Cyclical Tacit Collusion in the Hotel Industry

Christian Rojas‡

[Preliminary]

Abstract

I use daily price and occupancy data from the Manhattan lodging industry to test: a) the existence of tacit collusion, and b) the predictions of supergame models of collusion under cyclical demand and capacity constraints. First, I estimate a model that explores whether supracompetitive pricing exists among hotels of different quality tiers. Consistent with tacit collusion, I find that the luxury segment of the market is characterized by two distinct regimes: a high-price and low-occupancy ("collusive") regime and a low-price and high-occupancy ("competitive") regime (each occurring 50% of the time). Second, I use the model to investigate the role that the cyclicality of demand and the capacity constraints of the market play in: a) the probability with which the collusive regime occurs, and b) the intensity of collusion (i.e. the difference between the collusive and competitive prices). I find widespread support for the countercyclical collusion predictions set forth by Haltiwanger and Harrington (1991): collusion is more likely and more intense when demand is lowest and a boom is starting and least likely as the recession approaches.

‡ Professor, Department of Resource Economics, University of Massachusetts Amherst. I am grateful to Duane Vinson at STR for providing the hotel data.

1. Introduction

A large theoretical body has studied the mathematical conditions under which collusion can be sustained as a non-cooperative equilibrium in an infinitely repeated setting. Comparatively speaking, empirical work testing the existence of collusion in these dynamic environments has been notably limited. Part of the challenge for empiricists is the inability to control (or account) for the different market conditions and assumptions given in theoretical models. Importantly, the feasibility of a particular equilibrium (or equilibrium path) often depends crucially on the assumptions. Given the detrimental effects of collusion and its central role in antitrust discussions, empirical work geared at discerning the observationally likely noncooperative equilibria from the mathematically feasible set seems particularly valuable. This paper takes a step in filling this gap.

I focus on supergames that consider two important and frequent aspects of markets: cyclicality of demand and capacity constraints. Under demand cyclicality, theory (Haltiwanger and Harrington, 1991) predicts that collusion is easiest in periods right before a downturn occurs, while collusion is hardest at the beginning of an upturn. These predictions are reminiscent to those of Rotemberg and Saloner (1986), where booms are associated with collusion breakdowns. However, when capacity constraints are important, the Haltiwanger and Harrington predictions can be reversed (Fabra, 2006; Knittel and Lepore, 2010).

The hotel industry exhibits predictable demand cycles and, in some markets, capacity constraints are particularly binding. The Manhattan hotel market is a prime example of these features, as I detail in the data section. Further, I focus on the luxury segment of the market, which serves a particular niche of the market. I carry out a mixture model estimation that allows me to identify two regimes in the (price and occupancy) data. One regime is characterized by high-price and low-occupancy while the other is characterized by low-price and high-occupancy; importantly, when non-luxury segments of the market are exposed to the estimation technique, the regimes obtained are not consistent with the patterns observed in the luxury segment.

In addition to finding support for the existence of tacit collusion, the econometric procedure allows me to test whether the cyclical movement of prices and the probability of the collusive regime are consistent with the predictions of the theory. I find strong support for the predictions set forth by Haltiwanger and Harrington (1991): collusion is more likely and intense in the lowest demand periods right before an upturn begins whereas collusion is least likely and least intense in the peak demand periods right before a downturn occurs.

3. Data

The data is provided by STR, a firm that provides data analytics services to the hospitality sector worldwide. Hotels that subscribe to STR's services submit their daily price and occupancy data to STR which, in turn, uses the data to provide industry reports and specific consulting services to member hotels. STR makes its data available for academic purposes in an anonymized format. An important aspect is that STR offers near universal coverage of the lodging industry since (in most markets) almost all hotels submit their data to STR.

For our study, we obtained data on 178 hotels in the Manhattan district in New York City, with an aggregate room capacity of 62,199, comprising approximately 80% of the hotel room supply in Manhattan.[1] We data spans from January 1, 2013 through December 1, 2018 (2,190 days) and includes the average daily room price ("average daily rate" labeled as ADR by STR). STR also provides variables that capture several aspects of the hotels (total number of rooms, whether it is a franchised hotel, year of opening, etc.). An important variable for the analysis is the hotel's scale ("class"): economy (E), midscale (M), upper midscale (UM), upscale (U), upper upscale (UU), and luxury (L). There are few hotels in the economy and midscale tiers (2 and 4, respectively) and most hotels are distributed in the remaining categories (UM 26, U 57, UU 51, L 38). The analysis below focuses on the luxury segment.

