Can the World Cup Move Mexico City's Economy?

Lessons from F1, NFL, and Major Events

Joel J. D. Gonzalez Gomez

Independent Working Paper

June 2025

Pre-Tournament Analysis

Abstract

Written one year before the 2026 FIFA World Cup, this working paper studies whether major event months are associated with measurable short-term changes in Mexico City's tourism and hotel market. The pre-tournament analysis is motivated by the five matches scheduled for Mexico City Stadium, including the tournament's opening match. Using monthly hotel occupancy and tourist-arrival data for Mexico City from January 2019 to April 2024, I estimate whether Formula 1 and NFL event months are associated with higher hotel occupancy and tourists hosted in hotels, controlling for calendar seasonality, year effects, and the COVID-19 shock. The pre-tournament results should be interpreted as preliminary conditional associations rather than causal estimates. Event months are associated with visibly higher tourism indicators in descriptive comparisons. In the main regression, Formula 1 months are associated with a 6.7 percentage-point increase in hotel occupancy, while a combined major-event indicator is associated with a 7.0 percentage-point increase. Robustness checks keep the event coefficient positive, though statistical significance weakens when COVID-shock months are excluded. The effects on total tourists hosted in hotels are weaker and become negative after month and year controls, suggesting that occupancy captures short-run hotel-market pressure more clearly than aggregate tourist counts. The paper concludes that major events can move Mexico City's hotel market at the margin, but the magnitude and distribution of World Cup effects will depend on substitution, displacement, pricing, and whether visitors extend their stays beyond match days.

Contents

1 Introduction

Mega-events are often presented as economic accelerators. Governments and event organizers describe them as sources of visitors, hotel nights, consumption, international visibility, and long-term city branding. Yet the empirical question is more modest and more interesting: do these events actually show up in local economic data? This paper studies that question for Mexico City in June 2025, with one year remaining before the 2026 FIFA World Cup.

Mexico City is a useful case. It is already one of Latin America's largest tourism and business hubs, so any incremental event effect must be separated from ordinary seasonality, business travel, domestic tourism, and the city's broader post-pandemic recovery. It also regularly hosts internationally visible events, including the Formula 1 Mexico City Grand Prix and NFL regular-season games. These events provide a practical setting for studying whether event months are associated with changes in hotel occupancy and tourism activity. The World Cup cannot yet be evaluated ex post in the available monthly tourism series, but its expected local demand shock can be discussed using past major events as revealed evidence.

The title question - whether the World Cup can move Mexico City's economy - is intentionally broad. The empirical question in this paper is narrower: whether previous major sports-event months were associated with measurable changes in hotel-market indicators. The underlying idea is that hotel occupancy can serve as an observable, high-frequency proxy for short-run visitor pressure. Tourism is not the entire city economy, but hotels are one of the sectors most directly exposed to event-driven visitor demand.

The analysis uses monthly data on hotel occupancy and tourists hosted in hotels in Mexico City. The primary accessible series covers January 2019 to April 2024. Public CDMX Open Data resources identify the hotel occupancy dataset as monthly information on the average percentage of hotel occupancy in Mexico City and report a CSV resource updated in September 2024. The companion tourism dataset reports monthly domestic and international tourists hosted in hotels in Mexico City. Federal DataTur also provides a broader hotel-monitoring database for 70 tourist destinations, including variables such as hotel occupancy and related hotel-demand indicators. Because the accessible data in this version cover January 2019 to April 2024, the results should be read as a first empirical estimate rather than a definitive long-run assessment.

The paper contributes in three ways. First, it translates a timely public debate into a transparent empirical exercise. Second, it applies a simple event-month specification that is interpretable for non-specialist readers, making the paper suitable for a professional research note. Third, it establishes a compact empirical framework that can be updated as richer tourism and hotel-market data become available.

2 Literature Review

The literature on mega-events is divided between optimistic views of event-led development and more skeptical ex post evidence. Early public narratives often emphasize gross spending, visitor inflows, and media exposure, while academic work tends to stress opportunity costs, substitution, displacement, and the difference between gross and net effects.

