# Is There a Consensus? An Experimental Trial to Test the Sufficiency of Methodologies Used to Measure Economic Impact

**Daniel A. Rascher**
**University of San Francisco**

**Giseob Hyun**
**Facebook**

**Mark S. Nagel**
**University of South Carolina**

*This research utilizes local GDP of 383 MSAs in the U.S. to determine whether historical methods in the academic literature to measure the economic impact of sports are sensitive enough to generate conclusive results. An experiment is created and shows that commonly used methods fail to be able to detect the built-in-by-design injections of economic activity for the experimental group until very high levels of treatment of at least $300 million to $1 billion annually are present, thus providing evidence that Type I errors (rejecting a true null hypothesis) are likely to have occurred in some of the literature.*

*Keywords: economic impact, regression, type I error, designed experiment*

## INTRODUCTION

> *There are a lot of things economists disagree about, but the economic impact of sports stadiums isn't one of them. 'If you ever had a consensus in economics, this would be it,' says Michael Leeds, a sports economist at Temple University. 'There is no impact.'* (Bergman, 2015, paras 1-2)

As the above quote states, nearly all economists who have studied whether a facility, team, or major sporting event has a positive economic impact on a city or region have concluded that either there is no impact or that in a few cases there is a small positive impact (or a small negative impact). Economic impact is typically operationalized by measuring whether there is a net new change in the local economy (e.g., local Gross Domestic Product [GDP], employment, number of firms, tax revenues collected, overnight stays at hotels) because of a new facility, team, or sporting event. The academic literature supporting this consensus historically utilized regression techniques on panel data consisting of cities (or Metropolitan Statistical Areas [MSAs]) over a number of years with the use of indicator or dummy variables representing whether a market has a team, stadium, or hosted a major sporting event. More recent research has narrowed the time periods (to quarters, months, or days [when using hotel occupancy or tourist arrivals]) and/or the geographic region (from MSAs and counties to census areas and zip codes) to enable more fine-grained analysis, as some authors have noted that the annual analysis on county-level data, for instance, could be too insensitive to detect what is attempting to be detected (e.g., Matheson and Baade, 2006; Chikish, Humphreys, Liu, and Nowak, 2019).[1]

These regressions either use the level of GDP regressed on the dummy variables and control variables, or lagged dependent variables, or the growth in GDP as the dependent variable on the growth in the control variables along with the dummy variables of interest. The results often show the dummy variables as not being statistically significantly different from zero at conventional significance levels. Then, the authors tend to conclude that the existence of the team, stadium, or event has no net positive effect on the local economy. Occasionally, the authors note that the size of local economies may be too large relative to the events/teams/stadiums being studied to be detectable. Recent studies have specifically noted this problem and have focused on shorter time periods and smaller geographic areas (Baade, Baumann, and Matheson, 2008a; Baade, Baumann, and Matheson, 2008b; Baade, Baumann, and Matheson, 2011; Baumann and Matheson, 2017; Baumann and Matheson, 2018).

However, industry practitioners often utilize intercept surveys at sporting events, as opposed to the *ex post* studies from the academic literature, to directly measure whether there is net new incremental money being spent in the municipality due to a new stadium, team, or event. If attendees to a sporting event come from out of town, spend money at local hotels and restaurants, and would not have spent that money in town had the event not taken place (i.e., they are in town because of the event and it does not take the place of another trip to that same town at some other point in time), then that expenditure is deemed to be an incremental gain to the local economy. This is generally the methodology used by industry practitioners. These practitioner reports often show quite large impacts, e.g., more than $100 million per year because of a new team or stadium, and tens of millions or hundreds of millions of dollars from major sporting events, like the Super Bowl. These can be ten times (or more) larger than what the academic sports economists claim is likely the true impact (Kanell, 2019). For instance, economist Victor Matheson has concluded that one should "Take whatever number the sports promoter says, take it and move the decimal one place to the left. Divide it by ten, and that's a pretty good estimate of the actual economic impact" (Garofalo & Waldron, 2012, para. 9).

What explains the difference in the findings? The academic economists tend to claim that the industry reports either use the incorrect methodology (Crompton, 1995), and/or there are incentives to generate large impacts because the reports are often commissioned by teams, events, politicians and others who have a vested interest in showing large impacts (Jeanrenaud, 2007). Many of these studies do not control for possible negative effects of decreases in typical tourism, spending by locals, traffic congestion, etc., caused by the event/team/stadium, thus one might consider these to be measures of gross economic impact, not net economic impact (Agha & Taks, 2018).

While errors and embellishments in economic impact studies may certainly occur, an alternative to the results generated by academic economists is that their methods employed may not be sensitive enough to be able to measure the true underlying economic impact. Most of the early academic studies address the impact of teams in major sports leagues (e.g., Major League Baseball [MLB], National Basketball Association [NBA], National Football League [NFL], National Hockey League [NHL]) or major sporting events such as the Super Bowl or MLB All-Star Game on the host cities. In nearly all of those cases, the host cities are the largest such metropolitan areas in the United States, with annual GDP exceeding $250 billion.[2] Thus, if a team were to *truly* bring net new economic impact of $100 million annually to one of these markets, it would only be an increase of 1/2500 in GDP. This might not be detectable using annual panel data regressions. The average change in GDP from 2016 to 2017 in the top 50 U.S. markets (defined by size of GDP) exceeds $10 billion. Thus, even the *change* in GDP in a typical market is more than one hundred times the $100 million example, and thus might not be easily detected in panel regressions given the large variance in local GDP over time and across markets.

This research article introduces a novel experiment in order to test whether the historical methodologies used in the academic literature are powerful enough or sensitive enough to measure true impacts of varying sizes. The experiment involves taking the existing data on GDP across markets and over time as given and then injecting various amounts of economic impact into randomly selected markets. Then, panel regressions are estimated to see if those economic impact injections can be detected. In other words, a randomly chosen ten percent of the markets will have millions of dollars added to their annual GDP for one year. The year is randomly chosen for each market (using a six-year panel). Thus, there is not a question about whether or not there *is* a true economic stimulus coming from hosting a major event because that stimulus has been forcefully added to that market. If the resulting panel regressions (with dependent variables using levels of GDP both with and without lagged dependent variables on the right-hand side of the equation, and also using growth in GDP) show no statistically significant impact, then the methods are not suitable for measuring these sizes of impacts. Variations on how many markets get the stimulus (10% up to 50%), the dollar size of the stimulus (ranging from $25 million to $1 billion per year), separate analyses for large, medium, and small markets, and stimuli designed to represent a team, are analyzed in order to be sure the findings are robust. The remainder of this article discusses the relevant literature, methods, data, and results.

