45	-0.01	-0.008	0.031
	[0.2,0.4)	-0.042	-0.018	-0.017	0.014	0.645*	-0.013	-0.013	-0.004
	[0.4,0.6)	-0.009	-0.006	-0.007	0	0.428	-0.007	-0.008	-0.007
	[0.6,0.8)	-0.042*	-0.038**	-0.05*	-0.001	0.128	0.001	0.025*	-0.004
Transition	[0.0,0.2)	-0.015	0.014	-0.023	-0.013	-0.186	-0.034*	-0.006	-0.019
	[0.2,0.4)	-0.006	0.001	0.013	0.01	0.386	-0.01	-0.03*	-0.018
	[0.4,0.6)	-0.005	0.006	0.015	0.004	0.04	0.006	-0.021	-0.02
	[0.6,0.8)	-0.015	0.009	-0.027	-0.006	0.165	0.004	0	-0.038***
Post-Merger	[0.0,0.2)	-0.021	-0.036	-0.077*	-0.035***	-1.622***	-0.028	0.012	-0.015
	[0.2,0.4)	-0.043	-0.046**	-0.073*	-0.012	-0.624*	-0.012	-0.018	-0.022
	[0.4,0.6)	0.039	-0.009	0.044	0.001	-0.133	0.005	0	-0.015
	[0.6,0.8)	0	0.006	-0.036	-0.012	0.018	0.004	0.011	0

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table displays the estimated coefficients of each period in Equation (4) and is analogous to Figure 5a.

Table B.2: Estimates of Merger Effects - Rover

Period	Rover Share	Buyers (log)	Sitters (log)	Rover Stays (log)	Rover Match Rate	Price	Pr(Request Again)	Pr(Repeat Stay)	Pr(5 star)
3 Mons Before Baseline	[0.0,0.2)	-0.244***	-0.035	-0.253***	-0.011	-0.138	0.021	-0.008	0.008
	[0.2,0.4)	-0.113***	-0.03*	-0.148***	-0.01	0.646*	0.016	-0.012	-0.016
	[0.4,0.6)	-0.037	-0.026	-0.069**	-0.006	-0.115	0.013	0	-0.018
	[0.6,0.8)	-0.054**	-0.042**	-0.082***	-0.01	0.382	0	0.03*	-0.009
Transition	[0.0,0.2)	0.455***	0.24***	0.434***	0.077***	-0.253	0.048**	0.035	-0.015
	[0.2,0.4)	0.229***	0.133***	0.255***	0.019**	0.249	-0.005	-0.021	-0.01
	[0.4,0.6)	0.142***	0.093***	0.167***	0.015*	-0.215	0	-0.003	-0.014
	[0.6,0.8)	0.062***	0.072***	0.046*	-0.006	-0.097	-0.005	0.002	-0.029*
Post-Merger	[0.0,0.2)	1.192***	0.451***	1.301***	0.197***	0.724	0.08***	0.072***	-0.035
	[0.2,0.4)	0.56***	0.237***	0.699***	0.076***	0.687*	0.01	0.014	-0.029
	[0.4,0.6)	0.401***	0.178***	0.499***	0.05***	0.36	0.011	0.027*	-0.013
	[0.6,0.8)	0.192***	0.132***	0.188***	0.003	0.37	-0.005	0.016	0.001

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table is the same as Table B.1, except the outcomes are for Rover only. These results are analogous to Figure 5b.

Table B.3: Estimates of Merger Effects - New Users

Period	Rover Share	New Buyers (log)	New Sitters (log)	New Buyer Stays (log)	Match Rate (New Buyers)	Price (New Buyers)	Pr(Request Again) (New Buyers)	Pr(Repeat Stay) (New Buyers)	Pr(5 star) (New Buyers)
3 Mons Before Baseline	[0.0,0.2)	-0.01	0.025	0.074**	0.03	1.284	-0.018	0.007	0.021
	[0.2,0.4)	-0.021	0.055**	0.058*	0.025**	0.53	0.021	-0.019	-0.021
	[0.4,0.6)	0.006	0.005	0.015	0	-0.089	-0.001	-0.018	-0.002
	[0.6,0.8)	-0.027	0.018	-0.023	0.004	-0.084	0.038*	-0.007	0.028
Transition	[0.0,0.2)	-0.048	-0.01	-0.007	0.005	1.273	-0.015	0.011	0.003
	[0.2,0.4)	-0.038	0.023	-0.004	0.006	0.239	-0.025	-0.046**	-0.008
	[0.4,0.6)	-0.048	-0.004	-0.049	-0.007	-0.014	0.005	-0.029	-0.008
	[0.6,0.8)	-0.03	0	-0.06**	-0.013	-0.5	0.019	-0.02	-0.024
Post-Merger	[0.0,0.2)	0.016	-0.026	0.006	-0.013	0.129	0.037	0.023	-0.023
	[0.2,0.4)	-0.03	-0.019	-0.037	-0.008	0.524	0.014	-0.012	-0.026
	[0.4,0.6)	0.023	-0.015	0.013	-0.009	-0.25	0.027	-0.003	-0.002
	[0.6,0.8)	0.01	-0.014	-0.034	-0.014	0.368	0.045**	-0.005	0.005

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table is the same as Table B.1, except the outcomes are for Rover only. These results are analogous to Figure 6a

Table B.4: Estimates of Merger Effects - Rover Users

Period	Rover Share	Buyers (log)	Sitters (log)	Stays (log)	Match Rate	Price	Pr(Request Again)	Pr(Repeat Stay)	Pr(5 Star)
3 Mons Before Baseline	[0.0,0.2)	0.066	-0.051*	0.114***	-0.034	-0.446	-0.003	-0.039	0.029
	[0.2,0.4)	-0.058*	-0.09***	-0.034	-0.014	0.786	0.051*	-0.042	0.057
	[0.4,0.6)	0.006	-0.041**	-0.02	-0.035**	0.25	0.011	0.011	-0.019
	[0.6,0.8)	-0.002	-0.02	-0.037	-0.037**	1.157*	0.022	0.054	-0.019
Transition	[0.0,0.2)	0.047	0.122***	-0.053	-0.005	-1.356	0.01	0.022	-0.02
	[0.2,0.4)	0.072**	0.084***	0.025	-0.001	0.588	0.01	-0.053	0.026
	[0.4,0.6)	0.088***	0.08***	0.067**	0	-0.41	-0.004	-0.035	-0.02
	[0.6,0.8)	0.044**	0.057***	0.007	-0.023**	0.684	0.008	0.024	-0.002
Post-Merger	[0.0,0.2)	0.358***	0.257***	0.263***	0.041*	0.17	0.016	0.018	-0.042
	[0.2,0.4)	0.171***	0.116***	0.174***	0.039**	0.538	0.018	-0.005	0.055
	[0.4,0.6)	0.18***	0.125***	0.167***	0.013	0.053	-0.004	0.032	-0.024
	[0.6,0.8)	0.075***	0.081***	0.028	-0.024*	1.013*	-0.007	0.008	-0.022

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table displays the estimates of Equation (4) for users who engaged in a booking inquiry on Rover only in 2016. This table is analogous to Figure 6b.