Matheson (2006) provides a widely cited overview of the economics of sports mega-events and argues that ex ante impact estimates often overstate net local benefits. The mechanism is intuitive: some spending attributed to the event may have occurred anyway, local residents may substitute event spending for other local consumption, and ordinary visitors may avoid the city during periods of congestion or price increases. This distinction matters for Mexico City. An increase in hotel occupancy during a Grand Prix or NFL game does not automatically imply a proportional increase in total welfare or long-run GDP. It does, however, indicate that the hotel market experiences a measurable demand shock.

A more direct strand of literature studies hotel demand around local events. Depken and Stephenson (2018) examine daily hotel occupancy data in Charlotte, North Carolina, and find that major political conventions and NASCAR races are associated with large increases in hotel occupancy, room prices, and revenue, while many smaller events have more limited effects. This is closely related to the present paper because it focuses on hotel-market outcomes instead of broad macroeconomic aggregates. It also supports the methodological choice of using hotel occupancy as a direct indicator of the economic channel most likely to respond to event visitors.

Spatial displacement is another important theme. Humphreys and Zhou (2019) analyze events at the Staples Center in Los Angeles and find that nearby hotels may benefit while hotels farther away may experience weaker or even negative effects. Their findings imply that event effects can be highly localized. Since this paper uses citywide Mexico City indicators, the estimated coefficients may understate effects near the venue and overstate effects for areas less exposed to event demand. In other words, a city-level average may smooth out stronger neighborhood-level effects around Autodromo Hermanos Rodriguez, Estadio Azteca, or central hospitality districts.

The literature also shows that not every sporting event delivers the expected hotel-market gains. Sun (2012) studies a hallmark sporting event and asks why hotel rooms were not full despite expectations and planning. The paper highlights the possibility that mega-event demand can be weaker than official narratives suggest if high prices, crowding, travel frictions, or mismatched visitor expectations discourage ordinary tourists. This is relevant for the World Cup because strong global attention does not automatically translate into full incremental occupancy if regular visitors are displaced.

Formula 1 has a particularly relevant empirical literature because it is an annual, internationally mobile event. Recent work summarized by CEIBS reports evidence that F1 hosting status can raise tourism demand in host countries, while other regional studies find more mixed macroeconomic effects. The mixed evidence is useful: F1 may be strong enough to shift tourism and hotel outcomes without necessarily producing broad regional GDP gains. In the Mexico City context, this paper therefore focuses on hotel occupancy and tourists hosted in hotels rather than attempting to estimate overall GDP impact.

Finally, World Cup tourism studies often emphasize the difference between short-run visitor flows and long-run destination image. The World Tourism Organization's discussion of mega-events notes that mega-events can attract visitors and global attention to a destination. However, the long-run value depends on planning, integration with tourism strategy, and the host city's ability to convert temporary exposure into repeat travel and investment. Mexico City already has high baseline visibility, so the incremental value of the World Cup may be less about putting the city on the map and more about increasing international attention, hotel demand, and visitor spending during specific match windows.

3 Data

3.1 Tourism outcomes

The main dependent variables are monthly hotel occupancy and tourists hosted in hotels. Let t index months. Hotel occupancy is measured as a percentage of available hotel rooms occupied during month t:

$$Y_t^{occ} = \text{Hotel occupancy percentage in month } t. \tag{1}$$

The second dependent variable is total tourists hosted in hotels. Because tourist counts are right-skewed and multiplicative changes are easier to interpret, the regression uses the log transformation:

$$Y_t^{tour} = \log(\text{Tourists hosted in hotels}_t). \tag{2}$$

The accessible empirical sample runs from January 2019 to April 2024, for a total of 64 monthly observations. The sample has an important drawback: it begins after Formula 1 had already returned to Mexico City, meaning it cannot use the 2015 return as a separate treatment shock. It also includes the COVID-19 collapse and recovery period, which makes explicit controls necessary.