Table 1 provides summary statistics on the two key variables of interest. As expected, higher class hotels are pricier; however, the difference in mean price across two neighboring hotel classes is more or less similar (between $18 and $45) with the notable exception of the Luxury segment, which displays an average price of $170 over the closest class (UU). In addition, the Luxury segment has a wider price dispersion (Luxury's coefficient of variation is 0.51 compared to 0.27-0.35 for other classes). These figures suggest that the Luxury segment is a market that caters to a particular clientele and therefore may be significantly differentiated from other hotel classes. Another distinct aspect of Luxury hotels is given by the type of operation.

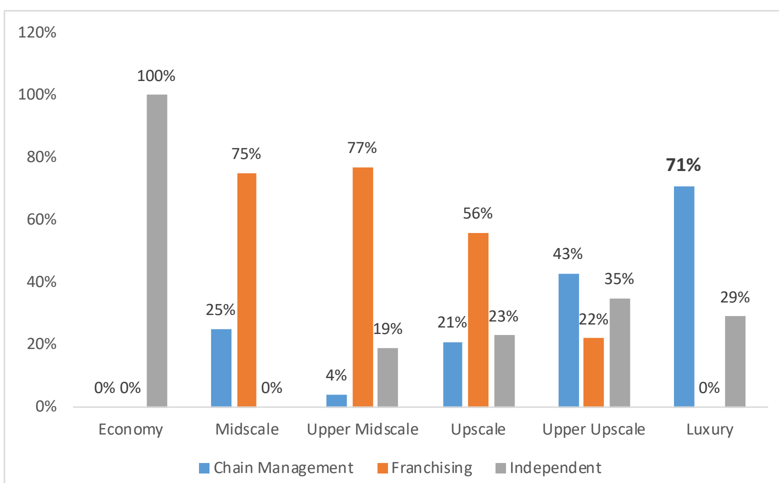
Table 1: Summary Statistics of ADR and Occupancy

Hotel Class	**ADR**					**Occupancy (%)**			
	# Obs	**Mean**	**S.D.**	**Min**	**Max**	**Mean**	**S.D.**	**Min**	**Max**
Economy	4,382	171.57	59.55	62.01	493	90.45	11.85	37.54	100
Midscale	8,764	191.12	51.25	74.34	474	91.11	11.11	25.25	100
Upper Midscale	51,961	209.54	63.94	30	753	91.96	11.40	1.42	100
Upscale	112,233	238.86	77.14	35.53	4,452	90.28	13.35	0.13	100
Upper Upscale	102,673	283.18	82.39	52.86	1,618	87.22	15.28	0.51	100
Luxury	76,428	452.19	230.95	90.76	6,180	82.11	17.23	0.61	100

[1] STR has data for the remaining 20% of the market, but for confidentiality reasons could not provide it for academic purposes. The hotels in our sample were not randomly selected; instead, we requested that the 80% coverage corresponded to that of the largest hotels in the data. The 62,199 figure is the average over time; total capacity increases over time (mostly because of entry, the remaining portion because of expansion) from a minimum of 54,918 to a maximum of 66,077.

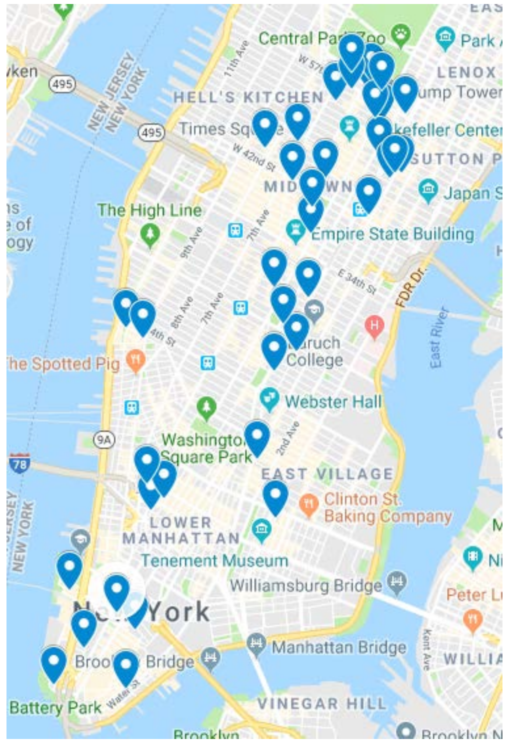
Another distinct aspect of Luxury hotels is given by the type of operation. Hotels can be operated under three different formats: chain management, franchising or independent. The first two formats are associated to hotel brands that belong to a large chain (e.g. Marriot, Sheraton, etc.), with the difference that management is carried out by the chain in the first format but delegated to a third party (a franchisee) in the second format. Independent operations correspond to hotels that do not belong to a hotel chain. Figure 1 shows the distribution of operation type by hotel class. The Luxury segment is characterized by having the large majority of hotels being operated under the chain management format (27 out of 38) and no franchised operations. This contrasts with other segments where most prevalent type of operation is franchising (M, UM, U) or independent (E); while chain management is also the most frequent type of operation in UU hotels, the rate is much smaller (also franchising has a sizable share in UU, at 22%). This evidence suggests that hotel chains have a policy of not delegating management of their hotels when operating in the Luxury segment and that competition occurs, by and large, among chain hotels (with independents accounting for less than a third of the hotels in the segment).[2] Luxury hotels are dispersed over the lower portion of Manhattan (south of Central Park), with some clustering in certain areas (notably in Midtown/Times Square district, adjacent to Central Park).

