Author Manuscript

Anticipating the Honeymoon: Event Study Estimation of New Stadium Effects in Major League Baseball Using the Imputation Method

Stefan Szymanski[1]

October 2022, revised January 2023

Abstract: This paper estimates the impact of new stadiums on attendance and revenues of Major League Baseball teams between 1970 and 2019. Recent studies reveal that two-way fixed effects (TWFE) models may produce biased estimates, proposing an "imputation" method instead. This paper uses the imputation method to generate a counterfactual estimate, based on untreated observations: the average treatment effect on the treated (ATT) equals the difference between actual and counterfactual estimates. The analysis shows that there were significant anticipation effects associated with new stadiums, up to three seasons before opening. It suggests previous estimates may significantly understate new stadium revenue gains.

Keywords: Event study, Imputation, Stadium attendance

JEL Classification: C21, L83, Z20

Abbreviations: TWFE: two-way fixed effects, ATT: Average treatment effect of the treated, MLB: Major League Baseball, MSA: Metropolitan Statistical Area, CH: Coates and Humphreys (2005), ZL: Zygmont and Leadley (2005), wpc: win percentage.

[1] University of Michigan, Sport Management Department, Ann Arbor, Michigan, 48109. stefansz@umich.edu

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/ecin.13149.

This article is protected by copyright. All rights reserved.

Author Manuscript

1. Introduction

Event studies are a widely used tool for economic policy analysis. Recent research has questioned the reliability of two-way fixed models (TWFE) commonly used to estimate treatment effects and offered new approaches to identification and estimation, e.g. Borusyak and Jaravel (2017), Sun and Abraham, (2021), de Chaisemartin and D'Haultfœuille (2020), Goodman-Bacon (2021), Callaway and Sant'Anna (2021). This paper focuses on the estimation procedure suggested in recent work by Borusyak et al (2022), Gardner (2021) and Liu et al (2021). This involves estimating a model using never treated and not yet treated units to generate a counterfactual case, which can then identify the causal impact of the treatment.

The impact of new stadia on attendance at games in professional sports leagues provides an almost ideal framework for testing event study estimation methods. Attendance data are available at the game level over many decades. Treated teams move to a new stadium at clearly identified dates, while teams that were not treated in the sample period constitute the control group. Unusually, it is possible to generate regression-free unconditional estimates of the average treatment effect on the treated (ATT) by using the games played between never-treated and treated teams as a matched sample, comparing games where the treated team is the home team, playing at the stadium in question, to those where it is the visiting team.

The policy relevance arises due to the large public subsidies often associated with the construction of new stadia. Zygmont and Leadley (2005), Coates and Humphreys (2005), Leadley and Zygmont (2006) and Soebbing et al (2016) all use TWFE models to estimate the impact of new stadia. These studies find significant attendance and price effects, sometimes described as "honeymoon effects". This paper extends the research in three ways. First, the estimation procedure addresses challenges posed by the conventional TWFE approach. Second, this paper considers anticipation effects not previously addressed in the literature: it shows that attendance rises significantly in the three years leading up to the opening of a new stadium, which coincides roughly with the construction duration. Third, unlike previous research, the paper uses game level rather than seasonal average attendance data, facilitating the almost ideal matching procedure.

The professional sport studied in this paper is Major League Baseball (MLB). Since 1970, the league has expanded from 24 to 30 teams, and in the half century since 1970, 24 new stadia have been built, including 20 than opened between 1992 and 2017. In these new stadia, attendance levels in opening season were significantly higher than in the seasons before ground was broken in the construction of a new stadium,

This article is protected by copyright. All rights reserved.

and also rose significantly in the construction period. The typical stadium broke ground two to three seasons prior to opening. Using the raw data, relative to average annual attendance in the four to ten seasons before opening, attendance was 54% higher in opening season, 26% higher in the season before opening, 18% higher two seasons before opening and 20% higher three seasons before opening. From the first year after opening, attendance trends back to the level seen in the year before opening, but does not fall significantly below this level. This pattern of results is confirmed in the regression analysis.

The post-opening estimates are close to those found in previous studies, but previous studies mostly neglected the anticipation effect. Since there is no evidence that, in the longer term, attendance falls below the level achieved in the anticipation period, the implication is that there may be a long term gain not identified in previous research. The data on ticket prices reveals an average price increase of 26% relative to the control group in the opening season of a new stadium, returning to parity by around ten seasons after opening. Based on a conservative estimate that the attendance effects also last for only ten seasons, the additional revenues created by a new stadium exceeded the cost in only five out of eighteen stadia built since 1992, but (given the large public subsidies) exceeded the private sector share of the cost in 13 out of 18 cases.