## LITERATURE

The literature concerning the economic impact of sports teams, facilities, and major sporting events is extensive and has been ongoing for decades. Based on reviews of individual stadia, their financing, and revenues generated, Okner (1974) concluded that most stadia produce a loss for their community. Baim (1990) analyzed 14 stadia and compared tax dollars spent and collected, concluding that the vast majority of the facilities produced financial losses. Baade (1987), in what appears to be the first of these studies to utilize regression techniques, studied nine markets over 19 years to analyze economic impact with a variety of dependent variables. Depending on the specification, there were positive, insignificant, or negative impacts from stadium construction or the introduction of a new team, but the general conclusion was that there were insignificant or negative economic impacts (once other controls were introduced).

Baade and Dye (1990) used the level of per capita personal income as the dependent variable in regressions on population, team and stadium dummy variables for nine individual markets over 19 years. They also pooled the data across markets (panel data), and tested personal income relative to the region's income as the dependent variable. They generally found insignificant or negative impacts using what appears to be the same general data as used in Baade (1987).

Baade (1994) created a study with a larger dataset (48 MSAs, 36 with a major professional team and 12 without) over the years 1958-1987. The author concluded that there is no evidence of a positive economic impact from new teams or stadia. In a separate model, controlling for regional economic growth, the results showed zero or negative economic impacts.

Baade (1996) employed what is now often termed a difference-in-differences regression in his study. Essentially, the analysis investigated the change in annual real per-capita income for cities experiencing a variation in the number of new stadia or teams compared to income changes in cities that did not have team or facility alterations. That dependent variable was regressed on the number of teams and stadia within the various markets over time using 48 markets over 30 years. The research noted that by not including additional control variables, the likelihood of getting statistical significance on the variables of interest (change in number of teams or stadia) is higher. In other words, if no significance is found without additional control variables, then it is not likely to be found when those additional variables are included. Individual city regressions tended to find no impact, as did the pooled regressions.

Porter (1999), Baade and Matheson (2001), and Matheson and Baade (2003) began the literature track analyzing major events, as opposed to teams or stadia, and their potential impacts. Porter (1999) used regression analysis of monthly sales tax data in Florida and Arizona for the counties that hosted Super Bowls during 1978-1996. Porter concluded there was no discernible impact. Baade and Matheson (2001) evaluated MLB's All-Star Game from 1973-1997 using the percent change in employment for an MSA regressed on the lagged dependent variable, other cities' change in employment, population, real per capita income, nominal wages, state and local taxes, a number of other dummy variables including an MSA dummy, annual trend, and finally an All-Star Game dummy variable. The sample consisted of 57 cities based on population. Additionally, the authors evaluated quarterly sales tax data from 1986-2000 in

California (in proportion to the sales tax collected in the rest of the state in which the city resides) using a panel dataset, and concluded that there is no economic impact.

The use of panel regression methodologies has indicated a consistent pattern of "no positive economic impact" across a number of studies (Coates, 2007; Coates & Depken, 2009; Coates and Humphreys, 1999, 2002, 2003). These studies, and a number of others, have led to the general consensus that sport teams, venues and events do not provide positive economic impact to a community. However, even when designing studies in the "traditional" manner, some events have been found to provide positive economic impact. Matheson and Baade (2003) analyzed whether the National Collegiate Athletic Association (NCAA) Women's and Men's Basketball Tournament's Final Four's impacted the local economy, finding no impact across 30 years of the Men's Final Four and 18 years of the Women' Final Four's except in the smallest city that hosted, College Park, MD. Agha (2013) also noted the positive change in local per capita income among a variety of minor league baseball teams and facilities. Santo (2005) found that in eight of nineteen situations, stadiums in downtown settings, especially for new teams (not replacement stadiums) were positively correlated with regional income share for the years 1984-2001.

In addition, though Coates and Depken (2011) examined monthly sales tax data for 23 cities in Texas from 1990-2008 and found that regular season professional games (NBA, MLB, NHL, NFL) had no discernible impact, they did note the Super Bowl and college football games do have a statistically significant positive impact. For the Super Bowl, the mega-event occurs not just on one day but typically offers an entire week of festivities. However, for college football games, the positive impact from these one-day events may be due to the games being played in a number of cities with populations less than 500,000 (College Station, Lubbock, Waco, etc.). In another study of college football in smaller markets (Gainesville and Tallahassee, FL), the authors concluded that each home college football game generated an additional $2 million in taxable sales. These teams tend to have 7 home games per season, thus implying about $14 million in increased taxable sales per year.

Agha and Rascher (2016) extensively discussed the potential for smaller cities to register a positive economic impact when utilizing traditional regression-based methodologies because an individual event is a larger portion of a smaller local economy. Interestingly, and somewhat ironically, some of the "consensus" "no-positive-economic-benefit" research does mention a key driver of this study - that the size of the economies studied in the bulk of the literature are so large as to possibly dwarf any true impact that may actually exist (thus, suffering a Type I error) and be detectable with long-utilized methods. For instance, Baade noted "Some analysts might argue that this model demands too much from professional sports. Since the professional sports industry is small relative to a large city's economy, it is unlikely that professional sports plays a statistically significant role in determining real per capita income" (1994, p. 12). In addition, Baade and Matheson stated:

> Because the level of employment in the cities hosting the All-Star Game ranged from about 700,000 workers to just more than 4,000,000 workers, a gain of 1,000 workers will be a small impact in relation to the whole citywide economy, representing a percentage gain of between 0.14% and 0.023% for the cities hosting games between 1973 and 1997…Although a gain of 1,000 jobs is still well within a 90% confidence interval for the predicted employment gains, the evidence clearly points against an economic impact that is "off the charts," as claimed by American League President Gene Budig. (2001, pp. 315-316)

Also, Baade and Matheson intimated some solutions, noting that "If the researcher can compress the time period, then it is less likely that the impact of the event will be obscured by the large, diverse economy within which it took place" (2001, p. 320). More recently, Baade, Baumann, and Matheson (2011, p. 374), utilizing monthly data, note a "valid criticism, however, of the existing body of work regarding the ex-post economic impact of sports is that these studies are trying to uncover the proverbial needle in a haystack." Using monthly data instead of annual data helps shrink the haystack, but it still potentially has the same problem of an event being relatively small compared to the local economy, as the authors suggest.