Figure 1: Distribution of Hotels' Operation Type, By Class



[2] The difference is more pronounced when using the number of rooms available: 18% independent v. 82% luxury.

Figure 2: Geographical Location of Luxury Hotels



Seasonality and Capacity Constraints

The Manhattan hotel industry is characterized by two features that are central to the analysis. The first is related to the predictable cyclicality of demand. The second has to do with the importance of capacity constraints. I provide descriptive evidence of these two features.

There are two types of demand cycles. The first one has a larger span and occurs over months of the year. The second one occurs over days of the week.[3] Figure 3 depicts the average monthly ADR and occupancy rates across all hotels over the six years of data. The Figure highlights several recurring patterns. First, price and occupancy move largely in tandem (correlation coefficient 0.72); the months for which this positive correlation seems weaker are the summer months of July and August. The positive correlation between price and occupancy is consistent with demand-side movements.

Second, the observed cycles are consistent with times of the year when one expects demand to be higher (lower). For instance, the cold winter months of January and February make it less appealing for travelers to visit the city (less so in December given the numerous festivities and shopping season); May and September are popular months because of ideal weather. July and

[3] The following figures using data from all hotels in our sample. Similar figures are obtained for the Luxury segment (and other segments).

August are popular months for tourists (which make the bulk of visitors), which makes occupancy increase while prices remain relatively low (price insensitive business travelers are least likely to visit the city during these months).

The hotel industry is characterized by a cost structure composed of large fixed costs that do not vary over time. In addition, marginal costs in this industry are unlikely to be too sensitive to output (occupancy rate). In any case, to the extent that supply-side shocks are responsible for these movements, there is no obvious explanation for why they should occur in the cycles portrayed by the data, nor are they consistent with a positive correlation between prices and occupancy.

Figure 3: Average Monthly ADR and Occupancy Rate, All Hotels

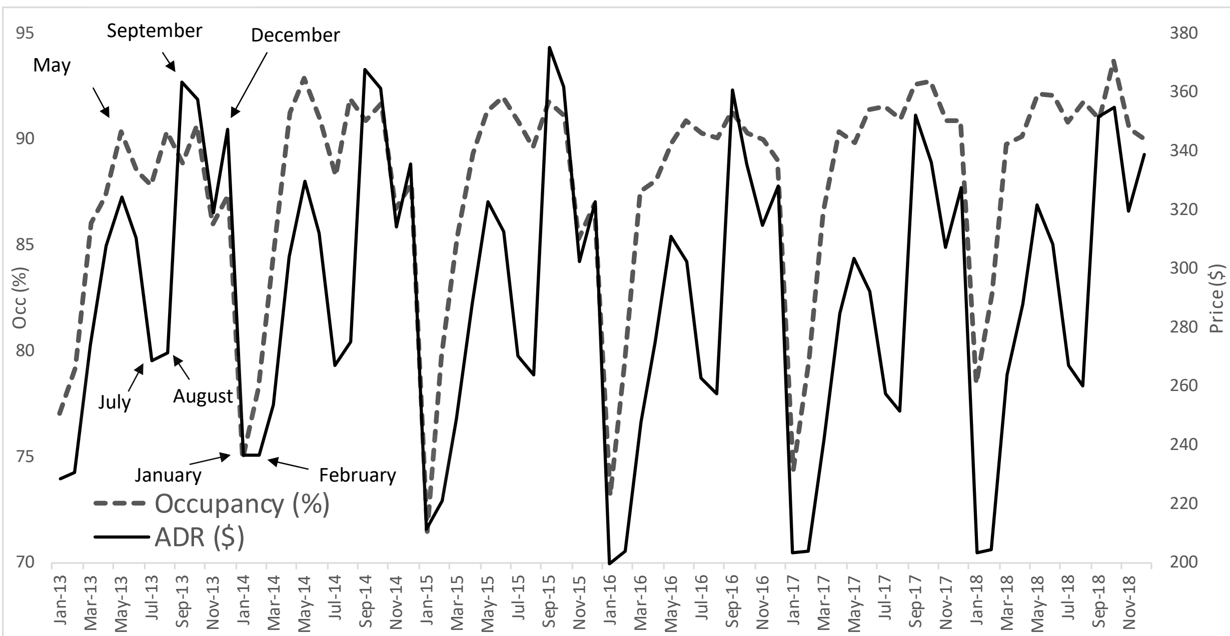


Figure 4 provides evidence for the second type of cycle. The Figure displays, for a subset of months (to conserve clarity), the monthly ADR (left panel) and occupancy over different days of the week. Several clear patterns are observed. First, Sunday consistently exhibits the lowest price and lowest occupancy over the week (although this is less evident for January), which is indicative that demand is the lowest in this day. Second, in all displayed months (except for December), Tuesday and Wednesday appear to be the days with the highest demand as prices and occupancy are the highest. Third, Saturday seems to be the most popular weekend day, although this is more clearly evidenced by the occupancy rate.