2. Literature review

2.1 Two-way fixed effects and the imputation estimator

The TWFE model has been described as the workhorse for the estimation of causal effects in event studies. Equation (1) illustrates the typical specification:

$$(1) \qquad Y_{it} = \alpha_i + \beta_t + \mathbf{D_{it}}\,\boldsymbol{\tau} + \mathbf{X_{it}}\,\boldsymbol{\gamma} + \varepsilon_{it}$$

where outcome Y_{it} and binary treatment D_{it} are measured in periods t and for units i, α_i are unit fixed effects, β_t are period fixed effects and X_{it} is a matrix of potential time-varying unit-specific confounding factors. The specification is meant to isolate the ATT from unit- and period-specific confounders. The dynamic treatment specification adopted here and in many papers allows for the ATT to vary by period, by specifying both leads and lags.

Recent research suggests that the TWFE model can be problematic – see e.g. Borusyak and Jaravel (2017), de Chaisemartin and D'Haultfoeillle (2020), Goodman-Bacon (2021), Callaway and Sant'Anna (2021),

This article is protected by copyright. All rights reserved.

Author Manuscript

Baker et al (2022). Borusyak et al (2022, henceforth BJS) highlight three main problems with the TWFE model. First, if there are no "never-treated" groups, the problem is under-identified and the treatment effect is collinear with the unit and period fixed effects (BJS, pp12-13). Second, when treatments are staggered, estimation of the treatment effect on later-treated units relies in part on a comparison with earlier treated units, the so-called "forbidden comparisons". This leads to an underweighting of the effect for later treated units, and can even imply that these observations are given a negative weighting in the estimation (BJS, pp. 13-15). Third, for later treated units, no valid comparators may exist to generate treatment effects if there are no untreated units. Roth (2022) and Sun and Abraham (2021) also show that the model presents problems for the consistent estimation of pre-trend effects which are used to test the parallel trends assumption.

Three recent papers, BJS, Gardner (2021) and Liu et al (2021) propose a novel imputation method to avoid these problems. As Imbens and Rubin (2015, p.141) observe, "At some level, all methods for causal inference can be viewed as imputation methods, although some more explicitly than others". BJS describe an explicit procedure for deriving a robust and efficient estimator as follows:

> First the unit and period fixed effects $\hat{\alpha}_i$ and $\hat{\beta}_t$ are fitted by regressions using untreated observations only. Second, these fixed effects are used to impute the untreated potential outcomes and therefore obtain an estimated treatment effect $\hat{\tau}_{it} = Y_{it} - \hat{\alpha}_i - \hat{\beta}_t$ for each treated observation. Finally, a weighted sum of these treatment effects estimates is taken, with weights corresponding to the estimation target.

This is essentially the same as the two stage procedure proposed by Gardner (2021, see pp 8-9) and the four step procedure by Liu et al. (2021, p.10). The imputation method is adopted in this paper, allowing in addition for the effect of confounders X_{it}. The logic of the imputation method is that it uses all untreated observations (whether for never treated or not-yet-treated units) to generate a counterfactual prediction of the target variable, absent treatment. The treatment effect is then identified as the difference between the counterfactual prediction and the actual value of the target variable.

A significant advantage of the imputation method is that it can be used to distinguish anticipation effects from pre-trends, which are used to test the parallel trends assumption. The approach with TWFE models is to test for the values of the pre-treatment coefficients. With the imputation method, it is possible to

This article is protected by copyright. All rights reserved.

Author Manuscript

specify anticipation effects in the counterfactual regression, while separately identifying pre-trend effects (BJS, pp 25-26).

A further advantage of the imputation method is that it can allow for unrestricted treatment effect heterogeneity, which poses a significant problem for TWFE models (see e.g. Sun and Abraham (2021)). Since the imputed effect is specific to each unit of observation (relative to the counterfactual), this allows for arbitrary treatment effect heterogeneity (BJS, Proposition 6, p21). While it remains a challenge to separate the causal effect from the error term, BJS propose a conservative approach to calculating standard errors, which attributes some variation in the treatment effects to the error terms (BJS, section 4.3, pp23-25).

2.2 Stadium effects

There are two broad strands in the sports stadium literature. The first concerns the impact of new stadia on the local economy, for a comprehensive survey see Bradbury et al (2022). The second, which is the subject of this paper, concerns the direct impact of new stadia on the sports business itself, in terms of attendance, prices and revenues. This literature[2] describes the impact of stadium age as a "honeymoon effect" (also more prosaically known as the "novelty effect"), e.g. Clapp and Hakes (2005), Leadley and Zygmont (2005) and Agha (2013): "Opening a new stadium or arena may result in an initial surge in attendance, but in time, the novelty of the new facility fades and attendance begins to decline", Leadley and Zygmont (2005), p203.