In other words, attempts to shrink the geographic region of study or the time period (for events, but not likely for stadia which tend to be operational over much of the year) can help make the analysis more likely to detect an impact when one actually exists. More recently, Matheson (2018, p. 275) made the case to create more sensitive tests, stating:

> For example, Baade and Matheson's (2006) examination of the Super Bowl found that its impact on the host economies was not statistically different from zero. However, the authors also noted that given the sensitivity of their model, the Super Bowl would have to generate at least $300 million in benefits before they would pick it up as statistically significant. Any impact level below that, no matter how real the benefits were, could not be differentiated from the statistical noise.

Given the potential that positive economic impacts may exist and yet be undetected with prior methods (a Type I error), the present study is the first to directly measure the threshold size of the economic impact of an event/team/stadium that is detectable using standard methods, and acts as a guideline for future researchers. In other words, how large must the actual economic impact be in order to be detected?

## METHODS

### Events

The methodology used in the present article is distinct from all of the previous literature concerning the economic impact of sports in that instead of testing whether an impact exists for facilities, teams, or actual sporting events, economic impact (or GDP of various amounts) is *randomly added* to actual GDP for a number of MSAs and then the typical regressions used in the literature are estimated to see if that impact is detectable. For instance, in one of the iterations approximately 10% of the markets are randomly chosen and $50 million is added to annual GDP in each of those markets. Treatments ranged from $25 million to $1 billion. The following models are then tested to see if they can detect this added economic activity. Equation (1) shows the traditional levels model of local GDP on year indicator variables and a dummy variable indicating a major sporting event. Let

$$GDPEXP_{it} = \alpha_i + \beta_1 Y2012 + \cdots + \beta_6 Y2017 + \beta_7 EVENT_{it} + \varepsilon_{it}, \tag{1}$$

where $GDPEXP_{it}$ (GDP experiment) is the actual GDP in MSA $i$ during year $t$ enhanced with an additional $50 million for the randomly chosen 10% of the markets, *Y2012* through *Y2017* are dummy variables indicating the year of the GDP value, and $EVENT_{it}$ represents the dummy variable for an event in the same MSA $i$ during year $t$ that would have caused the enhancement in $GDPEXP_{it}$.

In order to establish a baseline coefficient on the event dummy variable and thus a baseline impact, actual GDP ($GDP_{it}$) is first regressed on the right-hand side variables using panel fixed effects. It is expected that this coefficient is zero because the dummy variable of interest ($EVENT_{it}$) has been randomly chosen and no treatment has been added to the dependent variable. Then, the model described by Equation (1) is analyzed via a panel regression and the coefficient on $EVENT_{it}$ is compared to that of the baseline regression using a significance test. If there is no statistically significant difference between the two coefficients, then the model is not detecting the actual economic impact that has been built into the data.

As described above by Baade (1996), the lack of other explanatory variables tends to increase the likelihood that $EVENT_{it}$ would show a coefficient that is statistically significantly different from zero. In other words, if this model does not pick up the built-in economic impact, then a model that includes other explanatory variables is not likely to either.

Other models that have been tested (as described in the literature review) analyze the change in local GDP (i.e., $\Delta GDPEXP_{it} = GDPEXP_{it} - GDPEXP_{it-1}$) from one year to the next. The resulting model is shown in Equation (2). Let

$$\Delta GDPEXP_{it} = \alpha_i + \beta_1 Y2012 + \cdots + \beta_6 Y2017 + \beta_7 \Delta EVENT_{it} + \Delta\varepsilon_{it}, \quad (2)$$

where $\Delta EVENT_{it} = EVENT_{it}$ because if one event occurs in MSA $i$ during time $t$, but not during time $t\text{-}1$, then the result is simply $EVENT_{it}$. The analysis randomly chooses which MSAs receive an event and in which year.

Similarly, this model is tested first without the enhanced GDP to establish a baseline relationship between it and the randomly assigned $EVENT_{it}$. Then, the enhanced GDP ($\Delta GDPEXP_{it}$) model is estimated and a comparison of the two coefficients is made using a standard significance test.

These models are tested for all 383 MSAs as a whole, and also separately for the top 50 markets by GDP (large markets), the next 100 markets (medium-sized markets), and the remaining 233 markets (small markets).[3] Additionally, 30% and 50% of the markets are randomly chosen to host an event (not just 10% of the markets as described above). Moreover, larger treatments are administered: \$100 million, \$300 million, \$500 million, and \$1 billion. Finally, the data is re-created 20 times using different randomly selected markets and years. The results reported come from one of the sets of outcomes with any abnormalities described therein.

**Teams or Stadia**

The methodology above assessed markets that hosted a major sporting event either zero or one time during the six years of data. Much of the literature, however, is focused on sports facilities because often hundreds of millions of dollars of public money is used to pay for the construction of these facilities and industry stakeholders report large positive economic impacts that will ensue upon their completion. It is typically less controversial for an existing facility to host a sporting event (when presumably the costs to the public are much lower) than for a new venue to be built with public money. Of course, for cities or countries that host the Olympics or FIFA World Cup, respectively, significant construction projects are the norm. Yet, for an existing stadium to host an MLB All-Star Game, the public investment is much smaller.[4]

The only difference in analyzing events compared with facilities (or teams) is that new venues are built and then continue to be operational in the ensuing years. Thus, the dummy variable, $STADIA_{it}$ takes the value one for the opening year of a new facility and for all years beyond that, not just for one year as in the events analysis. Thus, the standard model is:

$$GDPEXP_{it} = \alpha_i + \beta_1 Y2012 + \cdots + \beta_6 Y2017 + \beta_7 STADIA_{it} + \varepsilon_{it}. \quad (3)$$