3.2 Event indicators

The treatment variables are monthly event dummies. Let $F1_t$ equal one if the Formula 1 Mexico City Grand Prix occurred during month t, and zero otherwise. Let NFL_t equal one if an NFL regular-season Mexico City game occurred during month t, and zero otherwise.

Let ME_t be a combined major-event indicator:

$$F1_t = \begin{cases} 1 & \text{if an F1 Mexico City Grand Prix occurred in month } t, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

$$NFL_t = \begin{cases} 1 & \text{if an NFL Mexico City game occurred in month } t, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

$$ME_t = \mathbb{1}(F1_t = 1 \text{ or } NFL_t = 1). \tag{5}$$

Formula 1 returned to the Mexico City calendar in 2015 after a 23-year absence, and the Mexico City Grand Prix is typically held in late October or early November. NFL regular-season games in Mexico City were hosted in 2016, 2017, 2019, and 2022 during the modern International Series period, with the 2018 planned game moved because of field conditions. In the accessible 2019-April 2024 sample, event months include F1 months in 2019, 2021, 2022, and 2023, and NFL months in 2019 and 2022. The 2024 F1 event is outside the accessible sample because the current series ends in April 2024.

3.3 COVID shock control

COVID-19 is the dominant structural break in the series. Mexico City's hotel occupancy fell dramatically in 2020 and then recovered gradually through 2021 and 2022. To avoid attributing this collapse to event timing or ordinary seasonality, I define a COVID shock indicator:

$$COVID_t = \begin{cases} 1 & \text{if } t \in [\text{April 2020}, \text{December 2021}], \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

This is a deliberately simple control. A more refined version could use mobility data, restrictions, airport passenger data, or interacted pandemic-recovery trends. For a short research note, the dummy is transparent and materially improves interpretability.

4 Descriptive Analysis

Figure 1 plots hotel occupancy over the accessible sample and marks major event months. Event months cluster around visible peaks in 2019, late 2021, late 2022, and late 2023. The post-COVID rebound is also clear: occupancy falls sharply in 2020, reaches a trough during the strict pandemic period, and returns toward pre-pandemic levels by 2022 and 2023.

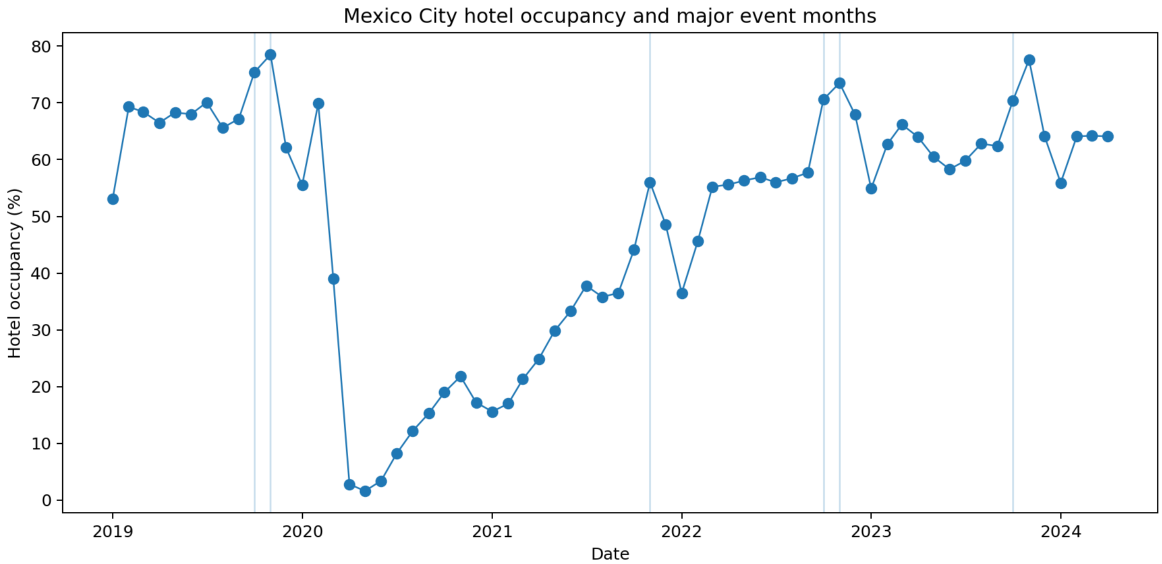

Figure 1: Mexico City hotel occupancy and major event months.