Table 1 summarizes the results drawn from seven studies that estimate season specific effects, covering eight different leagues.[3] The estimated impact in the year the stadium opens ranges from 9% to 50%, and then decreases linearly up to the 10th year after opening (which is the limit of most studies), by which time the effect is negligible. The last two columns of Table 1 report the seasonal averages, both for all estimates

[2] The literature started with Noll (1974), who estimated an attendance demand equation with a linear coefficient for stadium age up to ten years and found a significant effect

[3] These are not the only studies to have estimated stadium attendance effects. In addition to Noll (1974), McEvoy et al (2005) estimate age effects using a quadratic functional form while Poitras and Hadley (2006) estimate only an opening year dummy and new stadium dummy covering all subsequent years. These forms are unlikely to reliably pick up variations from season to season, and so comparison is omitted here. Depken (2006) examines the impact of stadium age on financial variables but not attendance. Hakes and Clapp (2006) use the data from their 2005 paper to estimate revenue effects.

.

This article is protected by copyright. All rights reserved.

Author Manuscript

and including only estimates that were statistically significant at the 5% level or better. Figure 1 is a chart of the first of these averages.

Of the studies cited in Table 1, four used the TWFE model - Zygmont and Leadley (2005), Coates and Humphreys (2005), Leadley and Zygmont (2006) and Soebbing et al (2016) - while the others omitted unit and time fixed effects. All of the studies cited here use season–level attendance averages rather than game level data.

Table 1 and Figure 1 here

Most of these studies also considered ticket prices. The more recent studies of Gitter and Rhoads (2014) and Soebbing et al (2016) examined minor leagues in the US and Canada, where pricing is considerably lower than in the major leagues. Some of these studies combined the attendance and price data to estimate revenue effects. Zygmont and Leadley (2005) find large revenue effects relative to construction costs, while Coates and Humphreys (2005) and Gitter and Rhoads (2014) find modest effects. These and other estimates are discussed further below.

The studies examining major league stadia effects used data ranging from around 1950 to 2002. They therefore gave limited coverage to the generation of new stadia built since the opening of Camden Yards as the baseball stadium of the Baltimore Orioles in 1992. Its fan-friendly design inspired the generation of stadia that followed, and it is on these stadia that this study will primarily focus.

3. Data and descriptive statistics

This study uses attendance and price data for MLB. Game level attendance data are available from www.retrosheet.org, season level estimates of ticket and car parking prices van be found in the Fan Cost Index produced by Team Marketing Reports (https://teammarketing.com/fancostindex/). Attendance data was collected for the half century from 1970-2019, while the price estimates are only available from 1991 on. Table 2 identifies the stadium age of MLB teams and divides them into two groups, a treatment group that opened at least one new stadium since 1973, and a control group consisting of teams that did

This article is protected by copyright. All rights reserved.

Author Manuscript

not open a new stadium.[4] The data includes 24 new stadia, 20 of which are from the "Camden Yards" era, i.e. the era of more innovative stadium construction.[5]

Each team in MLB plays approximately 81 home games in a season, and the attendance dataset consists of 110,163 games, of which 2,555 do not a have an attendance figure (most of these are doubleheaders, and two thirds of them are from the 1970s). Table 3 illustrates the distribution of games between treated teams (before or after stadium construction) and never-treated teams.

Of particular interest here are the games on the off-diagonal, where treated teams play never-treated teams. This presents an unusual, almost ideal, opportunity in the context of event studies with staggered treatments. Generally, there is no obvious matching with respect to timing between the treatment group and the control group. Here, it is possible to create a profile of attendance, based around the treatment date, for home games between treated teams and never-treated teams. In other words, there is a natural matching between the treatment group and a subset of the control group, where essentially the only difference between pairs in the sample is the location of the stadium.[6]

Tables 2 and 3 here

It is possible then to use these profiles to estimate the unconditional treatment effect without using regressions, as is shown in figure 2.

Figure 2 here

[4] In the period, there were four expansion franchises, two in 1993 and two in 1998, and it is assumed that there are no stadium specific attendance effects associated with expansion (i.e. while there may be a "franchise novelty" effect, this is a distinct effect from the effect associated with an existing franchise opening a new stadium). One of the expansion franchises (the Miami Marlins) has since opened a new stadium. Three teams opened two new stadia in this period. Given the long time series and the large gap between opening these stadia, it is possible to create pre-trends, anticipation and treatment effect dummies for both stadia.