The next model, analyzing how the change in GDP is affected by a new facility (or team), reduces to one in which the $STADIA_{it-1}$ is subtracted from $STADIA_{it}$ resulting in a dummy variable that resembles $EVENT_{it}$, yet the dependent variable still represents the treatment (e.g., \$50 million) for each year the stadium exists. In other words, differencing Equation (3) from its lagged self generates the key variable of interest, $\Delta STADIA_{it}$, which only takes the value one during the first year of opening of the stadium and is thus represented by $EVENT_{it}$. Thus, the model is:

$$\Delta GDPEXP_{it} = \alpha_i + \beta_1 Y2012 + \cdots + \beta_6 Y2017 + \beta_7 \Delta EVENT_{it} + \Delta\varepsilon_{it}. \quad (4)$$

As noted in the literature review, there are some more recent studies that have utilized quarterly or monthly data (often sales taxes collected as an indicator of spending), or daily hotel occupancy information, to measure economic impact. These shorter periodicities can be useful for investigating the impact of events, but less so for measuring the impact of facilities, which tend to be available for most of the year, or teams, which operate across multiple quarters and maintain yearlong business operations in the community. Thus, notwithstanding the same general principle that a given event may be too small relative to the size (and variance) of a local economy, the present analysis may leave open the possibility that the quarterly or monthly studies focused on short-term events are sufficiently sensitive enough to be able to detect the true underlying economic impact of most relevant events. Additionally, a facility hosts many individual events (e.g., tenant franchise games or concerts), and thus a study of daily hotel occupancy rates (e.g., Baumann and Matheson, 2017; Chikish, Humphreys, Liu, and Nowak, 2019) can be part of the process to analyze a facility by analyzing individual events.

## DATA

The data come from the U.S. Bureau of Economic Analysis and provide annual GDP for each of 383 metropolitan areas within the U.S. from 2012-2017. As described above, yearly dummy variables are added as well as the key variable of interest, either $EVENT_{it}$ or $STADIA_{it}$, which are randomly assigned to the metropolitan areas.

Even the smallest market of the 383 measured by GDP in 2013, Grants Pass, Oregon, generates nearly $2 billion per year and has a population of about 37,000. Even if it hosted an event that brought in $2 million in new spending by visitors, that would still be about 1,000 times smaller than the local economy, thus making it hard to detect using a time series, panel data, or other statistical tests.

The largest market, New York-Newark-Jersey City, generated over $1.7 trillion in GDP in 2017, and that was $55 billion more than the year before. It is likely that any event or perhaps even any business (e.g., Amazon's HQ2) that operates in that area might not show up in a statistical analysis simply because the market is so large compared to any one event or business.

## RESULTS

### Events

The number of fixed effects panel regressions testing whether the economic impact of events of various sizes can be detected is 432 based on:

- the percentage of random markets receiving the treatment (10%, 30%, and 50%),
- the size of the treatment ($25 million, $50 million, $100 million, $300 million, $500 million, and $1 billion),
- three model specifications (level of GDP, level of GDP including lagged GDP as an explanatory variable, and change in GDP),
- testing the full panel of MSAs (but also separately for large, medium, and small markets), and
- both the baseline models (utilizing actual GDP) and experimental models (utilizing the experimental GDPs with the additional economic impact treatment).

Then, a *z-score* test is conducted that compares the coefficient from the baseline model to the experimental model for each result (thus, 216 *z-score* tests). Therefore, a representative set of results will be shown through this test. A complete review of each model specification was completed with a summary of the circumstances that led to statistically significant effects displayed in Table 3 in the Conclusion. In general, the vast majority of the tests reject the null hypothesis that a true economic impact can be detected using the models and data that are commonly utilized in the literature. This is a classic Type I error in that a true underlying economic impact cannot be detected and thus is generally assumed to not exist even though it does exist, in this case by construction of the experiment.

Table 1 shows the results of applying a $50 million one-year treatment (a major event) to a random sample of 30% of MSAs, with the year chosen randomly within the six-year panel. As expected, the baseline models in the control groups (first, third, and fifth columns) show no significance from the *EVENT* variable. The control variables act as expected, with the effects of each year decreasing compared to the omitted year (2017). The effect of a $50 million event on the change in GDP shows a jump in the coefficient of about $60 million (consistent with expectations), but it is highly insignificant.

**TABLE 1**
**REGRESSIONS RESULTS FOR 30% OF MARKETS GETTING A $50 MILLION EVENT TREATMENT**

| | GDP Level | GDP Level with Treatment | GDP with Lagged | GDP Treatment with Lagged | Change in GDP | Change in GDP Treatment |
|---|---|---|---|---|---|---|
| Event Indicator | 892 | 942 | 0.535 | 59.976 | 1.87 | 61.31 |
| Lagged dependent variable | -- | -- | 1.001 *** | 1.001 *** | -- | -- |
| 2012 | -7983 *** | -7983 *** | -- | -- | -- | -- |
| 2013 | -6718 *** | -6718 *** | -537 *** | -536 *** | -546 *** | -545 *** |
| 2014 | -5010 *** | -5010 *** | -70 | -69 | -77 | -76 |
| 2015 | -3089 *** | -3089 *** | 108 | 108 | 104 | 103 |
| 2016 | -1794 *** | -1794 *** | -500 *** | -499 *** | -501 *** | -501 *** |
| Intercept | 45763 *** | 45763 *** | 1739 *** | 1735 *** | 1803 *** | 1800 *** |
| | ---------- MSA fixed effects not shown ---------- | | | | | |
| *F* | 32.66 | 32.68 | 4311 | 4311 | 6.36 | 6.38 |
| Number of observations | 2298 | 2298 | 1915 | 1915 | 1915 | 1915 |

Statistical significance: *** - 1% level, ** - 5% level, * - 10% level. 2017 is the omitted year.