Table B.5: Estimates of Merger Effects - DogVacay Users

Period	Rover Share	Buyers (log)	Sitters (log)	Stays (log)	Match Rate	Price	Pr(Request Again)	Pr(Repeat Stay)	Pr(5 Star)
3 Mons Before Baseline	[0.0,0.2)	0.008	0.057**	-0.044	-0.014	-3.375	0.043	-0.06	0.087
	[0.2,0.4)	0.03	0.062**	-0.006	0.012	-2.336	-0.006	0.157*	-0.043
	[0.4,0.6)	0.02	0.045**	-0.05*	-0.022	1.133	-0.018	0.038	-0.068
	[0.6,0.8)	-0.06**	-0.037	-0.073***	0	-2.234	0.04	0.134*	0.003
Transition	[0.0,0.2)	-0.001	-0.093***	0.074***	0.009	-0.327	-0.064	-0.144**	-0.006
	[0.2,0.4)	-0.008	-0.075***	0.077***	0.019	0.964	-0.029	-0.07	-0.052
	[0.4,0.6)	-0.032	-0.066***	0.038	-0.01	-0.279	-0.034	-0.016	0.057
	[0.6,0.8)	-0.01	-0.051**	0.043*	0.006	-1.053	0.057	0.041	-0.036
Post-Merger	[0.0,0.2)	-0.362***	-0.255***	-0.332***	-0.072**	-0.485	-0.085*	-0.126*	0.026
	[0.2,0.4)	-0.366***	-0.254***	-0.322***	-0.057**	1.037	-0.044	-0.106**	0.008
	[0.4,0.6)	-0.245***	-0.171***	-0.21***	-0.051**	-0.011	-0.032	-0.003	0.022
	[0.6,0.8)	-0.176***	-0.13***	-0.113***	-0.01	-1.346	0.04	-0.011	0.055

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This table displays the estimates of Equation (4) for users who engaged in a booking inquiry on Rover only in 2016. This table is analogous to Figure 6c.

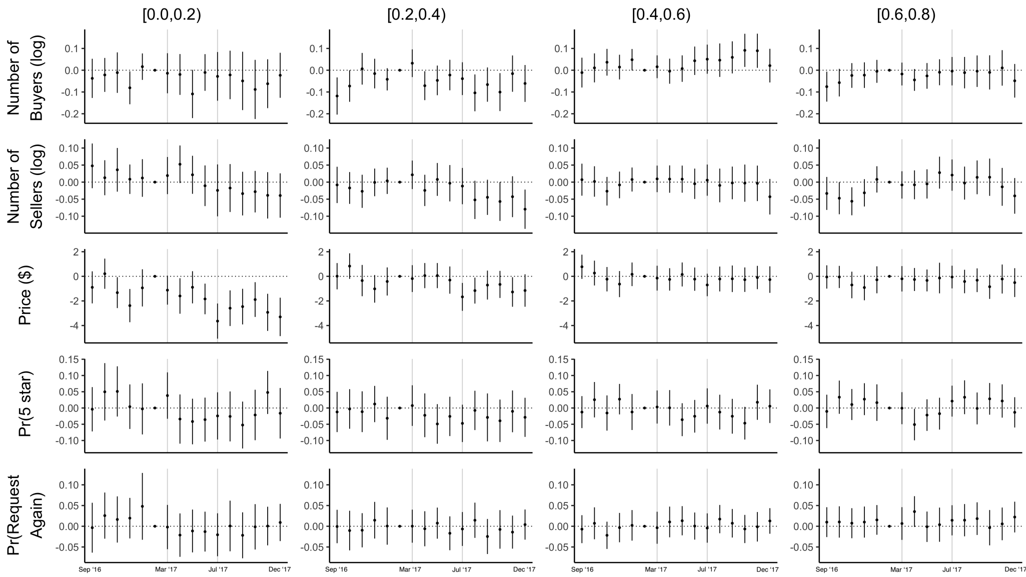
Table B.6: Estimates of Merger Effects for Conversion from Search to Request

Period	Rover Share	All	New	Rover	DogVacay
Transition	[0.0,0.2)	0.011*	0.003	-0.017	0.015
	[0.2,0.4)	0.012**	0.006	-0.02	-0.137
	[0.4,0.6)	0.007*	0.001	-0.007	-0.054
	[0.6,0.8)	0	0.001	-0.011	0.024
Post-Merger	[0.0,0.2)	0.026***	0.005	0	0.049
	[0.2,0.4)	0.021***	0.006	-0.015	-0.109
	[0.4,0.6)	0.013***	0.001	-0.01	-0.05
	[0.6,0.8)	0.006	0	-0.001	0.017

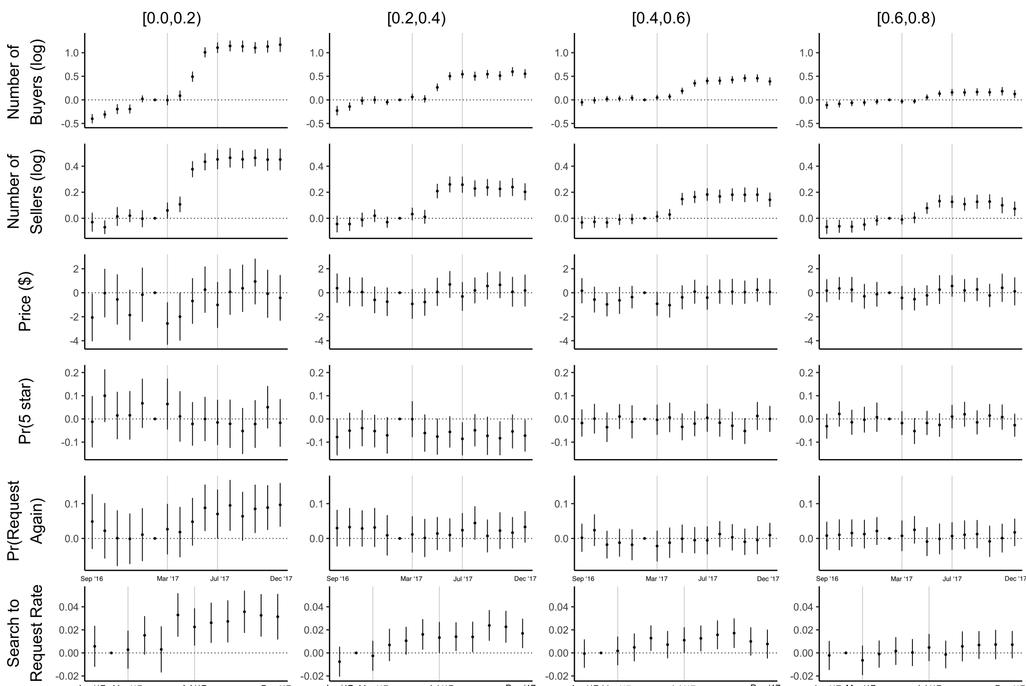
Note: *p<0.1; **p<0.05; ***p<0.01

This table displays the estimates of Equation (4). The outcome variables are the search to request rate for various types of users. This table is analogous to Figure C.1.