[5] In an ideal setting, the decision to build a new stadium would be randomly assigned to each team. Instead, there are possible selection effects. However, there is a good deal of uncertainty concerning the decision to build a new stadium. Moreover, replacement stadia have in the past been considered for stadia such as Wrigley Field (Chicago Cubs) and Fenway Park (Boston Red Sox). By contrast, Detroit's Tiger Stadium, which was equally iconic, was in fact replaced. Negotiations and planning over replacement stadia can span over decades, or can be wrapped in a short space of time, depending on the uncertain political and economic environment.

[6] Other differences are possible, including time of season, time of day and context of season, as well as differences in playing personnel. Most of these differences are used as controls in the regression estimates reported in the next section.

This article is protected by copyright. All rights reserved.

Author Manuscript

Visual inspection of Figure 2 suggests the following interpretation:

- Up to and including four seasons before the opening of a new stadium, the average gap in attendance between control group and treatment group teams is around 5500 per game, while the trends of attendance at the treatment group and control group are indistinguishable.
- In the three seasons prior to opening, there is a significant step increase in attendance at treatment group teams, equal to just under 3000 in the third season before opening and an additional 1400 in the season before opening. As a result, by the year before opening the confidence intervals of the control group and treatment group overlap.
- There is a large increase in attendance in the season that a stadium opens, equal to around 6,500 people per game, which raises attendance significantly above the level of the control group.
- In the following nine seasons attendance slowly declines to the point where it is indistinguishable from attendance at the control group.

This suggests that in addition to the short term impacts of opening a new stadium (the Honeymoon Effect), there is in fact a longer term effect factoring in the anticipation effect, which raises the level of attendance equal to that of the control group.

To date the literature has paid little or no attention to anticipation effects. Clapp and Hakes (2005) report a significant *negative* effect on attendance in the final season of an old stadium, except in cases where the stadium was first built prior to 1950. Gitter and Rhoads (2014) find a positive effect on attendance for the final year of a single A minor league baseball stadium, but significant negative effects in the final and penultimate years for AA and AAA stadia.

However, there are good reasons to believe that positive anticipation effects exist. At least four possible factors can be identified[7]:

(i) Nostalgia effects – longstanding fans who identify with a particular facility may have a stronger desire to attend games at a ballpark knowing that it will soon be demolished.

(ii) Separation effects – fans living close to an old stadium but a long distance from a new facility may be unwilling to travel to the new stadium, and may therefore have an increased demand

[7] A further effect might be a decision by ownership to invest in a better quality team in advance of moving to a new stadium. However, as we will see below, there is no evidence of any improvement in team performance in the years immediately before opening a new stadium.

This article is protected by copyright. All rights reserved.

Author Manuscript

for games at the old stadium (this could be particularly true for fans with low mobility and/or low incomes)

(iii) Commitment effects - fan commitment to a team might increase when owners commit to building a new stadium, since the prospect of relocation is diminished (for committed fans, relocation is an emotional trauma, and hence not-yet-committed fans may want to see owner commitment before making their own commitment)

(iv) Queueing effects – tickets at a new facility may be harder to obtain, and fans may be able to secure priority by going to games at the old stadium.

The key trigger for these anticipation effects is likely to be the date when ground is broken for the new construction, i.e. when sunk costs are incurred. Prior to that, considerable uncertainty remains, since there are usually several potential options, and negotiations take place in secrecy. Breaking ground can be used as a signal of owner commitment which affects fan attitudes to attendance.

Table 4 provides data on the gap between breaking ground and opening for the 22 MLB stadia completed between 1989 and 2017. In almost all cases, the first game in the new stadium took place at the start of the season, so the gap in years is estimated as the difference between the date of breaking ground and April 1st of the opening season. In almost all cases, ground was broken two to three years before opening. For the anticipation effect to bite, we need to assume that fans in a given season prior to opening are aware that a new stadium will be built. The last column of Table 4 shows the number of seasons spanned by the gap. In ten cases the gap spans two seasons, in ten cases three seasons, and in two cases four seasons.

Table 4 here

Author Manuscript

This article is protected by copyright. All rights reserved.