Not until the treatment reaches $500 million (with 30% of markets receiving it), does *EVENT* show up as being statistically significant, with a coefficient of $595 million (and p-value of 0.013) for the change in GDP model. It is not significant in the levels model. For the specification with 50% of markets receiving a treatment, statistical significance occurs at $300 million. Restricting the analysis to small markets (those outside of the top 150 in 2017 GDP), with 50% of those markets getting a $50 million treatment, the *EVENT* coefficient is statistically significantly different from zero and from the counterpart in the baseline regression.

To show the importance of the size of the markets and possible impacts, a $100 million Event in a large market (top 50) is not statistically significantly different from zero or the baseline coefficient. This is also true for a $300 and $500 million treatment. Not until the event generates $1 billion, does the model show statistical significance for top 50 markets.

### Stadium or Team

Other than the controversy surrounding the Olympics, FIFA World Cup, and other major international mega-events, where significant public money is typically spent on construction projects related to the events, most of the criticism of sports industry economic impact analyses relates to public money being used for new stadia and arenas for a local team. This analysis also includes 432 separate regressions in order to cover a range of possible types of economic impacts.

Table 2 shows the results of providing a $500 million injection to local GDP for 30% of the 383 MSAs with the time period of the facility opening randomly assigned over the six-year period. As expected, *STADIA* is not statistically significant from zero in the baseline model because it is randomly chosen so should generally have no correlation to the dependent variable. Once $500 million is added to the treatment group of MSAs, the coefficient on *STADIA* rises by $500 million, as expected, but it is not statistically significant.

For the model utilizing the change in GDP, while the application of the treatment is seen in the coefficient rising from -$81 million for the baseline to $212 million for the experimental regression, both coefficients are statistically significantly different from zero and from each other.

**TABLE 2**
**REGRESSIONS RESULTS FOR 30% OF MARKETS GETTING A $500 MILLION FACILITY TREATMENT**

| | GDP Level | GDP Level with Treatment | GDP with Lagged | GDP Treatment with Lagged | Change in GDP | Change in GDP Treatment |
|---|---|---|---|---|---|---|
| Stadium Indicator | -1035 | -535 | -80 | 213 | -81 | 212 |
| Lagged dependent variable | -- | -- | 1.001 *** | 1.002 *** | -- | -- |
| 2012 | -8251 *** | -8251 *** | -- | -- | -- | -- |
| 2013 | -6955 *** | -6955 *** | -554 *** | -502 *** | -563 *** | -513 *** |
| 2014 | -5169 *** | -5169 *** | -82 | -37 | -89 | -46 |
| 2015 | -3219 *** | -3219 *** | 99 | 124 | 95 | 118 |
| 2016 | -1866 *** | -1866 *** | -504 *** | -492 *** | -506 *** | -494 *** |
| Intercept | 46,117 *** | 46,117 *** | 1764 *** | 1694 *** | 1827 *** | 1772.128 *** |
| | ---------- MSA fixed effects not shown ---------- | | | | | |
| *F* | 32.71 | 33.57 | 4311 | 4311 | 6.39 | 6.58 |
| Number of observations | 2298 | 2298 | 1915 | 1915 | 1915 | 1915 |

Statistical significance: *** - 1% level, ** - 5% level, * - 10% level. 2017 is the omitted year.

It is not until the treatment is $1 billion, meaning that the new stadium generates an annual economic impact of $1 billion, that the *change in GDP* specification (not the levels model) becomes statistically significant with a coefficient of $504 million and a p-value of 0.046. When 50% of the markets are in the experimental group, a $300 million treatment shows an economic impact of $373 million, significant at the 10% level. When restricting the analysis to medium-sized cities with 10% of the markets receiving a new stadium (or team) at some point during the six-year period, an annual economic impact of $100 million results in a *z-score* (comparing to the baseline coefficient) of 2.10, which passes the one-sided test.

**Conclusion & Discussion**

Over 800 different regression specifications were analyzed in order to determine if an economic impact injection into a local economy representing a major event, team, or facility was actually detectable using standard estimation methods. Models whose dependent variable was the *change in local GDP* from one year to the next unsurprisingly showed statistically significant results at lower levels of treatment. However, the thresholds (shown in Table 3) whereby a treatment becomes large enough to show statistical significance tend to be higher than what many of the industry reports claim is the economic impact of an event, team, or new facility.

The most generous reading of Table 3 shows that if 50% of the markets host an event that is expected to generate $300 million to a market during a given year, then the methods used in the literature ought to be able to detect it. As an example, if a researcher were to analyze 50 cities over a 15-year period to see if hosting a Super Bowl (once per year) would drive a positive economic impact, the table shows that the event would have to actually generate around $500 million in order to result in a statistically significant coefficient.

There are vast differences between the largest and smallest MSAs in terms of GDP, population, etc. A more refined analysis that allows for separate models for each of the three sizes of MSA (i.e., Size 1 is top 50, Size 2 is next 100, and Size 3 is the remaining 233 markets) shows that in smaller markets the threshold size of an event for it to be detectable using the standard techniques in the literature is about $50 million. As shown in Table 4, for medium-sized markets, it is $500 million, and for large markets, it is $1 billion. Much of the previous, academic research has been focused on large markets hosting major events and has found that these events tend to have no discernible economic impact on those markets. Yet, according to the present study's results, those major events would have to generate $1 billion in net new spending in the community in order to generate positive and statistically significant results. Even the Super Bowl is generally not claimed by industry stakeholders to generate that size of an effect, meaning that many of the previous academic studies may have been insufficiently sensitive to be able to detect any actual economic impact (if there is one).