Figure B.1: Estimates of Merger Effects – Additional Outcomes



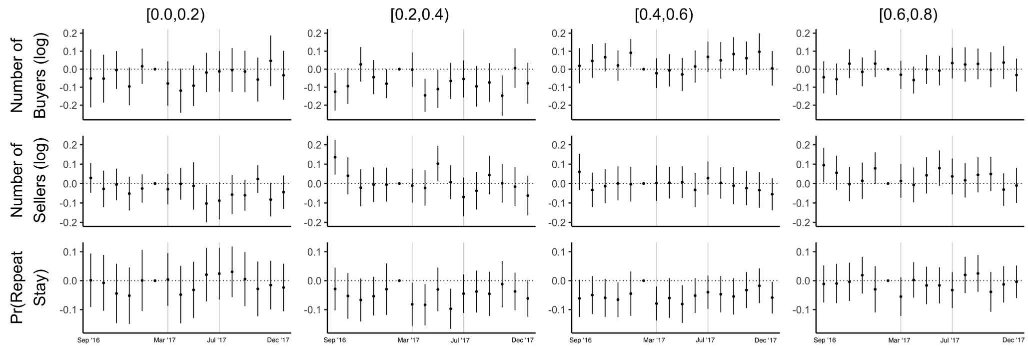
(a) Market Outcomes



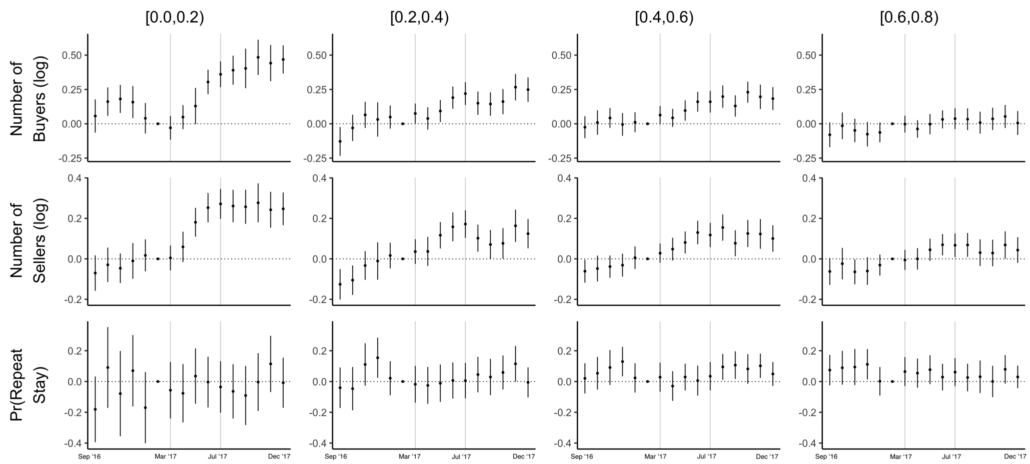
(b) Rover Outcomes

Regression estimates of Equation (3) for additional outcomes. Otherwise the figure is identical to Figure 5. Due to data limitations, we are not able to identify buyer and seller location precisely for DogVacay users, which limits our ability to use geographic distance between buyers and sellers as a proxy for match quality.

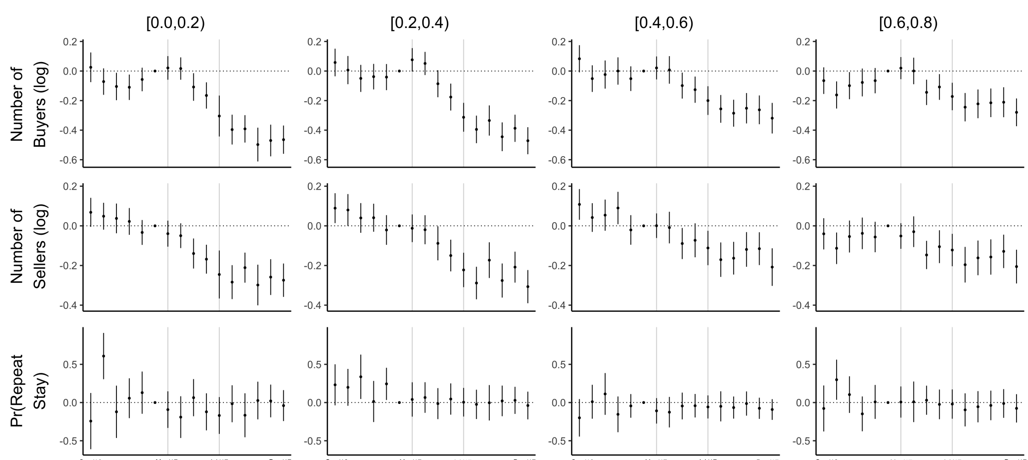
Figure B.2: Estimates of Merger Effects By User Type – Other Outcomes



(a) New Users



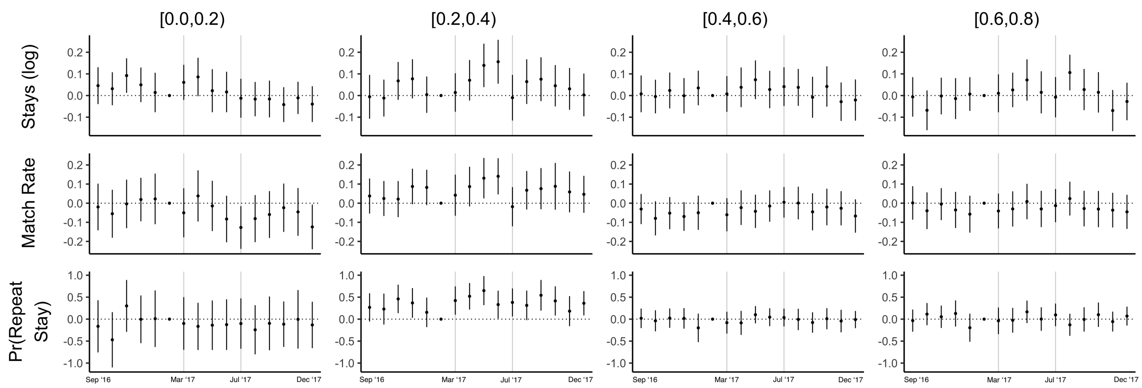
(b) Rover Users



(c) DogVacay Users

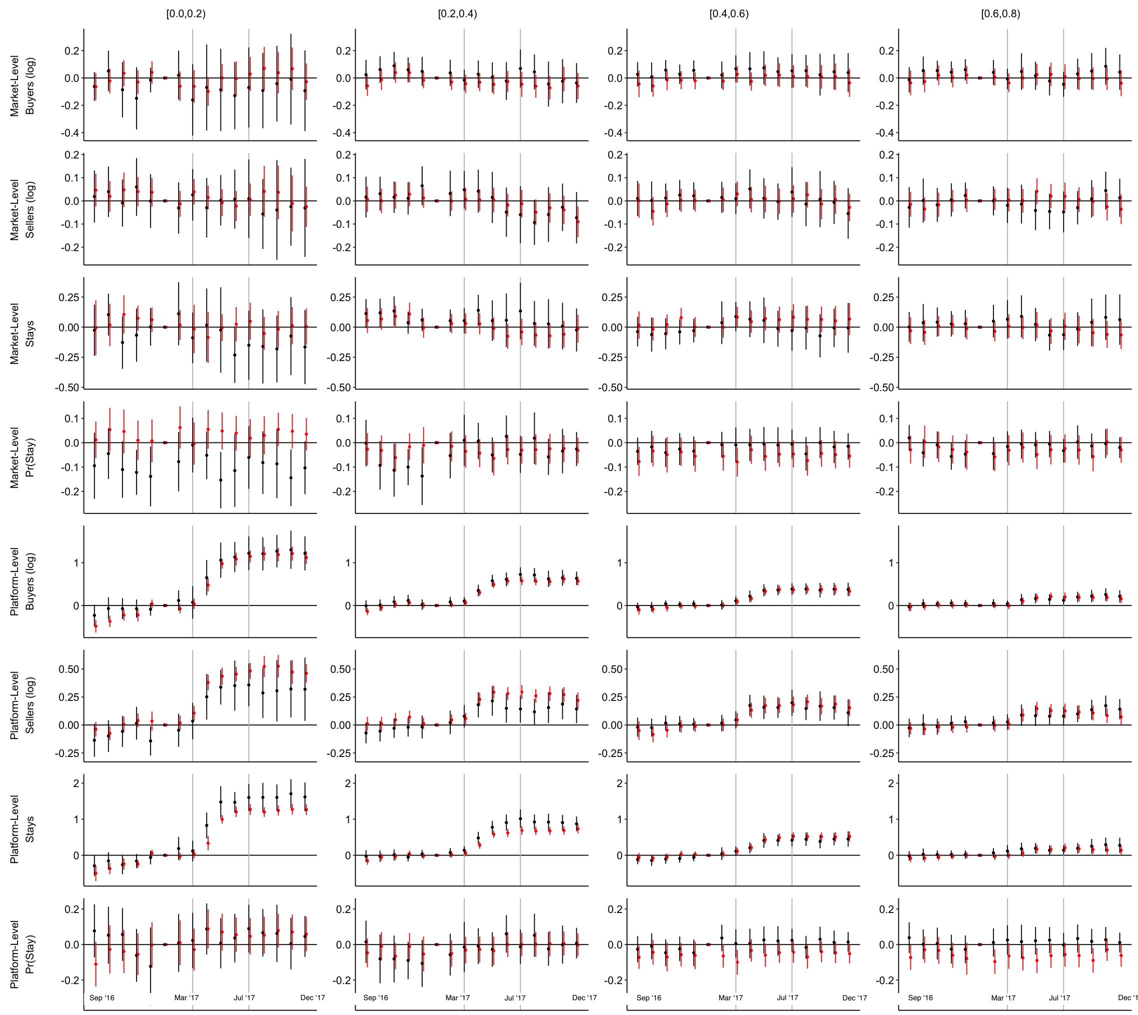
This figure displays results on additional outcomes for new users, Rover users, and DogVacay users. Otherwise the figure is identical to Figure 6. Outcomes for multi-homing users are in Appendix Figure B.3.