Author Manuscript

4. Estimation and results

4.1 Attendance

We now compare the imputation based estimates with the results from a conventional TWFE model. For the TWFE models, treatment dummies are included for seasons spanning ten years before and up to 20 seasons after the stadium opened. For the imputation model, seven seasons of pre-trends are specified, starting ten years before opening a new stadium, and treatment effects are estimated for the following 23 seasons, where the first three are anticipation effects (anticipation effects longer than three seasons were not statistically significant).[8]

Figure 3 illustrates the two sets of estimates based on the full set of game observations. The coefficients are reported in the first two columns of Table A1. The regression models included a number of additional controls based on

(a) Timing: month, night games, weekend games[9], holiday games[10], and the first game of the season
(b) Team success: win percentage in the season to date of the home team and visiting team, the home team's success in the previous season (winning percentage over 50%, participation in the post-season, and performance in the World Series)
(c) Context: interleague play and whether the home team is playing in a temporary stadium
(d) Demographics: population and per capita income in the MSA of the home team[11]
(e) New franchise effects: dummies for each of the first eight seasons of the four new franchises added in the 1990s.

All of the control variables have the expected signs, and all except the per capita income variables are statistically significant. The coefficients are reported in the first two columns of Table A2.

[8] The TWFE model is estimated using the reghdfe command and the imputation model is estimated using the did_imputation Stata module (Borusyak (2021)). In both cases standard errors are clustered by home team.
[9] Weekend games are defined as Saturday and Sunday games. An anonymous referee suggested that Friday games could also be considered weekend games – experimenting with this alternative made no significant difference to the coefficients of interest.
[10] Labor Day, 4th of July, Memorial Day
[11] Bureau of Economic Analysis data for US Metropolitan Statistical Areas. Data for Montreal and Toronto based on Canadian GDP per capita (World Bank) and city population estimates from UN World Population Prospects.

This article is protected by copyright. All rights reserved.

Author Manuscript

Figure 3 here

Figure 3 shows that both imputation and TWFE estimates indicate a sharp increase in attendance in the season that a stadium opens, followed a significant but declining effect up to around twelve seasons after opening. The TWFE estimates are larger than the imputation estimates, by around one to two thousand fans per game, but the TWFE estimates are also much less well defined, and the 95% confidence intervals are about twice the size of those of the imputation model up to around ten years after the stadium opens.

Identification of anticipation effects is one of the main advantages of the imputation approach to event studies (alongside the ability to adopt a principled approach to testing for pre-trends). In a TWFE model a lead dummy variable (i.e. a dummy for a season before a new stadium opened) could represent either an anticipation effect or part of a pre-treatment trend – there is no identifiable difference. With the imputation approach, if a period prior to treatment is deemed subject to the anticipation effect, the sample used to estimate the auxiliary (counterfactual) regression is altered, i.e. it won't include this observation, and the anticipation effect will be calculated as the difference between the actual value (of attendance in this case) and the counterfactual estimate. Thus pre-trends are identified separately from anticipation effects (Borusyak et al (2022), pp25-27).

This is of practical significance. Two studies cited in Table 1 examined attendance effects of new stadia for MLB teams -- Coates and Humphreys (2005, CH) and Zygmont and Leadley (2005, ZL). Neither of these identified anticipation effects. However, the estimated ATT reported here in the first year after opening, relative to the season prior to opening, is close to the figure reported in CH and ZL. In other words, if the anticipation effects are subtracted from the post opening effects, then the values are very similar to those of CH and ZL. The imputation model estimates reveal that game attendance in the season before opening has increased by 4,407 relative to the pre-anticipation baseline, while in the season the stadium opens the increase is 10,069 (also relative to the pre-anticipation baseline). The difference (6,282) is remarkably close to the estimated first year impact of a new MLB stadium reported in CH (5,887) and ZL ("approximately 6000"). This implies that omitting the anticipation effects may significantly understate the ATT.

This also has implications for statistical significance. If the three seasons prior to opening are taken to be the baseline, then treatment effects beyond the first ten seasons would not only be smaller, they would

This article is protected by copyright. All rights reserved.

appear statistically insignificant. The coefficients in CH and ZL are statistically significant up to year 8, compared to 12 years after opening based on the imputation estimates.

The interpretation suggested by the imputation model is consistent with the unconditional regression-free estimates described in Figure 2, where there is a step increase in attendance three seasons before opening. If the TWFE model estimates were relied upon, the anticipation effect would be associated only with the last season of the old stadium rather than three seasons. An effect covering the last three seasons of the old stadium is consistent with the gap between opening of the new stadium and breaking ground, as shown in Table 4.

The "parallel trends" assumption is a necessary condition for identification of the ATT. Here that is tested using the pre-trends, and the fact that the coefficients on the pre-trend variables are not statistically different from zero supports the parallel trends assumption. However, since the TWFE estimates fail to identify significant pre-treatment effects apart from the last season prior to opening, it would be easy to draw the mistaken conclusion that there is a much more limited anticipation effect.