**TABLE 3**
**P-VALUES SHOWING HOW LARGE THE ECONOMIC IMPACT OF AN EVENT MUST BE IN ORDER TO BE DETECTED USING HISTORICAL ECONOMETRIC METHODS**

*10% of cities*

| | GDP | GDP with Lag | First Difference GDP |
|---|---|---|---|
| No Adjustment | 0.814 | 0.435 | 0.437 |
| Additional 50M | 0.835 | 0.361 | 0.363 |
| Additional 100M | 0.855 | 0.296 | 0.297 |
| Additional 300M | 0.938 | 0.115 | 0.116 |
| Additional 500M | 0.978 | 0.035 | 0.036 |
| Additional 1,000M | 0.772 | 0.001 | 0.001 |

*30% of cities*

| | GDP | GDP with Lag | First Difference GDP |
|---|---|---|---|
| No Adjustment | 0.367 | 0.998 | 0.994 |
| Additional 50M | 0.341 | 0.802 | 0.798 |
| Additional 100M | 0.316 | 0.618 | 0.614 |
| Additional 300M | 0.228 | 0.136 | 0.135 |
| Additional 500M | 0.160 | 0.013 | 0.013 |
| Additional 1,000M | 0.056 | **0.000** | **0.000** |

*50% of cities*

| | GDP | GDP with Lag | First Difference GDP |
|---|---|---|---|
| No Adjustment | 0.630 | 0.436 | 0.434 |
| Additional 50M | 0.582 | 0.264 | 0.263 |
| Additional 100M | 0.536 | 0.146 | 0.145 |
| Additional 300M | 0.373 | **0.005** | **0.005** |
| Additional 500M | 0.245 | **0.000** | **0.000** |
| Additional 1,000M | 0.066 | **0.000** | **0.000** |

**TABLE 4**
**P-VALUES SHOWING HOW LARGE THE ECONOMIC IMPACT OF AN EVENT MUST BE IN ORDER TO BE DETECTED, DISAGGREGATED BY MARKET SIZE**

*50% of cities by size*

| | GDP | GDP with Lag | First Difference GDP |
|---|---|---|---|
| No Adjustment | | | |
| Size 1 | 0.225 | 0.403 | 0.424 |
| Size 2 | 0.082 | 0.342 | 0.641 |
| Size 3 | 0.522 | 0.244 | 0.183 |
| Additional 50M | | | |
| Size 1 | 0.220 | 0.377 | 0.397 |
| Size 2 | 0.125 | 0.521 | 0.868 |
| Size 3 | 0.057 | **0.001** | **0.001** |
| Additional 100M | | | |
| Size 1 | 0.216 | 0.352 | 0.371 |
| Size 2 | 0.185 | 0.739 | 0.893 |
| Size 3 | **0.002** | **0.000** | **0.000** |
| Additional 300M | | | |
| Size 1 | 0.200 | 0.264 | 0.279 |
| Size 2 | **0.618** | 0.370 | 0.183 |
| Size 3 | **0.000** | **0.000** | **0.000** |
| Additional 500M | | | |
| Size 1 | 0.184 | 0.192 | 0.204 |
| Size 2 | **0.741** | 0.035 | 0.012 |
| Size 3 | **0.000** | **0.000** | **0.000** |
| Additional 1,000M | | | |
| Size 1 | 0.149 | 0.077 | 0.083 |
| Size 2 | **0.017** | **0.000** | **0.000** |
| Size 3 | **0.000** | **0.000** | **0.000** |

The summarized results on the impact of a new stadium or team are shown in Table 5. The bar is even higher than for events, in that if 30% of the cities examined acquired a new team or facility, the actual impact would need to be around $1 billion annually in order for the methods traditionally used in the academic literature to detect it. If 50% of the cities examined opened a new facility, the actual impact would need to be around $300 million in order to justify the use of the standard techniques.

**TABLE 5**
**P-VALUES SHOWING HOW LARGE THE ECONOMIC IMPACT OF A FACILITY MUST BE IN ORDER TO BE DETECTED**

*10% of cities*

| | GDP | GDP with Lag | First Difference GDP |
|---|---|---|---|
| No Adjustment | 0.492 | 0.800 | 0.803 |
| Additional 50M | 0.510 | 0.753 | 0.756 |
| Additional 100M | 0.529 | 0.707 | 0.710 |
| Additional 300M | 0.607 | 0.535 | 0.538 |
| Additional 500M | 0.690 | 0.388 | 0.389 |
| Additional 1,000M | 0.913 | 0.141 | 0.142 |

*30% of cities*

| | GDP | GDP with Lag | First Difference GDP |
|---|---|---|---|
| No Adjustment | 0.293 | 0.751 | 0.748 |
| Additional 50M | 0.317 | 0.841 | 0.838 |
| Additional 100M | 0.342 | 0.933 | 0.930 |
| Additional 300M | 0.455 | 0.704 | 0.707 |
| Additional 500M | 0.587 | 0.399 | 0.401 |
| Additional 1,000M | 0.972 | **0.046** | **0.046** |

*50% of cities*

| | GDP | GDP with Lag | First Difference GDP |
|---|---|---|---|
| No Adjustment | 0.746 | 0.373 | 0.373 |
| Additional 50M | 0.795 | 0.293 | 0.293 |
| Additional 100M | 0.844 | 0.226 | 0.226 |
| Additional 300M | 0.954 | **0.064** | **0.064** |
| Additional 500M | 0.755 | **0.013** | **0.013** |
| Additional 1,000M | 0.343 | **0.000** | **0.000** |

Hence, it is not surprising that an industry report might conclude that a new stadium in a top 50 market (which averages over $250 billion in annual GDP) would have an economic impact of $200 million per year, and the academic literature would conclude the opposite, that those same markets receive no net economic impact from a new facility based on panel regressions similar to those in this study because the thresholds to get a statistically significant result from those regressions are at $300 million or more. Accounting for the size of the markets (Table 6) shows that even if a new team or facility in a large market generates $1 billion per year in net new economic impact, the standard methodologies used in the academic literature will not detect it and will thus conclude that no impact exists. Small market economic impacts tend to become detectable at $50 million.