Figure 2 plots total tourists hosted in hotels. The pattern is similar but noisier. Tourist counts collapse in 2020 and recover, but the relationship between event months and tourist counts is less visually stable than for occupancy. One reason is that hotel occupancy reflects both volume and capacity pressure, while tourists hosted in hotels may be affected by stay length, room sharing, domestic/international mix, and reporting conventions.

Figure 2: Tourists hosted in Mexico City hotels and major event months.

Table 1 summarizes event and non-event months. In the raw descriptive comparison, event months have higher average hotel occupancy and higher tourist counts. Major-event months average 70.7 percent occupancy compared with 47.5 percent in non-event months, a raw difference of 23.2 percentage points. The raw difference in tourists is also positive. These simple differences are useful, but they are not causal: events occur in particular months, especially October and November, and those months may already have different tourism patterns.

Table 1: Descriptive comparison: event vs. non-event months

Outcome	Event	Event avg.	Non-event avg.	Difference
Hotel occupancy (pp)	F1	68.1	48.5	19.6
Hotel occupancy (pp)	NFL	76.0	48.8	27.2
Hotel occupancy (pp)	Major event	70.7	47.5	23.2
Tourists hosted in hotels	F1	1,223,578	856,560	367,018
Tourists hosted in hotels	NFL	1,356,645	864,107	492,538
Tourists hosted in hotels	Major event	1,267,934	839,316	428,618

Notes: Calculations use the accessible January 2019-April 2024 monthly sample. Occupancy is measured in percentage points. Tourist counts refer to tourists hosted in hotels.

5 Econometric Methodology

5.1 Research objective and estimand

The purpose of the empirical section is to estimate whether months with major international events are associated with higher tourism activity in Mexico City. The available data are monthly and citywide, so the estimand is not an event-day causal treatment effect. It is a monthly reduced-form association after controlling for ordinary seasonality, annual shocks, and the pandemic period. Formally, the preferred estimand is:

$$\beta = E[Y_t \mid ME_t = 1, X_t] - E[Y_t \mid ME_t = 0, X_t], \tag{7}$$

where Y_t is a tourism-market outcome in month t, ME_t is a major-event-month indicator, and X_t is the set of controls. The coefficient should be read as the average conditional difference between event and non-event months, not as a fully causal effect of hosting an event.

This distinction is important because Formula 1 and NFL games are not randomly assigned across months. They typically occur in October or November, which may differ from the rest of the year even in the absence of the event. To address this, the empirical design uses month fixed effects. The model also controls for year fixed effects to absorb broad annual conditions, including the post-pandemic recovery path. The COVID shock is included separately because the pandemic generated an unusually large discontinuity in both hotel occupancy and tourist arrivals.

5.2 Outcome variables

Two dependent variables are used. The first is hotel occupancy, measured in percentage points:

$$Y_t^{occ} = 100 \times \frac{OccupiedRooms_t}{AvailableRooms_t}. \tag{8}$$

This is the preferred outcome because major events can create short-run pressure on the hotel market even if the number of tourists does not increase one-for-one. Occupancy captures the intensity with which existing hotel capacity is used.

The second outcome is the logarithm of tourists hosted in hotels:

$$Y_t^{tour} = \log(Tourists_t), \tag{9}$$

where $Tourists_t$ is the number of tourists registered in Mexico City hotels during month t. The logarithmic transformation makes the coefficient approximately interpretable as a percentage change. For exact interpretation, I use:

$$\%\Delta Tourists = 100 \times \left(e^{\hat{\beta}} - 1\right). \tag{10}$$

5.3 Event-month variables

The event variables are binary indicators. Let $F1_t$ equal one if the Formula 1 Mexico City Grand Prix occurred in month t, and zero otherwise. Let NFL_t equal one if an NFL Mexico City regular-season game occurred in month t, and zero otherwise. The combined major-event indicator is defined as:

$$ME_t = \mathbb{1}(F1_t + NFL_t > 0). \tag{11}$$

The combined indicator is useful because the sample contains only a small number of event months. Separating F1 and NFL is informative, but the NFL estimate in particular is based on very few treated observations in the accessible sample. The combined specification therefore improves statistical power by pooling comparable international sports-event months.