Figure 4: Monthly ADR (left panel) and Occupancy (right panel) by Day of Week

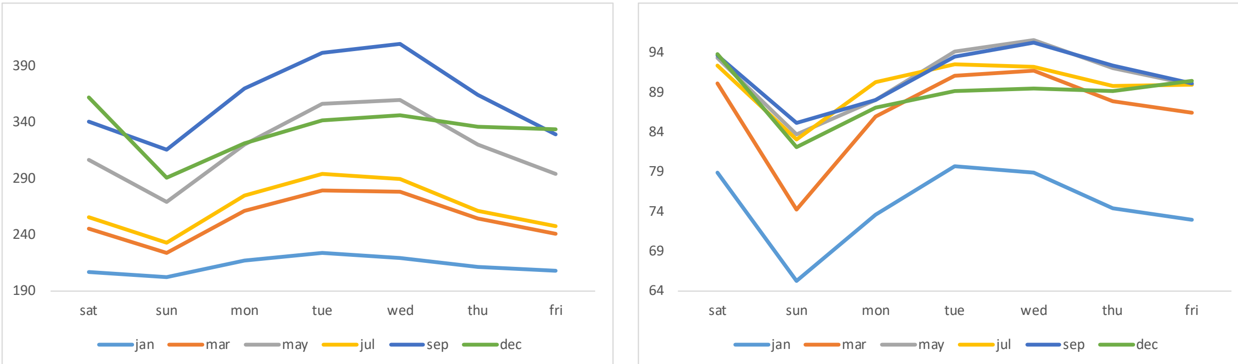


Table 2: Daily Sell-Out Rates and Average Monthly Idle Market Capacity, by Month

Month	*% Hotel-Days Sold Out*	*Average Monthly Idle Capacity of Market*	*% Hotel-Days Sold Out*	*Average Monthly Idle Capacity of Market*
	All Hotels		Luxury Hotels	
January	7.6%	14.4%	2.9%	24.6%
February	9.7%	12.6%	3.4%	26.4%
March	14.9%	12.0%	6.8%	6.6%
April	20.1%	6.4%	9.7%	4.3%
May	25.7%	5.4%	13.7%	16.5%
June	23.3%	7.3%	11.0%	5.8%
July	23.6%	9.4%	11.8%	25.4%
August	18.0%	2.3%	7.5%	9.5%
September	21.6%	6.3%	11.2%	18.4%
October	23.7%	3.7%	11.9%	12.5%
November	17.8%	7.4%	8.1%	12.0%
December	19.9%	4.8%	9.7%	16.0%

4. Methodology

I propose to use a regime switching (also known as a mixture) model. Since the interest is identifying whether the data is consistent with the existence of a collusive regime and competitive regime, I restrict the analysis to two regimes, denoted by the superscript $s = \{C, NC\}$; $C = collusive; NC = non - collusive$. Further, in each regime, I specify two equations, one for price (p) and one for occupancy (o); this allows me to gauge whether, conditional on the data supporting the existence of two regimes, one regime is high-price and low-occupancy (consistent with collusion) and the other is low-price and high-occupancy (consistent with collusion). I first explain how the model is set up in the absence of regime switching (which I explain at the end of this section).

Given the key role of day-of-the-week and month-of-the-year in the cyclicality of demand, the explanatory variables the two equations I consider are month and day fixed effects. The two equations are given by:

$$\ln(p_{it}) = \delta_1^s + day_t\boldsymbol{\delta}_d^s + month_t\boldsymbol{\delta}_m^s + \varepsilon_{it}^s$$

$$\ln(o_{it}) = \alpha_1^s + day_t\boldsymbol{\alpha}_d^s + month_t\boldsymbol{\alpha}_m^s + u_{it}^s$$