**TABLE 6**
**P-VALUES SHOWING HOW LARGE THE ECONOMIC IMPACT OF A FACILITY MUST BE IN ORDER TO BE DETECTED, DISAGGREGATED BY MARKET SIZE**

*50% of cities by size*

| | GDP | GDP with Lag | First Difference GDP |
|---|---|---|---|
| No Adjustment | | | |
| Size 1 | 0.766 | 0.393 | 0.392 |
| Size 2 | 0.478 | 0.712 | 0.707 |
| Size 3 | 0.765 | 0.369 | 0.264 |
| Additional 50M | | | |
| Size 1 | 0.774 | 0.381 | 0.379 |
| Size 2 | 0.369 | 0.848 | 0.815 |
| Size 3 | 0.369 | **0.038** | **0.040** |
| Additional 100M | | | |
| Size 1 | 0.782 | 0.369 | 0.367 |
| Size 2 | 0.276 | 0.988 | 0.928 |
| Size 3 | **0.036** | **0.001** | **0.003** |
| Additional 300M | | | |
| Size 1 | 0.814 | 0.323 | 0.322 |
| Size 2 | **0.065** | 0.487 | 0.631 |
| Size 3 | **0.000** | **0.000** | **0.000** |
| Additional 500M | | | |
| Size 1 | 0.847 | 0.281 | 0.281 |
| Size 2 | **0.009** | 0.160 | 0.294 |
| Size 3 | **0.000** | **0.000** | **0.000** |
| Additional 1,000M | | | |
| Size 1 | 0.929 | 0.194 | 0.194 |
| Size 2 | **0.000** | **0.002** | **0.015** |
| Size 3 | **0.000** | **0.000** | **0.000** |

The small number of times the academic literature has concluded that a sporting event or stadium (team) has generated a positive economic impact can be explained by the findings here and are not some unexpected abnormality, but instead driven by the context or size of the market in which teams participate, facilities are built, or events occur. Coates and Depken (2011), using monthly sales tax data, found that a season of Baylor University football games (held in Waco, Texas) generates an additional 1.8% to the local economy, or $175 million. Their finding is entirely consistent with the present results suggesting that smaller markets, like Waco, get discernible economic impacts from some sporting events or teams. Similar analyses of college football in smaller markets show positive and statistically significant results under some circumstances (Baade, Baumann, and Matheson, 2011).

The present research is also consistent with Agha's findings (2013). Her study focused on minor league baseball teams in 269 metro areas, some with populations lower than used in the present analysis. Noting that these smaller markets might allow for statistically significant findings unlike the majority of the previous academic research - which focused on the largest markets - Agha found that some minor league teams had positive and statistically significant effects of 0.2% to 0.7% of local per capita income, or \$67 to \$118 in per capita income. This translates to about \$30 million on average for each of the 233 small markets in the present data, which is consistent with the findings presented in Table 6. In other words, minor league teams in small markets that generate at least \$30 million per year in economic impact may marginally be detected as being statistically significant. In addition, Agha found significance for some levels of minor league baseball in many markets, but not across the board. In other words, there may be positive impacts from more of the markets studied by Agha, but those impacts could be \$20 million or less and thus not detectable by the standard academic methods utilized.

Given the large economies of the cities that tend to host major teams or sporting events, this paper demonstrates that the likelihood of committing a Type I error, not finding a positive economic impact when in fact there is one, is substantial. This can also be seen by noting the often large size of the confidence intervals around the coefficients of interest. E.g., in Baade, Baumann, and Matheson (2008) on the impact of major sporting events, Table 4 shows a coefficient on hosting the Super Bowl (with the dependent variable as the natural log of local taxable sales) to be 0.0336, a standard error of 0.0312, and a t-statistic of 1.08. The upper bound of a 95% confidence interval (not shown) would be 0.0948, implying that it also cannot be rejected that hosting the Super Bowl increases taxable sales by nearly 10 percent. In other words, the precision is very low for the sort of conclusions drawn.

More detailed and fine grained methods, such as the use of monthly data (Coates and Depken, 2011; Baumann and Matheson, 2018), daily hotel occupancy rates or tourist arrivals (Baumann and Matheson, 2017; Chikish, Humphreys, Liu, and Nowak, 2019) or a focus on smaller geographical regions is warranted (Feng and Humphreys, 2012).

These types of studies are occurring with more frequency, but will still have similar issues to solve albeit smaller. In other words, if monthly data is likely to provide a context that is twelve times smaller, there is still the possibility that an event impact will be subsumed by the size of the local monthly economy. Thus, future research should create similar guidelines for monthly data or daily occupancy rates. Additionally, shorter time periods have other potential problems such as if the timing of when monthly taxes are paid occurs months after an event (Coates and Depken, 2011), or if payments by event owners to local vendors occurs 90 days after an event (a personal communication with a major event owner by one of the authors confirms that this is often the case). The use of daily occupancy rates is even more fine-grained, but also may suffer from other measurement issues such as whether visitors stay in facilities captured by the data (e.g., is Airbnb data included, or do many visitors stay at a friend's residence?), or whether the location studied (e.g., within 4 miles of an arena) is where people stay when coming from out of town to attend an event at the arena. Yet, these same sets of data with specified periodicities and geographies can also be utilized to create threshold impacts that the models can sufficiently detect. Moreover, industry reports that attempt to account for the various errors presented by Crompton (1995) should not so easily be dismissed as clearly being incorrect simply because the results differ from the previous findings in the academic literature.

The present study demonstrates that much of the early seminal academic literature uses statistical analyses that are not sensitive enough to be able to detect economic impacts in the hundreds of millions (or even billions) of dollars in the large markets that tend to host major sports teams and events. This research helps provide guidelines for the size of events/teams/facilities that are detectable using standard techniques, and provides one explanation for the gap between the academic literature and industry reports that generate opposite findings. An industry report showing \$200 million in annual economic impact from a stadium might be a fair measure and yet that amount would not be detectable in the academic literature. Further research should focus on these same types of experimental trials focused on monthly data and daily occupancy rates. Moreover, testing any econometric models on any topics, not just sports economics, that fail to find statistically significant coefficients using experimental trials of this sort (where impacts are randomly injected into the dependent variable) can help provide thresholds and guidelines of how Type I errors may occur.

## ENDNOTES

[1] The authors would like to thank Victor Matheson for valuable feedback on an earlier version of the paper. Any and all errors are our own.

[2] The average 2017 GDP of the top 50 largest MSAs in the U.S. exceeds $250 billion.

[3] Variations that allow for different growth rates across the three market sizes are also tested by allowing for separate yearly indicator variables across market sizes. In other words, if smaller markets tend to grow faster than larger markets, the more refined yearly dummy variables will incorporate that.

[4] There are certainly costs to the public and local government when hosting a major event. These often include the effects of traffic congestion, overtime pay for police, fire, and other city services.

## REFERENCES


Agha, N. (2013). The economic impact of stadia and teams: the case of minor league baseball *Journal of Sports Economics*, *14*(3), 227-252.