Figure 4 illustrates a different comparison – between the full sample imputation estimates and a restricted sample based on the matched teams used to construct figure 2. In the matched sample, games within the treatment group (i.e. treated team vs treated team) and within the control group (never treated team vs never treated team) are excluded (see Table 2). This represents a check on selection bias of the matchups between teams. If there were characteristics of the treated group that made games among them (only) more attractive than games against never treated teams, then the coefficients for the treatment effects would be significantly different (the matched sample coefficients are reported in column 3 of Table A1 and A2). These characteristics might be associated with the structure of league competition itself (e.g. the placement of teams in leagues or divisions).

Figure 4 here

In fact, inspection of figure 4 shows that the estimates are very similar. About eight seasons after opening, some divergence in estimates emerges, and by twelve seasons after opening, the matched sample estimates become statistically insignificant. However, for the ten seasons before and after opening the

This article is protected by copyright. All rights reserved.

Author Manuscript

differences are small, suggesting that there is no material bias arising from the structure of the team matchups in the data.

The models so far have been estimated using game level data, unlike previous studies which used seasonal averages. Figure 5 illustrates the difference when attendance figures are aggregated to the seasonal level (column 5 of Tables A1 and A2). Using only seasonal data represents a significant restriction. All within season effects are suppressed, while the degrees of freedom are substantially reduced. This also restricts the nature of the comparisons that can be made, e.g. controlling for game matchups discussed above is not possible. Figure 5 suggests that, in this case, this ends up making little difference – and the two sets of estimates are almost indistinguishable.

Figure 5 here

For completeness, Figure 6 reports the point estimates for the six model variants reported in Table A1 (adding the TWFE model variants for the matched sample and season level data to the four models described above). Overall, the models suggest a fairly high degree of consistency in the estimates of the attendance effects, although there is more variance in the estimates the further removed in time from the opening of the stadium. Importantly, none of the pre-trend estimates are significantly different from zero in any of the models, suggesting that the parallel trends assumption cannot be rejected (the coefficients are reported in Table A1).

The imputation model estimates in figure 6 are shown in bolder type. The estimated treatment effects from the imputation model appear consistently lower than the TWFE estimates. The size of the difference, when comparing each pair of estimates (e.g. full sample imputation and full sample TWFE) averages around 1,000 spectators per game. By the logic of the imputation approach, the TWFE models are overestimating the impact of a new stadium by including in the estimation observations of attendance at games where the new stadium has already opened.

Figure 6 here

Author Manuscript

This article is protected by copyright. All rights reserved.

A number of variations used to test robustness are reported in Table A3. Column R1 is estimated excluding the four expansion franchises (Miami, Colorado, Tampa Bay and Arizona). Column R2 in addition excludes the Washington Nationals, who played as the Montreal Expos until 2004. Column R3 excludes all stadia built before 1992. Column R4 drops all cases where capacity utilization exceeded 100% (about 2% of all games).[12]

In all cases the size and significance of the estimates are similar to the central model (full sample imputation) up to ten seasons after the opening of the new stadium, but after that point there is some divergence. In columns R1, R2 and R4, the estimates are all very close to the central model- none of the stadium effects from the anticipation period (three seasons before opening) up to the tenth season after opening being more than 10% different in magnitude, and the majority being within 5%. In columns R3 the new stadium estimates are similar to the other models for the first ten years after opening but are significantly larger from ten years on. This seems consistent with the argument that the new generation of stadia after Camden Yards was particularly attractive to fans. Some caution is required when interpreting estimates beyond ten years after stadium opening, especially since many of these new stadia have been open for little more than ten years. However, the robustness tests show that the estimates are stable from the pre-treatment period and up to around ten years after opening.

4.2 Team success

A surprising feature of the data are that there is no evidence that a new stadium has any effect on team success. Team success is itself a significant driver of attendance, and hence one might have expected to observe some complementarity between a new stadium, attendance and team success. Quinn et al (2003) find evidence of this complementarity for MLB, but not for the other North American major leagues, while Clapp and Hakes (2005) find no evidence of such complementarity for MLB. Complementarity might be expected to come through one of two mechanisms. First, increased attendance could lead directly to better results through a kind of encouragement effect (often discussed, particularly in terms of the

[12] It is perhaps surprising that there is so little evidence in the data of capacity constraints, given that sell outs are commonplace in many other professional leagues. Baseball has always been different, given the very large number of games played in the season. Moreover, the attendance statistics may actually overstate the extent of capacity utilization, given that reported figures are based on tickets sold, and do not account for no-shows. Furthermore, capacity figures are often indicative, and stadium managers often have flexibility in determining admissions. This is reflected in the fact that in many cases attendance is in excess of reported capacity. Overall, capacity constraints appear to be a relatively limited problem in MLB, certainly in comparison with other professional league sports.