where i indexes hotels and t is time; $month$ and day are vectors containing indicator variables for day-of-the-week and month-of-the-year (respectively); $\boldsymbol{\delta}_d$ and $\boldsymbol{\delta}_m$ are their corresponding parameter vectors. To simplify notation, I henceforth define: $\delta_1^s + day_t\boldsymbol{\delta}_d^s + month_t\boldsymbol{\delta}_m^s = X_t\delta^s$ and $\alpha_1^s + day_t\boldsymbol{\alpha}_d^s + month_t\boldsymbol{\alpha}_m^s = X_t\alpha^s$ (i.e. X_t is a vector that contains all fixed effects and the constant, and δ^s, α^s are the corresponding vectors of parameters). I assume that the error terms (ε and u) are bivariate normal: $(\varepsilon_{it}^s, u_{it}^s) \sim N_2(0,0,\sigma_\varepsilon^{s2}, \sigma_u^{s2}, \rho^s)$, where $\rho^s = \sigma_{\varepsilon u}^s / \sigma_\varepsilon^s \sigma_u^s$. The bivariate normal density is given by:

$$\varphi^b(\varepsilon^s, u^s) = \frac{1}{2\pi\sigma_\varepsilon^{s2}\sigma_u^{s2}\sqrt{1-\rho^{s2}}} exp\left(\frac{-z^s}{2(1-\rho^{s2})}\right)$$

where $z^s = \frac{\varepsilon^{s2}}{\sigma_\varepsilon^{s2}} + \frac{u^{s2}}{\sigma_u^{s2}} - \frac{2\rho^s\varepsilon^s u^s}{\sigma_\varepsilon^s\sigma_u^s}$. The object of my estimation is to obtain the model parameters ($\theta^s = \delta^s$, α^s, $\sigma_\varepsilon^{s2}, \sigma_u^{s2}, \rho^s$). I carry out this exercise via maximum likelihood, following the steps described below.

One detail needs to be taken care of before writing the likelihood function that is taken to estimation. The occupancy variable is censored since hotels often sell out. This censoring from above ($o_{it} = 1$) is dealt with a bivariate Tobit specification that allows for one of the two variables to be censored. I define an indicator variable γ_{it} which takes a value of 1 if hotel i sells out in period t and zero otherwise. For a given set of parameters θ^s in regime s, an observation's contribution to the likelihood is given by the joint density (f) evaluated at the observed data:

$$\begin{aligned} L_{it}^s &= f(\varepsilon_{it}^s, u_{it}^s; \theta^s) \\ &= [\varphi^b(\varepsilon_{it}^s, u_{it}^s; \theta^s)]^{(1-\gamma_{it})}[\underbrace{Prob(u_{it}^s \geq -X_t\alpha | \varepsilon_{it}^s; \theta^s)}_{f\left(u_{it}^s|\varepsilon_{it}^s;\theta^s\right) \text{ when selling out}} \times \underbrace{\varphi(\varepsilon_{it}^s; 0, \sigma_\varepsilon^{s2})}_{f\left(\varepsilon_{it}^s;\theta^s\right)}]^{\gamma_{it}} \end{aligned}$$

$$= [\varphi^b(\varepsilon_{it}^s, u_{it}^s; \theta^s)]^{(1-\gamma_{it})} \left\{ \left[1 - \Phi\left(\frac{-X_t\alpha - \frac{\sigma_u^s}{\sigma_\varepsilon^s}\rho^s\varepsilon_{it}^s}{\sqrt{1-\rho^{s2}}\sigma_u^s} \right) \right] \times \varphi(\varepsilon_{it}^s; 0, \sigma_\varepsilon^{s2}) \right\}^{\gamma_{it}}$$

where Φ is the standard normal CDF and φ is the univariate normal density.[4]

To complete the specification of the likelihood, I incorporate the regime switching aspect by specifying a probability h, which determines that an observation belongs to the competitive regime $(1-h)$ and to the collusive regime with probability h. Further, I allow h to be a function of t; since my interest is in determining whether tacit collusion varies over the demand cycle in ways consistent with the theory, I use the same month and day fixed effects as in the price and occupancy equations. Specifically:

$$h_t = \frac{\exp(X_t\beta)}{\exp(1 + X_t\beta)}$$

where X_t includes a constant and the fixed effects just described and β is a vector of parameters to be estimated. Then, the contribution of observation i, t to the log-likelihood is:

$$LL_{it} = ln\left[h_t L_{it}^C + (1-h_t)L_{it}^{NC}\right]$$

[capacity constraints: ongoning]