Agha, N., & Rascher, D. (2016). An explanation of economic impact: Why positive impacts can exist for smaller sports. *Sport, Business and Management*, *6*(2), 182-204.

Agha, N., & Taks, M. (2018). Modeling resident spending behavior during sport events: do residents contribute to economic impact? *Journal of Sport Management*, *32*(5), 473-485.

Baade, R.A. (1987, February 23). *Is there an economic rationale for subsidizing sports stadiums?* Policy Study No. 13. The Heartland Institute. Retrieved from https://www.heartland.org/_template-assets/documents/publications/17280.pdf

Baade, R.A. (1994, April 4). *Stadiums, professional sports, and economic development: Assessing the reality*. The Heartland Institute. Retrieved from https://www.heartland.org/_template-assets/documents/publications/8828.pdf

Baade, R.A. (1996, March). Professional sports as catalysts for metropolitan economic development. *Journal of Urban Affairs*, *18*(1), 1-17.

Baade, R.A., & Dye, R.F. (1990). The impact of stadiums and professional sports on metropolitan area development. *Growth & Change*, *21*(2), 1-14.

Baade, R.A., & Matheson, V.A. (2001, November). Home run or wild pitch? Assessing the impact of Major League Baseball's All-Star Game. *Journal of Sports Economics*, *2*(4), 307-327.

Baade, R.A., Baumann, R.W., & Matheson, V.A. (2008a). Selling the game: estimating the economic impact of professional sports through taxable sales. *Southern Economic Journal*, *74*(3), 794-810.

Baade, R.A., Baumann, R.W., & Matheson, V.A. (2008b, December). Assessing the economic impact of college football games on local economies. *Journal of Sports Economics*, *9*(6), 628-643.

Baade, R.A., Baumann, R.W., & Matheson, V.A. (2011, March). Big men on campus: estimating the economic impact of college sports on local economies. *Regional Studies*, *45*(3), 371-380.

Baim, D.V. (1990, November 26). *Sports stadiums as "wise investments": An evaluation*. The Heartland Institute. Retrieved from https://www.heartland.org/_template-assets/documents/publications/27054.pdf

Baumann, R.W., & Matheson, V.A. (2017). Many happy returns? The Pro-Bowl, mega-events, and tourism in Hawaii. *Tourism Economics*, *23*(4), 788-802.

Baumann, R.W., & Matheson, V.A. (2018, April) Mega-events and tourism: the case of Brazil. *Contemporary Economic Policy*, *36*(2), 292-301.

Bergman, B. (2015, March). *Are pro sports teams economic winners for cities?* Marketplace. Retrieved from https://www.marketplace.org/2015/03/19/are-pro-sports-teams-economic-winners-cities/

Chikish, Y., Humphreys, B., Liu, C., & Nowak, A. (2019). Sports-led tourism, spatial displacement, and hotel demand. *Economic Inquiry*, *57*(4), 1859-1878.

Coates, D. (2007, October). Stadiums and arenas: Economic development or economic redistribution? *Contemporary Economic Policy*, *25*(4), 565-577.

Coates, D., & Depken, C.A. (2009, December). Do college football games pay for themselves? The impact of college football games on local sales tax revenues. *Eastern Economic Review*, *35*(4), 531-547.
Coates, D., & Depken, C.A. (2011). Mega-events: Is Baylor football to Waco what the Super Bowl is to Houston*? Journal of Sports Economics*, *12*(6), 599-620.
Coates, D., & Humphreys, B. R. (2002). The economic impact of postseason play in professional sports. *Journal of Sports Economics*, *3*(3), 291-299.
Coates, D., & Humphreys, B.R. (1999). The growth effects of sport franchises, stadia, and arenas. *Journal of Policy Analysis and Management*, *18*(4), 601-624.
Coates, D., & Humphreys, B.R. (2003). Professional sports facilities, franchises, and urban economic development. *Public Finance and Management*, *3*(3), 335-357.
Crompton, J.L. (1995, Jan). Economic impact analysis of sports facilities and events: Eleven sources of misapplication. *Journal of Sport Management*, *9*(2), 14-35.
Feng, X., & Humprheys, B. (2012). The impact of professional sports facilities on housing values: evidence from census block group data. *City, Culture and Society*, *3*, 189-200.
Garofalo, P., & Waldron, T. (2012, September 7). If you build it, they might not come: The risky economics of sports stadiums. *The Atlantic*. Retrieved from https://www.theatlantic.com/business/archive/2012/09/if-you-build-it-they-might-not-come-the-risky-economics-of-sports-stadiums/260900/
Jeanrenaud, C. (2007). Sports events: Uses and abuses of economic impact studies. *Finance & the Common Good*, *1*(26), 99-104.
Kanell, M.E. (2019, February 1). Super Bowl economic bonanza? Depends who's talking. *Atlanta Journal Constitution*. Retrieved from https://www.ajc.com/business/super-bowl-economic-bonanza-depends-who-talking/jXaVJa2ekoIFstBHJ27HLO/
Matheson, V. (2018). Is there a case for subsidizing sports stadiums? *Journal of Policy Analysis and Management*, *38*(1), 271-277.
Matheson, V.A., & Baade, R.A. (2003, September). *An economic slam dunk or March Madness? Assessing the economic impact of the NCAA Basketball Tournament*. College of the Holy Cross, Department of Economics Faculty Research Series. Retrieved from https://www.nku.edu/~lipping/PHE385/ncaa.pdf
Matheson, V.A., & Baade, R.A. (2006) Padding required: assessing the economic impact of the Super Bowl. *European Sport Management Quarterly*, *6*(4), 353-374.
Okner, B.A. (1974). Subsidies of stadiums and arenas. In R. Noll (Ed.), *Government and the sports business* (pp. 325-348). Washington, DC: Brookings.
Porter, P. (1999). Mega-sports events as municipal investments: A critique of impact analysis. In J. Fizel, E. Gustafson, & L. Hadley (Eds.), *Sports economics: Current research* (pp. 61-74).Westport, CT: Praeger.
Santo, C. (2005). The economic impact of sports stadiums: recasting the analysis in context. *Journal of Urban Affairs*, *27*(2), 177-191.