5.4 Preferred specification

The preferred specification is:

$$Y_t = \alpha + \beta ME_t + \theta COVID_t + \sum_{m=2}^{12} \delta_m M_{m,t} + \sum_{y} \lambda_y Year_{y,t} + \varepsilon_t, \tag{12}$$

where Y_t is either Y_t^{occ} or Y_t^{tour}, $COVID_t$ is a pandemic shock dummy, $M_{m,t}$ are month fixed effects, and $Year_{y,t}$ are year fixed effects. January is the omitted month category. The coefficient of interest is β.

This specification is preferred because it matches the structure of the data and the research question. Month fixed effects compare event months against the usual seasonal pattern for that month. Year fixed effects control for broad changes in the city economy, tourism recovery, and travel conditions. The COVID dummy isolates the extraordinary pandemic collapse that is not well summarized by ordinary year effects alone.

5.5 Separate-event specification

I also estimate a separate-event model:

$$Y_t = \alpha + \beta_1 F1_t + \beta_2 NFL_t + \theta COVID_t + \sum_{m=2}^{12} \delta_m M_{m,t} + \sum_{y} \lambda_y Year_{y,t} + \varepsilon_t. \tag{13}$$

Here β_1 captures the conditional association between F1 months and the outcome, while β_2 captures the same relationship for NFL months. This model is useful for interpretation, but the coefficients should be read with caution because the number of event months is small.

5.6 Uncertainty and HAC standard errors

Monthly tourism outcomes are likely to display serial correlation. A weak month can be followed by another weak month, and the recovery from COVID unfolded gradually over many months. For this reason, ordinary OLS standard errors may understate uncertainty. The coefficients are estimated by OLS, but inference is based on heteroskedasticity and autocorrelation consistent standard errors:

$$\widehat{Var}_{HAC}(\hat{\beta}) = (X'X)^{-1} X'\widehat{\Omega} X (X'X)^{-1}, \tag{14}$$

where $\widehat{\Omega}$ allows the residuals to be heteroskedastic and serially correlated. I use a Newey-West style HAC estimator with a short monthly lag window. Given the sample size, the goal is not to claim precise causal identification, but to avoid overstating statistical significance.

5.7 Robustness checks

The main robustness checks test whether the occupancy result depends on the exact control structure or on a single event month. I estimate four alternative models. First, I replace year fixed effects with a linear time trend. Second, I omit year effects and keep month fixed effects plus the COVID dummy. Third, I restrict the sample to the post-COVID recovery period. Fourth, I drop the COVID-shock months entirely. I also perform a leave-one-event-month-out exercise in which the preferred model is re-estimated after excluding each event month one at a time.

These checks are designed to answer a narrow but important question: is the main positive occupancy result entirely driven by one event month or by the COVID adjustment? If the coefficient remains positive across alternatives, the qualitative finding is more credible. If significance weakens when the pandemic months are excluded, that is also informative because it shows that the short accessible sample limits the strength of the conclusion.

6 Empirical Results

6.1 Main estimates

Table 2 reports the main regression results. The strongest and most interpretable result is for hotel occupancy. In the separate-event specification, Formula 1 months are associated with a 6.68 percentage-point increase in hotel occupancy, statistically significant at the 5 percent level. NFL months are associated with a 7.77 percentage-point increase, but the estimate is not statistically significant at conventional levels. The lack of statistical significance for NFL should not be read as evidence of no effect; it mostly reflects the fact that there are only two NFL event months in the accessible sample.