5. Results

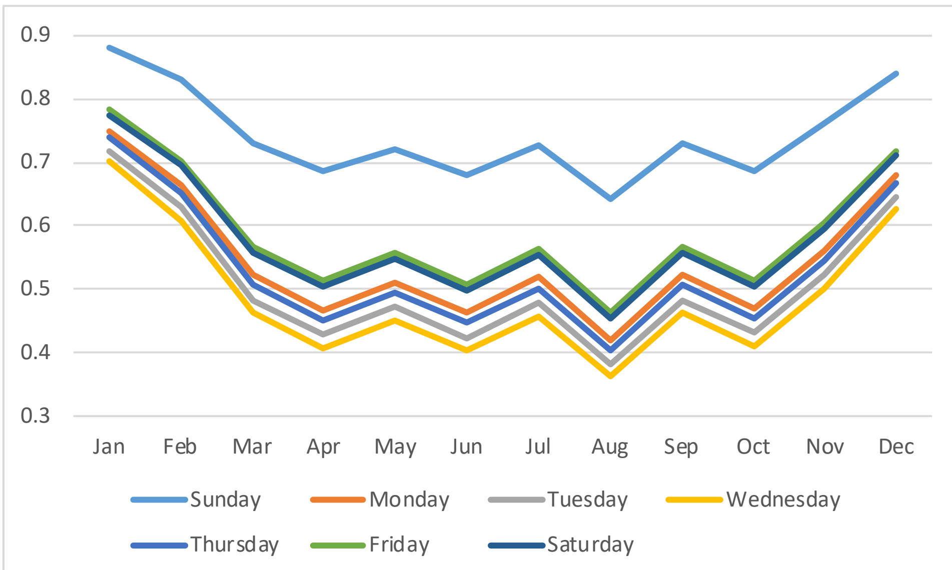
Table 3 displays the results of the estimation. The first thing to note is that the price and occupancy constant coefficients are consistent with the existence of a regime characterized by high price and low occupancy and another by low price and high occupancy. To ease the interpretation of the numerous coefficients. Figures 5, 6 and 7 plot the predicted values. Figure 5 displays the probability of collusion by month (over days of the week). I can be seen that the day with the highest probability of collusion is Sunday (the weakest demand day, right before the busy weekdays start) whereas the day with the lowest probability of collusion is a Wednesday (the peak of the week). Similarly, it can be seen that January and December are the months with the highest probability of collusion; of the two, January is the month with the weakest demand (preceding the upturn over the coming months; I discuss December below).

---

[4] The last equality follows from the fact that if $\varepsilon_{it}^s, u_{it}^s$ are distributed bivariate normal, then $u^s|\varepsilon^s \sim N(\frac{\sigma_u^s}{\sigma_\varepsilon^s}\rho^s\varepsilon^s, (1-\rho^{s2})\sigma_u^{s2}$.

Figures 6 and 7 display the predicted values for price and occupancy. One thing to note is that prices under the collusive regime do not follow closely the cyclicality of demand, a pattern that one would expect if hotels are not competing (prices are known to be less volatile when collusion exists). Second, the difference between the collusive price and the competitive price is smallest in the month of December, the month that precedes the largest "recession" in the year. Finally, the occupancy data suggests that the greatest impact that collusion has is on January; this, again, consistent with the predictions of Haltiwanger and Harrington.

Figure 5: Probability of Collusion over Months, by Day of Week



Capacity Constraints

[in progress]

6. Concluding Remarks

This work in progress studies empirically the possible existence of tacit collusion in the luxury hotel market in Manhattan. In addition to finding important support for the existence of tacit collusion, the evidence suggests that the collusive pricing patterns and the probability of collusion are consistent with the predictions set forth by the Haltiwanger and Harrington (1991).

[capacity constraints: on goning]

Table 3: Maximum Likelihood Results

	δ^C		δ^{NC}		α^C		α^{NC}		h	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	5.736	0.009	5.507	0.008	-0.636	0.005	-0.235	0.004	2.005	0.056
Mon	0.105	0.008	0.185	0.006	0.102	0.005	0.108	0.003	-0.906	0.042
Tue	0.139	0.008	0.237	0.005	0.216	0.005	0.151	0.003	-1.066	0.042
Wed	0.144	0.008	0.239	0.005	0.231	0.005	0.151	0.003	-1.149	0.043
Thurs	0.124	0.008	0.155	0.006	0.156	0.005	0.113	0.003	-0.970	0.042
Friday	0.078	0.008	0.053	0.006	0.119	0.005	0.102	0.003	-0.729	0.042
Sat	0.044	0.008	0.082	0.006	0.207	0.005	0.138	0.003	-0.766	0.043
Feb	0.031	0.010	0.023	0.008	0.072	0.006	0.000	0.003	-0.415	0.055
Mar	0.130	0.011	0.123	0.007	0.179	0.006	0.019	0.003	-1.005	0.052
Apr	0.260	0.011	0.227	0.007	0.240	0.007	0.036	0.003	-1.227	0.052
May	0.307	0.010	0.345	0.007	0.278	0.006	0.055	0.003	-1.051	0.052
June	0.293	0.011	0.294	0.007	0.249	0.007	0.045	0.003	-1.249	0.052
July	0.136	0.010	0.159	0.007	0.207	0.007	0.059	0.003	-1.025	0.052
Aug	0.194	0.011	0.112	0.007	0.221	0.007	0.036	0.003	-1.422	0.051
Sept	0.456	0.010	0.456	0.007	0.298	0.006	0.048	0.003	-1.006	0.053
Oct	0.421	0.010	0.428	0.007	0.286	0.007	0.049	0.003	-1.216	0.052
Nov	0.332	0.010	0.371	0.008	0.234	0.006	0.041	0.003	-0.850	0.053
Dec	0.345	0.010	0.456	0.010	0.260	0.006	0.052	0.003	-0.343	0.056