Figure B.3: Estimates of Merger Effects By User Type – Multihoming Users



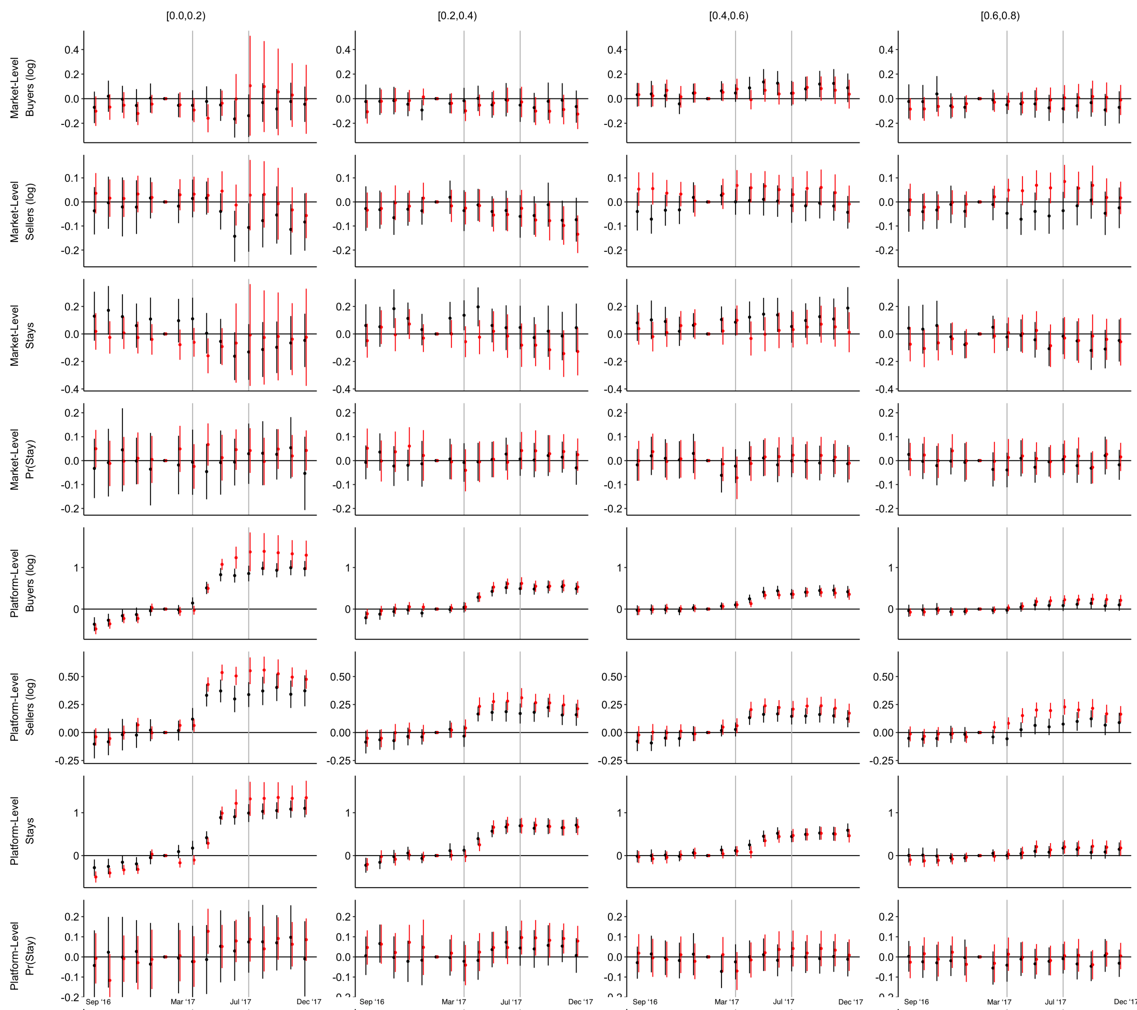
This figure displays results for multi-homing users. Multi-homing users are defined as those who engaged in booking inquiries on both Rover and DogVacay in the previous year. Otherwise the figure is identical to Figure 6 and Appendix Figure B.2.

Figure B.4: Estimates of Merger Effects – Heterogeneity by Market Size, Other Outcomes



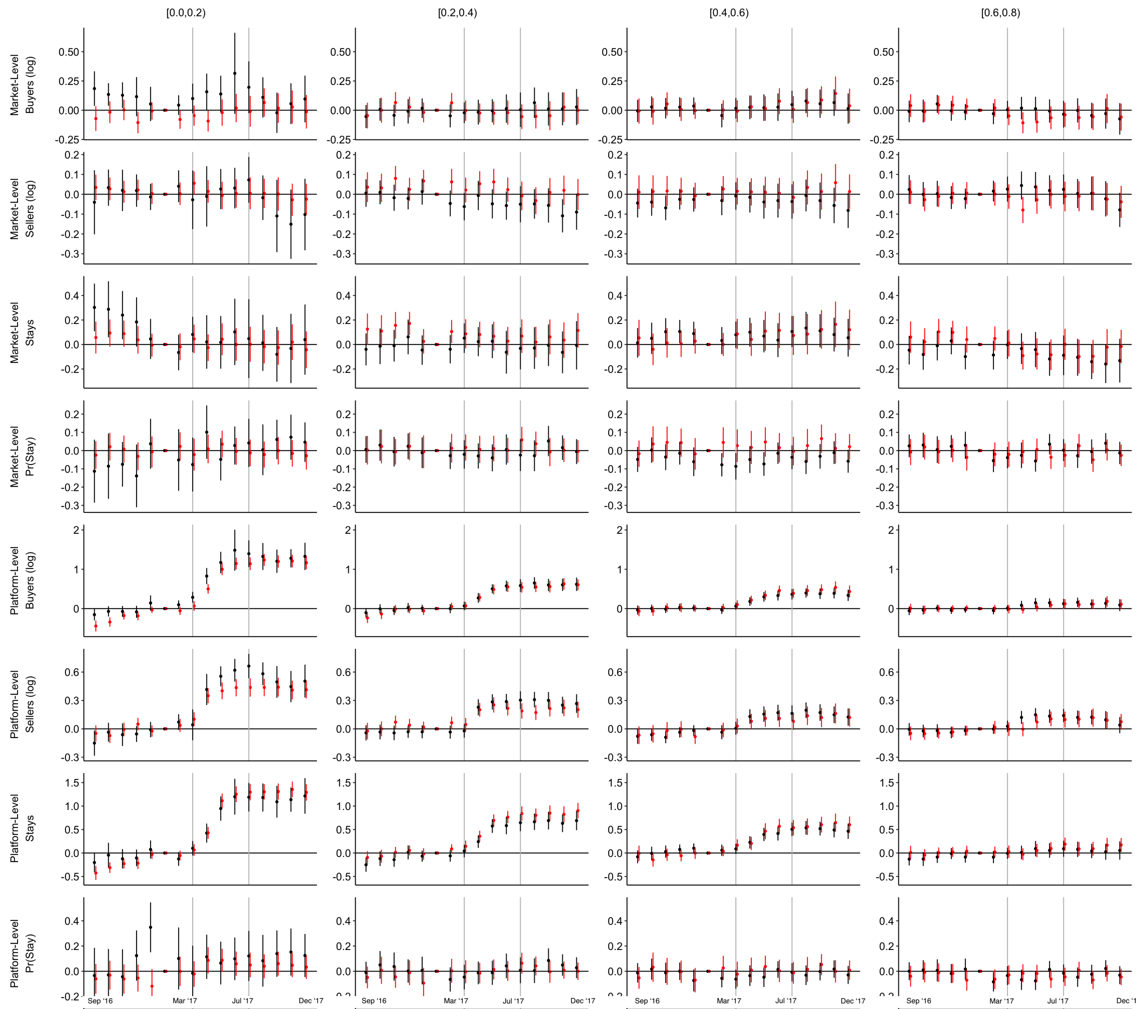
The figure is identical to Panel (a) of Figure 7 for small versus large markets, except that it provides results for additional outcomes.