The combined event model estimates that major-event months are associated with a 7.02 percentage-point increase in hotel occupancy. This estimate is statistically significant at the 5 percent level using HAC standard errors. Economically, the magnitude is meaningful. In the sample, non-event months average 47.5 percent occupancy, while major-event months average 70.7 percent occupancy in the raw data. After controlling for seasonality, year effects, and the COVID shock, the model still estimates an event-month premium of about seven percentage points.

Table 2: Main regression results

Specification	Outcome	Term	Coef.	HAC SE	p-value	N	R^2
Separate events	Hotel occupancy (pp)	F1 event month	6.679	(3.235)	0.039	64	0.890
Separate events	Hotel occupancy (pp)	NFL event month	7.771	(5.156)	0.132	64	0.890
Separate events	Hotel occupancy (pp)	COVID shock	-29.907	(7.777)	< 0.001	64	0.890
Combined event	Hotel occupancy (pp)	Major event month	7.017	(3.514)	0.046	64	0.890
Combined event	Hotel occupancy (pp)	COVID shock	-29.864	(7.748)	< 0.001	64	0.890
Separate events	Log tourists	F1 event month	-0.188	(0.238)	0.430	64	0.730
Separate events	Log tourists	NFL event month	-0.130	(0.264)	0.622	64	0.730
Separate events	Log tourists	COVID shock	-0.922	(0.363)	0.011	64	0.730
Combined event	Log tourists	Major event month	-0.170	(0.233)	0.466	64	0.730
Combined event	Log tourists	COVID shock	-0.920	(0.362)	0.011	64	0.730

Notes: All models include month fixed effects and year fixed effects. HAC standard errors are reported in parentheses. For hotel occupancy, coefficients are percentage points. For log tourists, coefficients are log points. The exact percentage conversion for the combined tourist coefficient is $100(e^{-0.170} - 1) = -15.6\%$, but the estimate is statistically insignificant and should not be interpreted as evidence of a negative effect.

The COVID coefficient is large and negative in all models. In the occupancy model, the pandemic shock is associated with an approximately 30 percentage-point reduction in hotel occupancy, conditional on the fixed effects. This validates the decision to control explicitly for the pandemic period. Without this adjustment, the recovery pattern could be incorrectly attributed to event timing or annual differences.

6.2 Interpretation of the tourist-count result

The tourist-count models are weaker than the occupancy models. The major-event coefficient for log tourists is negative and statistically insignificant after controls. This should not be interpreted as evidence that events reduce tourism. Rather, it shows that aggregate monthly tourist counts do not contain a clean event signal in this short sample.

There are several plausible reasons. First, event visitors may displace ordinary visitors, raising occupancy and prices without increasing total tourist counts proportionally. Second, a monthly tourist count is a coarse measure for an event that may last only a few days. Third, attendees may share hotel rooms, stay with friends or relatives, or use short-term rentals not fully captured in hotel statistics. Fourth, large events may affect room rates and occupancy more strongly than the number of registered hotel tourists. For this reason, hotel occupancy is treated as the preferred outcome in this version of the paper.

6.3 Robustness of the occupancy result

Table 3 reports the robustness checks for the combined major-event indicator using hotel occupancy as the outcome. The coefficient is positive in all specifications. The preferred estimate is 7.02 percentage points. When year fixed effects are replaced by a linear trend, the estimate rises to 8.56 percentage points but becomes statistically weaker. When only month fixed effects and the COVID dummy are included, the estimate is 8.91 percentage points. When the sample is restricted to the post-COVID recovery period or when COVID-shock months are excluded, the coefficient remains positive but falls to roughly 2.4-2.8 percentage points and is not statistically significant.

Table 3: Robustness checks: major-event coefficient on hotel occupancy

Specification	Coef.	HAC SE	p-value	N	R^2
Preferred: month FE + year FE + COVID	7.017	3.514	0.046	64	0.890
Alt. 1: month FE + linear trend + COVID	8.563	5.010	0.087	64	0.823
Alt. 2: month FE + COVID only	8.906	5.087	0.080	64	0.822
Alt. 3: post-COVID recovery only	2.448	1.525	0.108	40	0.937
Alt. 4: excludes COVID-shock months	2.762	1.979	0.163	48	0.903

Notes: The dependent variable is hotel occupancy in percentage points. All models include month controls. The preferred model includes year fixed effects and a COVID-shock dummy. HAC standard errors are reported.