#Obs: 75,703. LL = -10076.97

Figure 6: Predicted APR over Days of the Week, by Month

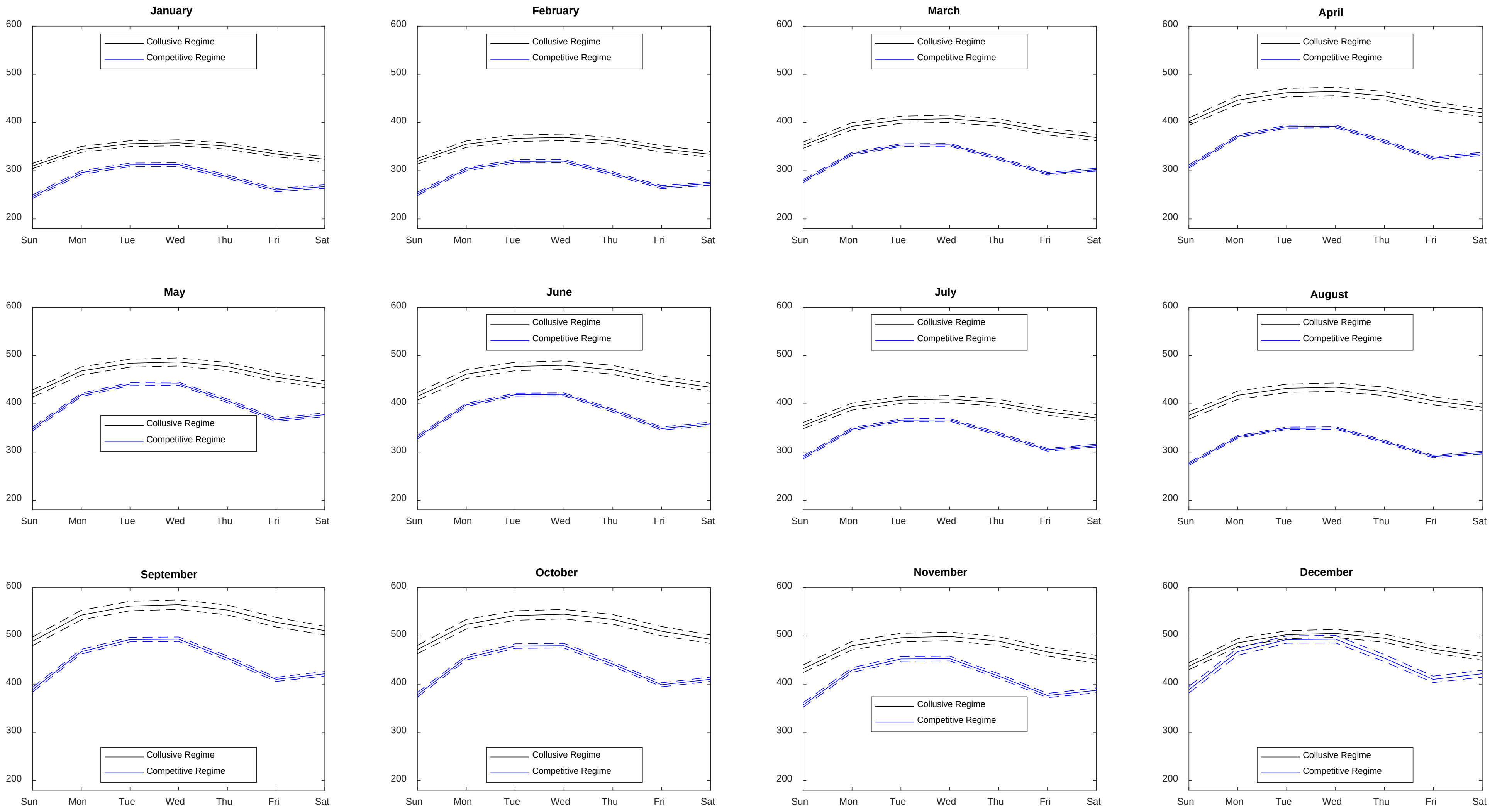
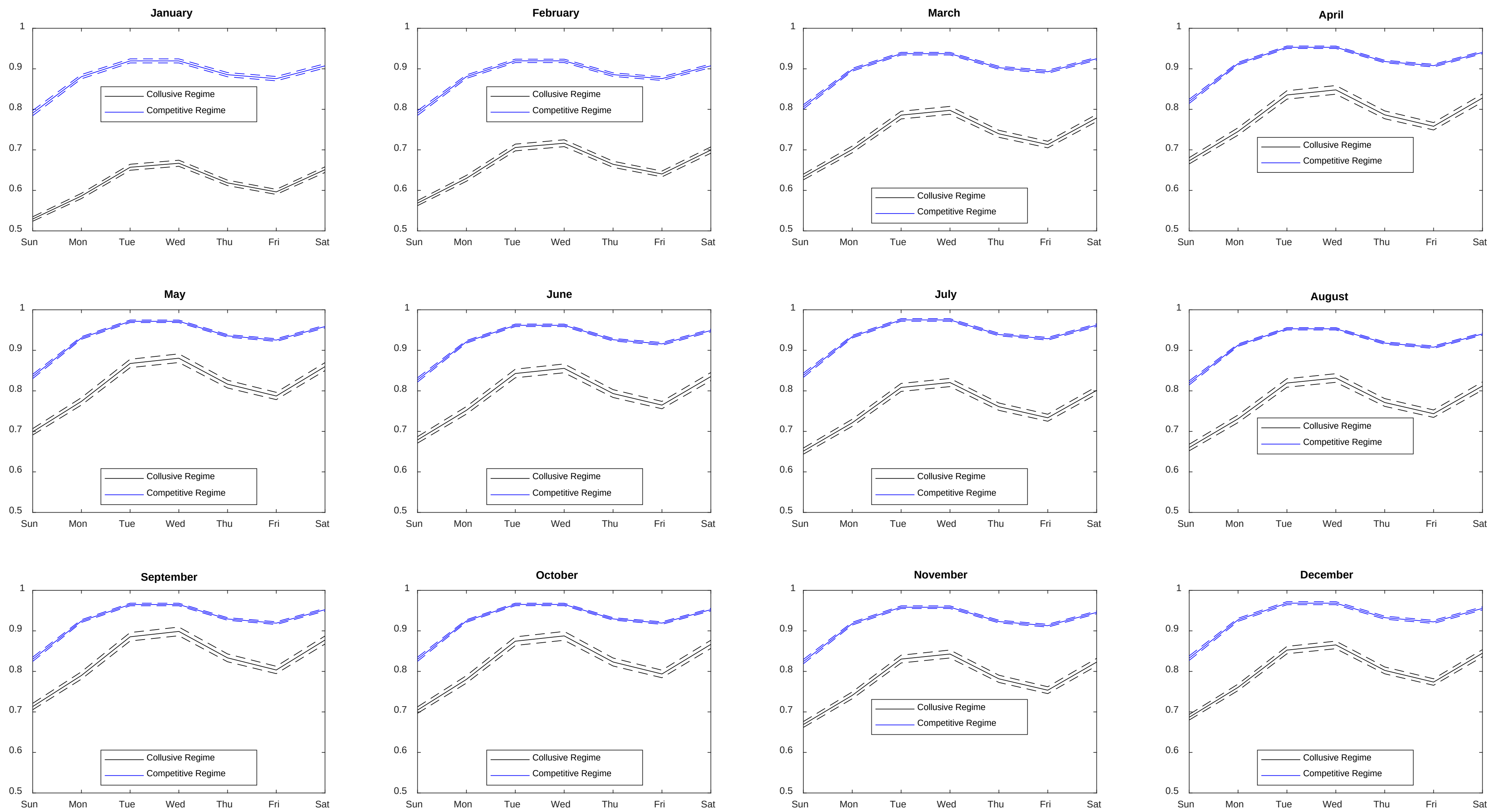