Figure B.5: Estimates of Merger Effects – Heterogeneity by Multihoming Propensity, Other Outcomes



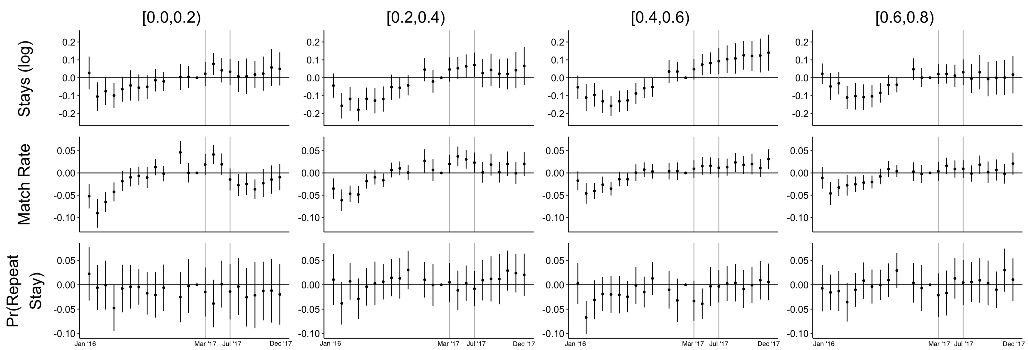
The figure is identical to Panel (b) of Figure 7 for markets with little versus substantial multihoming, except that it provides results for additional outcomes.

Figure B.6: Estimates of Merger Effects – Heterogeneity by Buyer to Seller Ratio, Other Outcomes

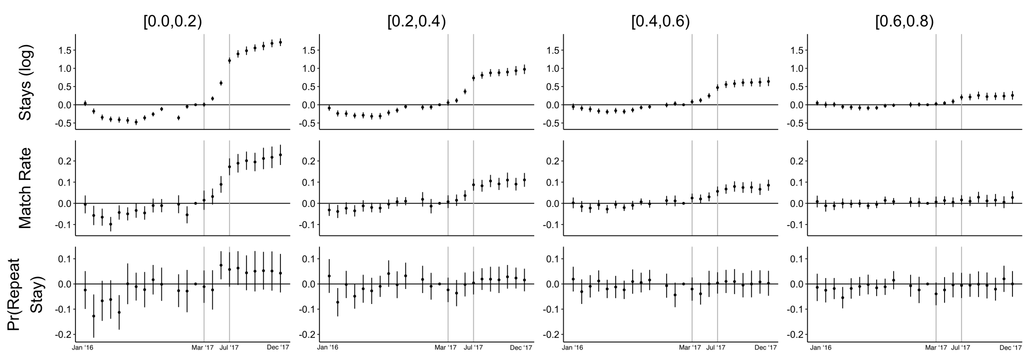


The figure is identical to Panel (c) of Figure 7 for markets with large versus small differences between Rover and DogVacay in the number of buyers relative to sellers, except that it provides results for additional outcomes.

Figure B.7: Estimates of Merger Effects – Unmatched



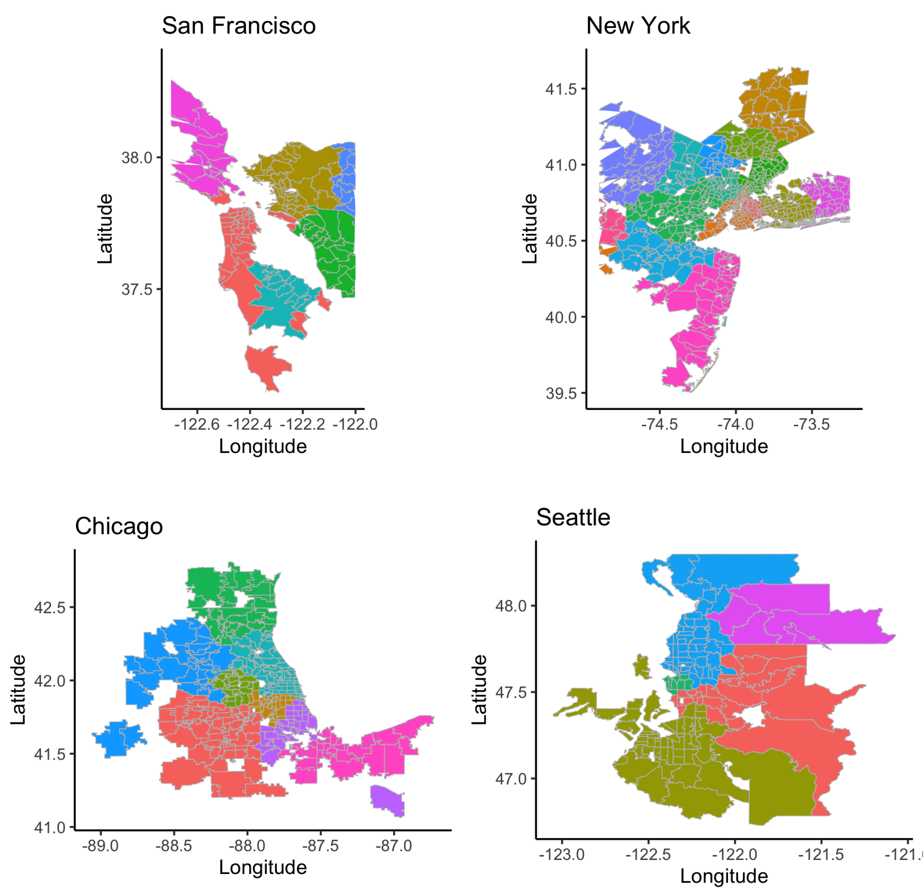
(a) Market Outcomes



(b) Rover Outcomes

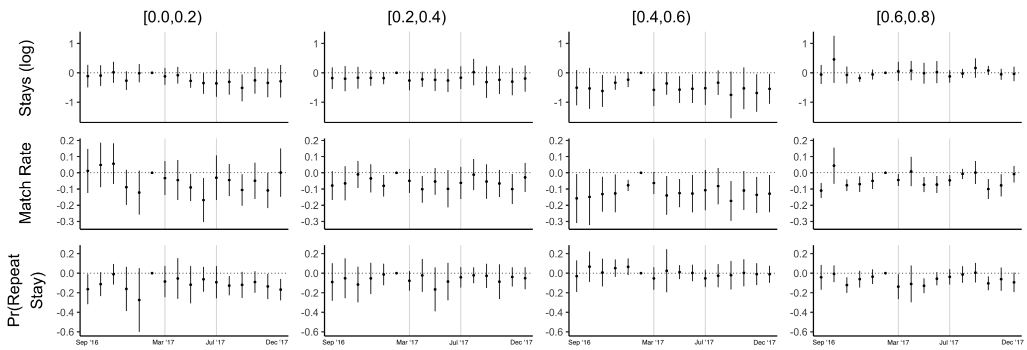
Regression estimates of Equation (5). Otherwise the table is identical to Figure 5.

Figure B.8: Cluster Maps - CBSAs

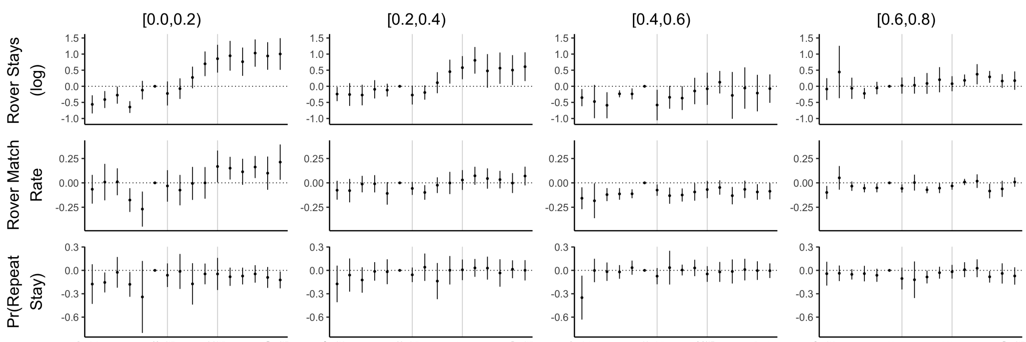


The figures plot the clusters for four Core-Based Statistical Areas (CBSAs) formed by aggregating zip codes using hierarchical clustering with geographic constraints.