These robustness checks strengthen the qualitative conclusion but also discipline the interpretation. The positive sign is stable, which supports the idea that major events are associated with higher hotel occupancy. However, statistical significance depends on the treatment of the pandemic and the short sample. The most defensible conclusion is therefore not that major events causally raise occupancy by exactly seven percentage points, but that the available data show a positive and economically meaningful event-month association in hotel occupancy.

The leave-one-event-month-out exercise further supports this interpretation. Re-estimating the preferred model while excluding each event month one at a time produces event coefficients between 5.69 and 8.82 percentage points. This implies that the main result is not mechanically driven by a single event month. The p-values vary around conventional thresholds, again reflecting the small number of treated observations.

Table 4: Leave-one-event-month-out estimates

Excluded event month	Coef.	HAC SE	p-value
2019-10	6.949	3.685	0.059
2019-11	7.513	3.370	0.026
2021-11	8.821	3.948	0.025
2022-10	5.691	3.071	0.064
2022-11	6.300	3.444	0.067
2023-10	7.379	3.608	0.041

Notes: Each row re-estimates the preferred combined-event model after removing the listed event month. The dependent variable is hotel occupancy in percentage points.

6.4 Result summary

The best reading of the evidence is threefold. First, the hotel-market signal is positive: major-event months are consistently associated with higher hotel occupancy. Second, the statistical evidence is suggestive rather than definitive because the accessible sample is short and includes the pandemic shock. Third, the tourist-count evidence is inconclusive, which implies that future versions should add higher-frequency data and price variables such as average daily rate or revenue per available room. Those variables would better capture whether event months increase net economic value or simply reallocate existing tourism demand.

7 Discussion: What This Means for the World Cup

The results suggest that major events can move Mexico City's hotel market, but in a specific way. They show up more clearly in hotel occupancy than in aggregate tourist counts. For the World Cup, this distinction matters. A World Cup match day can raise room demand, prices, and venue-area activity without necessarily increasing monthly tourist counts one-for-one if regular visitors are displaced or if fans concentrate their stays around short windows.

Using the combined-event coefficient as a rough benchmark, the model implies that a major-event month is associated with approximately seven additional percentage points of hotel occupancy. This should not be mechanically applied to the World Cup, for at least four reasons. First, World Cup matches differ from F1 and NFL games because they involve multiple matches, international fan bases, and global media attention. Second, the World Cup's timing in June and July differs from the October-November timing of F1 and NFL events. Third, World Cup pricing and hotel booking behavior may generate displacement of ordinary tourists. Fourth, the host-city effect may depend on Mexico's performance, visiting-team fan bases, public safety, and transportation conditions.

A conservative pre-tournament interpretation is that Mexico City should expect clear hotel-market pressure during World Cup match windows, especially around the opening match. However, the broader economic effect will depend on whether the city increases net visitor nights or merely reallocates existing tourism demand. The key policy and business question is therefore not just whether hotels fill up, but whether the event extends stays, raises spending outside stadium zones, and converts global attention into repeat tourism.

8 Limitations

This paper has four main limitations. First, the accessible sample is short. It includes 64 monthly observations from January 2019 to April 2024. This limits statistical power and makes the estimates sensitive to the pandemic period. Second, the analysis uses monthly data even though event effects are likely concentrated in a few days around each event. Daily hotel data would be substantially better. Third, the specification is associational. It controls for seasonality and year effects, but it does not fully solve selection bias or displacement. Fourth, the estimates are citywide, meaning they cannot distinguish hotel-market effects near the venue from effects in other parts of the city.

A longer 2015-2024 sample would be valuable because it would add more pre-pandemic observations, include the 2015 return of Formula 1 to Mexico City, and increase the number of event months used for estimation. The current results should therefore be interpreted as preliminary evidence based on the accessible monthly sample.