This article is protected by copyright. All rights reserved.

Author Manuscript

potential to impact the decisions of umpires, see e.g. Garicano and Palacios-Huerta (2005)). Second, team owners might invest in a better team in order to attract more fans to the new stadium.[13]

Figure 7 is based on the matched sample used for Figure 2. It shows the trends of treatment and control group home win percentages from ten seasons before opening up to twenty seasons after opening. If there were any effect, one would expect it to obervable in performance of team when playing in its home stadium. The time series clearly suggest there is no overall trend, before or after the opening of a new stadium. Nor is there any evidence of a shift associated with the opening season.

This finding is further corroborated by evidence on the number of winning seasons and appearances in World Series. In the ten seasons prior to opening there were 106 winning seasons (win percentage greater than .500) out of 232 treated team seasons, while in the ten seasons after there were 118 winning seasons out of 231. The t statistic for the difference in proportions is 1.16, implying there was no statistically significant difference. From this sample, 15 made it to the World Series with six winners before opening, while 16 made it after opening, also with 6 winners. Once again, there appears to be no significant change in team performance after opening a new stadium.

Figure 7 here

4.3 Prices

Ticket price data were obtained from Team Marketing Report's Fan Cost Index (FCI), which provides estimates of typical ticket prices at each ballpark going back to 1991. In reality, the price paid for a ticket can vary considerably based on the location of the seat and on each game, but the (FCI) price provides some insight into trends over time and comparisons between stadia. In addition to ticket prices, FCI provides estimates of related costs of attending an event, including car parking fees.

Figure 8 illustrates the average cost of a ticket plus car parking fee for the treated teams and the matched sample control group. The latter shows a clear linear trend, while the ticket price of the treated teams jumps significantly on opening of the new stadium. There appears to be no anticipation effect related to

[13] As one anonymous referee suggested, they might also invest in more playing talent in the anticipation period.

This article is protected by copyright. All rights reserved.

Author Manuscript

ticket prices. Prior to opening of a new stadium, the ticket price of the treatment group appears to run parallel to the control group, at a level that is about $3-5 lower. From 11 to 17 years after opening this relative trend appears to resume. Thus what appears to occur is that opening a new stadium produces a step increase, so that price becomes significantly higher than in the treatment group, but over time this premium declines and disappears around ten years after the stadium opens, resuming the pre-opening trend.

Figure 8 here

4.4 Revenues

The ticket price and attendance data can be combined to estimate the additional revenue attributable to a new stadium. ZL estimate that the combined effect of attendance and ticket price increases would increase revenues by $228 million over 15 years (this is the present value based on a 5% discount rate), which they contrast with "an approximate cost of $250 million for constructing a new facility".[14] CH, who estimate the revenue increases for three ballparks built in the 1990s, find that the increases were in the range of $30-40 million, undiscounted over the first eight years, insufficient to cover interest payments on long term bonds issued to finance the stadia.[15] Hakes and Clapp (2006) use the same attendance data as CH and combine it with FCI ticket price data to estimate that a new MLB stadium opened in 2004 would generate an increase in total revenues of $272 million over ten years in present value terms.

To estimate the revenue impact, we identify the additional fans (actual less counterfactual, based on the full sample imputation estimates (figure 3)) and the counterfactual number of fans. The additional fan estimate is based on an average derived from the imputation for each game.[16] For these fans, it is assumed that they each pay the ticket price plus 50% of the FCI car parking fee, but no allowance is made for additional stadium related expenditures. For the remaining fans, a price effect is derived from the FCI data. The ticket and car parking price premium associated with a new stadium is calculated on the

[14] ZL, p496. Their data was for the period 1970-2000 and included 20 stadia built since 1956 (a stadium was included if was up to 15 years). This was without including luxury box revenues.

[15] See CH Table 6, p453. The data used by CH covered the period 1969-2001 but included only eighteen stadia, considering only those in the data period that were up to ten years old.

[16] Using saveestimates in the did_imputation Stata routine. For each observation the imputed treatment effect is $\hat{\tau}_{it} = Y_{it} - \hat{Y}_{it} - \varepsilon_{it}$ where $\hat{Y}_{it}$ is the counterfactual estimate derived from the auxiliary model. Individual estimates are not consistent, but given the number of observations per unit, the weighted sums will be, see Borusyak et al (2022) section 4.3.