Figure B.9: Estimates of Merger Effects – Geographic Clusters



(a) Market Outcomes



(b) Rover Outcomes

Regression estimates of Equation (3) with geographic clusters as markets instead of zip codes. Otherwise the table is identical to Figure 5.

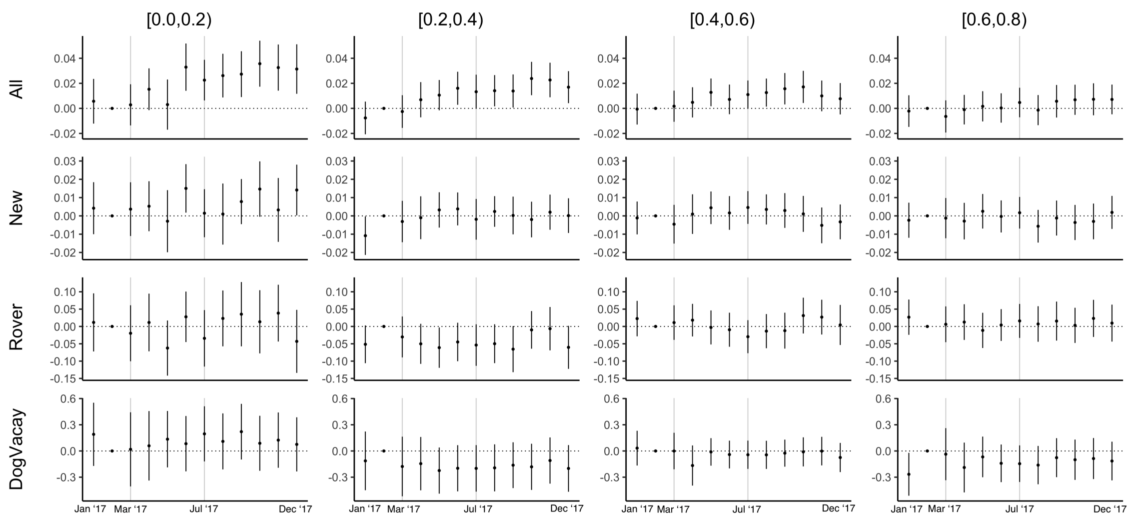
C Additional Results Based on Search Data

The discussion in the main body of the paper has focused on booking inquiries and stays. However, there may also be effects of the merger on the probability that a search leads to a booking inquiry. Intuitively, searchers should be more likely to find suitable sitters in markets with more sitters. We have data on search behavior only starting in 2017 and only for the Rover platform. As a result, we can only compute platform level rather than market level outcomes related to search. This limits our ability to say how search conversion changed at a market level, but does allow us to measure changes in platform efficiency.

We observe data on search requests, which are queries into the Rover search engine, and search results, which are results returned for those queries. We are also able to observe the mapping between a search and a user in the database for a subset of queries. For other queries, we cannot map the search to a user, either because the user did not have an account or because the platform was not able to successfully map the search to a user. We attribute the search to a location by using the first zip searched by the searcher in a given month. Lastly, we define a conversion (either to a booking inquiry or to a stay) as a binary variable that takes the value of 1 when a searching user has at least one booking inquiry or stay initiated in that year-month.

Using the above definitions and matched sample, we estimate the effect of merging the two platforms on platform conversion rates (Figure C.1) from search to booking inquiry. The first row shows that conversion rates increase by up to 3 percentage points in markets with the lowest Rover market share pre-acquisition (first plot on the first row). This is a large effect relative to the baseline conversion rate across year-month-zip codes in 2017 and confirms the finding in Figure 5b in the main paper. The increase is mostly driven by compositional changes—DogVacay users migrating to Rover—given that we do not see differences post-acquisition in conversion rates for existing or new users (last three rows of Figure C.1).

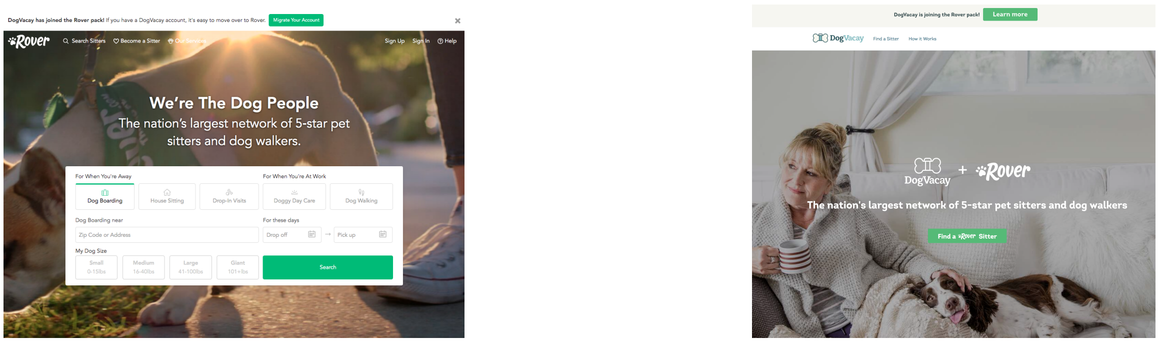
Figure C.1: Merger Effects for Conversion from Search to Booking Inquiry



Regression estimates of Equation (3). The first row displays results where the outcome is the conversion rate of searches to booking inquiries for all searchers. The second row displays results only for users who have not previously made a request or searchers who are unknown. The third row displays results only for users who made requests exclusively on Rover in 2016. The fourth row displays results for users who made requests exclusively on DogVacay in 2016.

D Additional Figures and Tables

Figure D.1: Rover's and DogVacay's Landing Pages After the Merger

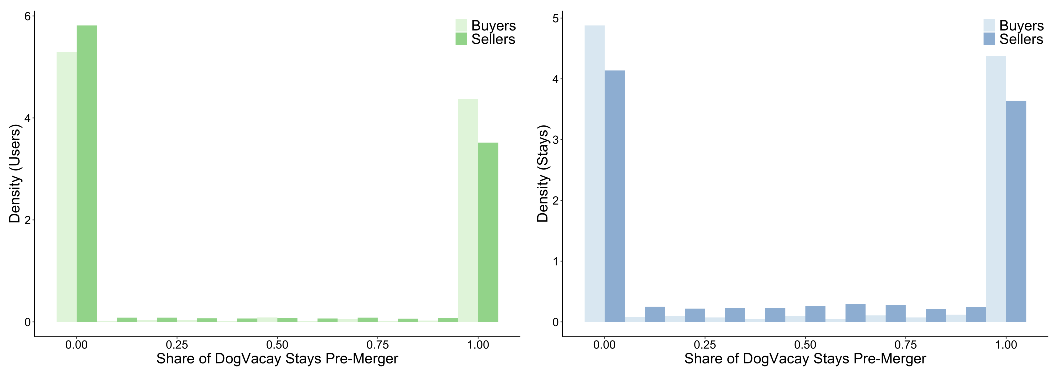


(a) Rover.com, July 2017.

(b) Dogvacay.com, July 2017.