9 Conclusion

This working paper asks whether major events can move Mexico City's economy, using hotel occupancy and tourist arrivals as the measurable first layer of economic activity. The answer is cautiously yes for hotel occupancy. In the accessible 2019-April 2024 sample, Formula 1 months and combined major-event months are associated with higher hotel occupancy after controlling for seasonality, year effects, and the COVID shock. The magnitude is economically meaningful: roughly seven percentage points in the combined specification. The evidence is weaker for total tourists hosted in hotels, suggesting that aggregate tourist counts may be too coarse to capture short-run event pressure.

For the 2026 World Cup, this June 2025 pre-tournament analysis does not imply that the tournament will automatically transform the city economy. Rather, the evidence suggests that major events can create measurable pressure in hotel markets, especially when they attract non-local visitors and global attention. The net economic effect will depend on how many visitors are incremental, how long they stay, whether ordinary tourists are displaced, and whether spending spreads beyond stadium-adjacent areas.

The most practical conclusion is that the World Cup should be evaluated with humility and data. Ex ante forecasts should be treated as scenarios, not certainties. Ex post analysis should compare hotel occupancy, tourists, room revenue, local spending, and mobility around match windows against credible counterfactuals. This paper provides a first version of that framework for Mexico City and a structure that can be updated as richer data becomes available.

References

CDMX Open Data.
Ocupacion hotelera en la Ciudad de Mexico. Monthly hotel occupancy percentage dataset, CSV resource updated September 2024.

CDMX Open Data.
Turistas en la Ciudad de Mexico. Monthly domestic and international tourists hosted in hotels in Mexico City, CSV resource updated September 2024.

DataTur / SECTUR.
Ocupacion hotelera en los 70 destinos principales monitoreados en DataTur. CSV resource for hotel occupancy in tourist-class lodging establishments, updated May 2025 metadata.

Depken, Craig A., and E. Frank Stephenson (2018).
Hotel Demand Before, During, and After Sports Events: Evidence from Charlotte, North Carolina. *Economic Inquiry.*

Humphreys, Brad R., and Li Zhou (2019).
Sports-led tourism, spatial displacement, and hotel demand. *Economic Inquiry.*

Matheson, Victor A. (2006).
Mega-Events: The effect of the world's biggest sporting events on local, regional, and national economies. College of the Holy Cross working paper.

Sun, Y. Y. (2012).
Why hotel rooms were not full during a hallmark sporting event. *Tourism Management.*

Formula 1.
Mexico City Grand Prix circuit and timetable information.

NFL.
NFL International Series and Mexico City game history.

FIFA.
FIFA World Cup 2026 Mexico City host-city and hospitality information.

A Event-month coding

The event-month approach uses the month in which the main race or game occurred. This is appropriate for monthly data, but it introduces measurement error because event demand is likely concentrated over a few days. A daily dataset would allow the preferred event-study form:

$$Y_d = \alpha + \sum_{k=-K}^{K} \beta_k \mathbb{1}(d - d^* = k) + \gamma X_d + \delta_{dow} + \lambda_y + \varepsilon_d, \tag{15}$$

where d indexes days, d^* is the event date, and β_k traces pre-event, event-day, and post-event hotel-market effects.

B Future version

The next version of this paper should add three improvements. First, it should use daily hotel data or at least weekly hotel data. Second, it should estimate a difference-in-differences model comparing Mexico City with other large Mexican tourism destinations not hosting the event in the same period. Third, it should add outcome variables for average daily rate, revenue per available room, or tourist spending, because occupancy alone captures volume pressure but not price effects.

$$Y_{i,t} = \alpha_i + \lambda_t + \beta(Host_i \times Event_t) + \gamma X_{i,t} + \varepsilon_{i,t}, \tag{16}$$

where i indexes cities or destinations, $Host_i$ identifies Mexico City, and $Event_t$ identifies event windows. This design would improve causal interpretation by comparing Mexico City to destinations exposed to the same national shocks but not directly hosting the event.