Figure 7: Predicted Occupancy over Days of the Week, by Month



References

Borenstein, S. and Shepard, A. 1996. "Dynamic Pricing in Retail Gasoline Markets." *RAND Journal of Economics*, 27: 82-106.

Gan, L. and M. Hernandez. 2013. "Making Friends with your Neighbors? Agglomeration and Tacit Collusion in the Lodging Industry," *Review of Economics and Statistics*, 95: 1002-17.

Fabra, N. 2006. "Collusion with Capacity Constraints over the Business Cycle." *International Journal of Industrial Organization*, 24: 69-81.

Haltiwanger, J. and Harrington, J. 1991. "The Impact of Cyclical Demand Movements on Collusive Behavior." *RAND Journal of Economics*, 22: 89-106.

Knittel, C. and J. Lepore. 2010. "Tacit collusion in the presence of cyclical demand and endogenous capacity levels," *International Journal of Industrial Organization*, 28: 131-44.

Knittel, Christopher R., and Victor Stango. 2003. ''Price Ceilings as Focal Points for Tacit Collusion: Evidence from Credit Cards,'' *American Economic Review*, 93: 1703–29.

Rotemberg, J. and Saloner, G., 1986. "A Supergame-Theoretic Model of Price Wars during Booms." *American Economic Review*, 76: 390-407.