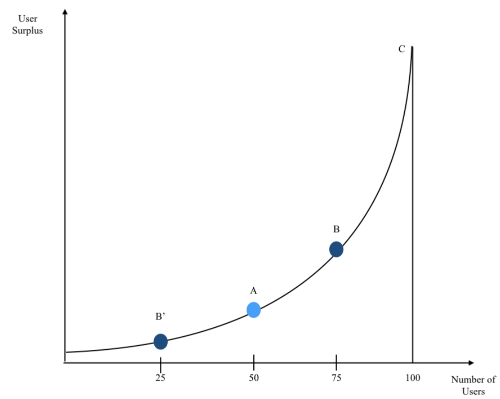
The figures show the landing page of Rover and DogVacay after the merger of the two platforms was completed. The screenshots are accessible on Wayback Machine (`https://web.archive.org/web/20170714115852/https://www.rover.com/` *and* `https://web.archive.org/web/20170704144306/https://dogvacay.com/`*). In July 2017 (right panel), DogVacay users could migrate to Rover by clicking on "Migrate Your Account" at the top.*

Figure D.2: Multi-Homing



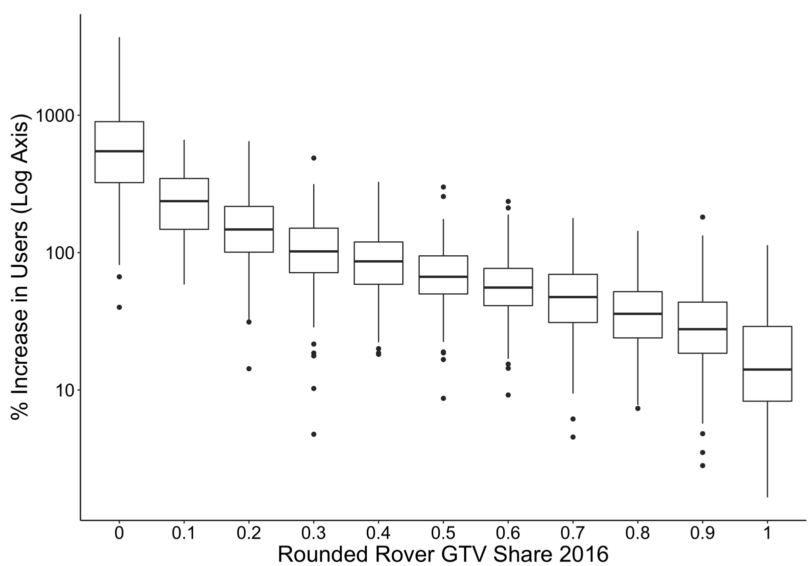
The figures plot the distribution of transactions between Rover and DogVacay for users active before the acquisition. On the left panel, an observation is a user (buyer in light, seller in dark). The histogram plots the share of users' transactions occurring on DogVacay. Users at 1 are those who only transacted on DogVacay prior to the acquisition, while those at 0 only transacted on Rover. Those in between multi-home, i.e. transact on both platforms prior to the acquisition. The right-hand panel weighs each seller by the number of transactions. The comparison between the left and right plots shows that multi-homing users transact more than single-homers.

Figure D.3: Network Effects and Aggregate Surplus



The figure plots user surplus as a function of market size, in a setting with network effects. Network effects imply that the surplus curve is increasing and convex. Point A denotes a city where Rover has 50 users, while point B denotes a city where Rover has 75 users. Both cities have a total of 100 users, so in city A DogVacay also has 50 users, while in city B DogVacay has 25 users (point B'). The merger leads Rover to have 100 users in both cities (point C), but the change in aggregate surplus is bigger in city A than in city B. Let y denote surplus. Because of network effects we have that $y_{B'} + y_B > 2y_A$, i.e. the sum of the surplus generated by DogVacay and Rover pre-merger is bigger in city B than in city A. So $y_C - 2y_A > y_C - y_{B'} - y_B$.

Figure D.4: Transactions from DogVacay Users as Share of Prior Rover Users



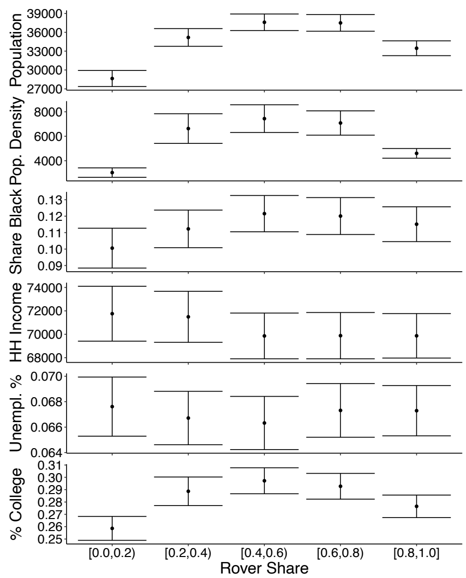
Box plot of the percentage change in the number of transacting users post-acquisition due to DogVacay users switching to Rover as a function of Rover market shares in 2016. Specifically, the percentage change in users is the number of DogVacay users who migrated their profiles to Rover and transacted after '2017-04-01' over the number of Rover users transacting between '2016-01-01' and '2017-04-01'. The zip code's Rover market share is defined using gross transaction volume and is rounded to the nearest 0.1.

Table D.1: First Movers and Rover Market Share

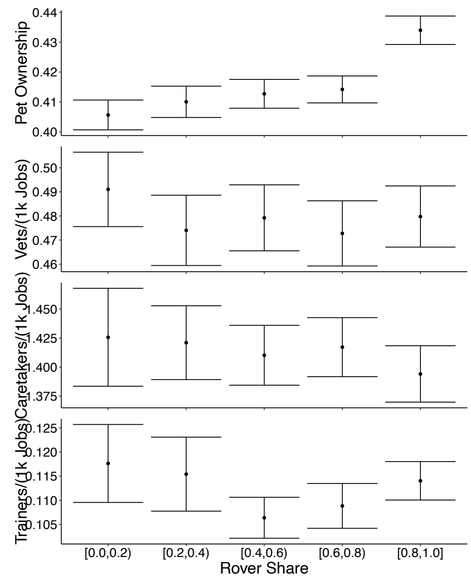
	Dependent variable:			
	2016 Rover Market Share			
	(1)	(2)	(3)	(4)
1{First Mover = Rover}	0.081***	0.078***	0.069***	0.071***
	(0.007)	(0.007)	(0.007)	(0.007)
State FE	N	Y	N	N
CBSA FE	N	N	Y	Y
Year Month FE	N	N	N	Y
Observations	8,200	8,200	8,200	8,200
R^2	0.017	0.055	0.155	0.162
Adjusted R^2	0.017	0.049	0.124	0.125
Note:				*p<0.1; **p<0.05; ***p<0.01

*The table displays the OLS estimates of Rover's market share in 2016 on whether Rover was the first mover in the market for all markets where both Rover and DogVacay had at least one transaction before 2016 and the market had more than 50 transactions during 2016. Each market is a zip code. Rover is defined to be the first mover in the market if the first transaction was booked on Rover. Results also hold for when the first mover is defined to be the first platform to reach 10 transactions in the market. *p<0.1; **p<0.05; ***p<0.01.*

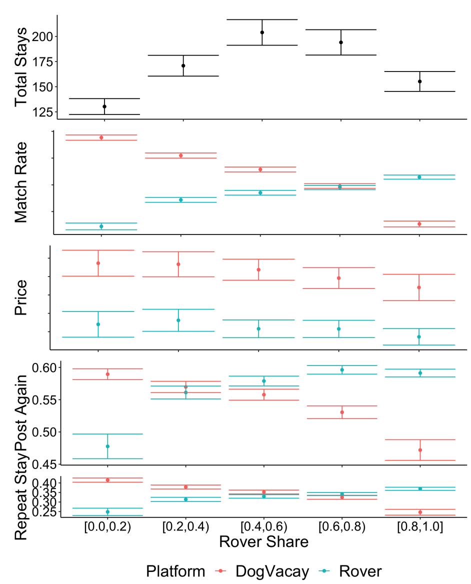
Figure D.5: Differences Across Zip Codes



(a) Population Demographics.



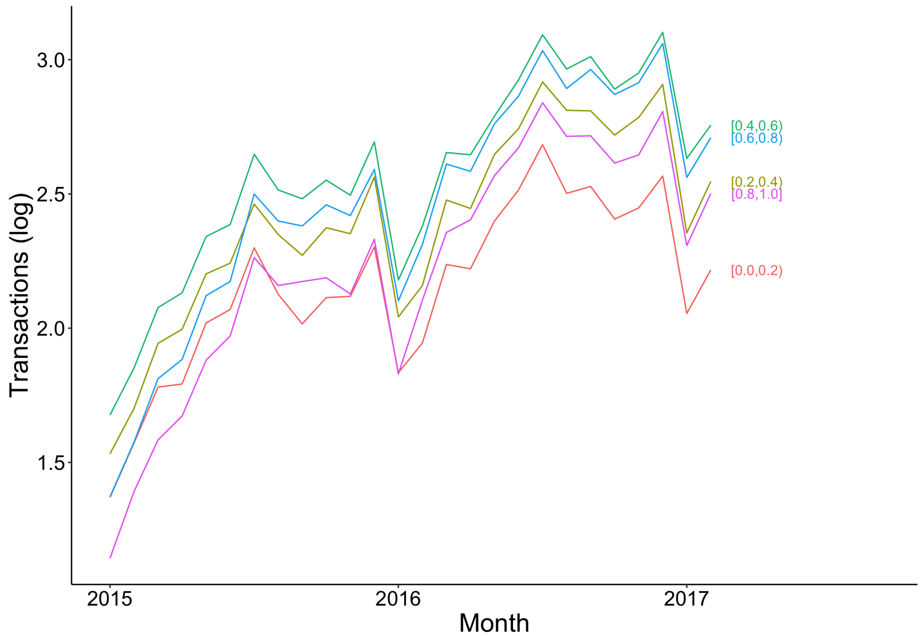
(b) Pet Ownership and Services.



(c) Platforms' Performance.

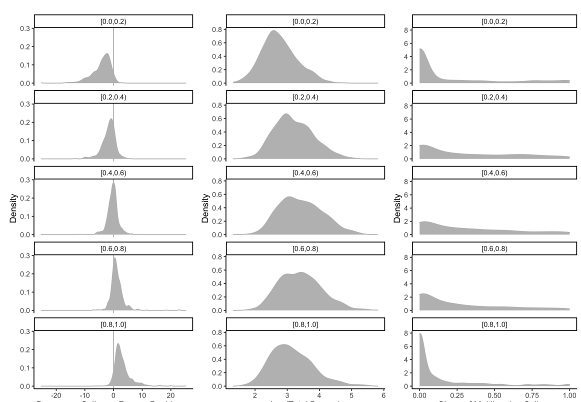
Differences across zip codes in population demographics (left), pet ownership and services (right), and Rover and DogVacay's performance (bottom). Each zip code is grouped by market share – the groupings are defined in Figure 4. The plot on the left shows average population demographics within each market group: population and population density, share of black residents, median household income, unemployment rate, share of the population with a college degree. The plot on the right shows the share of households with pets, as well as jobs related to pet services: number of veterinarians, animal caretakers, and animal trainer per 1,000 jobs. Data come from the 2016 American Community Survey and Bureau of Labor Statistics Occupational Employment Statistics. The plot on the bottom shows average (Rover + DogVacay) stays, ad well as other performance metrics broken down by platform: price; match rates; share of buyer requesting again within 3 months; share of buyers transacting again with the current seller (conditional on the current transaction being a new relationship). Vertical bars correspond to 95% confidence intervals. The absolute levels of price and match rates are omitted to protect company information.

Figure D.6: Matches over Time



The figure plots the average number of monthly stays across market share groups.

Figure D.7: Heterogeneity Across Market Share Groups



The figure plots the density of three dimensions of heterogeneity across markets. An observation is a zip code, and zip codes are divided across rows depending on Rover's market share in 2016. The left column plots the 2016 monthly average difference in the number of active buyers per seller on Rover compared to DogVacay in a zip code. The center column plots the monthly average number of (Rover + DogVacay) buyers in a zip code. The right column plots the share of sellers in a zipcode who transacted on both platforms in 2016.

Table D.2: Comparison Across Market Share Groups

	[0.8,1.0]	[0.0,0.2)	[0.2,0.4)	[0.4,0.6)	[0.6,0.8)
Panel A: Population Demographics					
Population	33,463	-4,815***	1,717*	4,131***	4,032***
Land Area (sq. miles)	22.58	10.54***	1.67	-0.17	-3.12
Population Density	4,600	-1,572**	2,028***	2,839***	2,482***
Share Asian	0.09	-0.03***	-0.01***	-0.01**	-0.00
Share Black	0.12	-0.01*	-0.00	0.01	0.00
Share White	0.70	0.07***	0.02**	0.01	-0.01
Average Income ($)	87,898	2,496	2,787	-179	-292
Median Income ($)	69,872	1,888	1,621	-11	12
Unemployment Rate	0.07	0.00	-0.00	-0.00	0.00
Share Uninsured	0.10	-0.00	0.00	0.01**	0.01*
Share Non Citizen	0.09	-0.02***	-0.00	0.00	0.01***
Share with College	0.28	-0.02**	0.01*	0.02***	0.02**
Share Poor	0.04	-0.00	-0.00	-0.00	-0.00
Share with Pets††	0.43	-0.03***	-0.02***	-0.02***	-0.02***
Vets/1,000 jobs††	0.48	0.01	-0.01	-0.00	-0.01
Animal Caretakers/1,000 jobs††	1.39	0.03	0.03	0.02	0.02
Animal Trainers/1,000 jobs††	0.11	0.00	0.00	-0.01*	-0.01
Panel B: Market Performance					
Stays	155	-25***	16**	49***	39***
Nightly Price (log $)†	–	0.09***	0.07***	0.05***	0.03***
Match Rate†	–	0.11***	0.03***	-0.01	-0.02***
Share Repeat Transactions	0.48	0.09***	0.00	-0.02***	-0.03***
Share Requesting Again	0.58	-0.00	-0.02***	-0.01***	-0.00
Share Transacting with Same Sitter	0.36	0.03***	-0.01	-0.02***	-0.02***
Panel C: Rover Performance					
Stays	141	-128***	-86***	-35***	-4
Nightly Price (log $)†	–	0.01	0.02**	0.01	0.01
Match Rate (rel. to Panel B)†	0.02	-0.18***	-0.09***	-0.06***	-0.04***
Share Repeat Transactions	0.49	-0.22***	-0.09***	-0.05***	-0.03***
Share Requesting Again	0.59	-0.11***	-0.03***	-0.01*	0.00
Share Transacting with Same Sitter	0.37	-0.12***	-0.06***	-0.04***	-0.03***
Panel D: DogVacay Performance					
Stays	14	103***	102***	83***	42***
Nightly Price (log $)†	–	0.04***	0.04***	0.04***	0.03**
Match Rate (rel. to Panel B)†	-0.16	0.32***	0.26***	0.20***	0.14***
Share Repeat Transactions	0.27	0.32***	0.25***	0.21***	0.14***
Share Requesting Again	0.47	0.12***	0.10***	0.09***	0.06***
Share Transacting with Same Sitter	0.25	0.17***	0.13***	0.11***	0.08***
N	793	577	560	639	692

The table compares zip code-level demographics and platform performance across market share groups. Demographics data are obtained from the US Census Bureau. For each of the characteristics, the first column displays the average value in the control group. The other columns display the difference of a particular market share bin compared to the control group, and whether the difference is statistically significant at standard confidence levels. Panels separate variables into the following 4 groups: population demographics; aggregate platform performance (Rover + DogVacay); Rover performance; and DogVacay performance. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

† *:The level of nightly price is not displayed for the control group to protect company information. We only show log differences across market share groups. Analogously, the match rate is not displayed for the control group in Panel B. For Panel C and D the control group column displays the percentage point difference in match rates between the zip code average match rate and the match rates in each of the two separate platforms.*

†† *: CBSA-level variables. Each zip code is assigned the value of its CBSA, and then mean and standard deviation are computed with zip code as units of observation.*