This article is protected by copyright. All rights reserved.

Author Manuscript

assumption that there is a counterfactual price, which (a) equals the actual price both in seasons earlier than three years from opening and ten years after opening, and (b) follows a linear trend between these start and endpoints. The price premium is then the difference between actual price and proposed counterfactual price. We adopt the conservative assumptions that all new stadium effects are exhausted ten seasons after opening. The revenue effects for 18 stadia opened between 1992 and 2010 are shown in Table 5.

The estimates in Table 5 suggest that there are substantial revenue gains from new stadia, figures that are more in line with ZL than CH. The average revenue increase is $260 million, with almost three quarters of this coming from the additional fans. This estimate looks similar to the $272 million estimate of Hakes and Clapp (2006), although their estimate is a present value in real 2000 dollars.

Table 6 compares these additional revenues to the estimated cost of stadium construction. The average construction cost was $398 million. This implies that the additional revenues covered around two thirds of the construction cost on average. In only five cases did the revenues exceed the costs, and all these were stadia built before the millennium. However, the figures also show that the revenues were generally far in excess of the private contribution to funding stadia, suggesting that public subsidies generally made these stadia hugely profitable for the team owners and their investors. In only five cases were the additional revenues lower than private funding contribution to construction, and in three of those the contribution exceeded 90% of the private cost.

Tables 5 and 6 here

Finally, Table 7 shows the profile of additional revenues by season, aggregated across all 18 new stadia. An important contribution of this paper is the estimation of anticipation effects. The three seasons prior to opening generate 11% of the total estimated revenue gain, which is almost entirely attributable to increased attendance rather than price increases.

Table 7 suggests that the attendance impact in season 9 (i.e. the tenth season since opening) is roughly equal to the impact in the season before the stadium opens. As figure 3 and Table A1 (column 1) show, the impact after year 12 appears to stabilize around a point estimate of around 2-3,000, but that is also insignificantly different from zero. Given the very large confidence intervals, these figures cannot be considered reliable. Nonetheless, it is possible that the revenue estimates reported here are material understatements, if these longer term effects are present.

This article is protected by copyright. All rights reserved.

Author Manuscript

The figures presented in this section provide an impression of the financial impact of new stadia, but there are potentially significant omissions. As mentioned above, there are potential additional revenue sources that could arise if the treatment effects last longer than ten seasons. Luxury boxes and concessions also add significant revenues, as do the naming rights of a new stadium.

Table 7 here

The counterfactual costs of maintaining the old stadium rather than building a new one are even harder to quantify. On the one hand, maintenance costs over any given timespan are likely to be lower simply due to the age of the facility. With discounting, a new stadium also reduces the present value of the cost of the next replacement stadium, since it is pushed further into the future. Advances in stadium technology should also be capable of achieving improvements in labor efficiency and thus reduce labor costs. On the other hand, new stadia that enable owners to charge higher prices by offering better services are also likely to incur additional labor costs. Better technologies might also increase costs in other ways, especially if they involve more maintenance. The figures presented here are not discounted - the present value of the financial benefit would be smaller. The figures are also not adjusted for inflation.

While it would be possible to generate more complex calculations in relation to costs and benefits, this would require the addition of more and more heroic assumptions for which there is limited empirical support. Rather than go down that route, it seems wiser to conclude that the evidence suggests that new stadia generate substantial revenues which would cover a significant fraction, but possibly not the entirety of new stadium costs. Furthermore, it appears that the extent of public subsidies for most stadia opened since 1992 has ensured that the team owners were able to reap substantial profits from their limited investment.[17]

[17] Another approach, not considered here but suggested by an anonymous referee, would be to compare the revenue estimates to a model based on franchise revenue estimates published by Forbes (and previously, Financial World) since 1990. This approach was adopted by Depken (2006) and Bradbury (2019). This would have the advantage of including other potential revenue sources associated with a new stadium such as naming rights and sponsorships. However, it would also exacerbate potential confounding factors, such as local broadcast rights deal, which are significant in MLB. Another problem with the Forbes data is that these estimates are based on an unspecified proprietary method. Audited financial statements are generally not available for MLB franchises, but the Forbes data are sometimes inconsistent with those few audited financial statements which have entered the public domain. For example, the leaked financial statement for the Los Angeles Angel in 2009 shows a total revenue figure of $241 million (downloadable at Rodney Fort's Sports Business Pages:

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

5 Conclusions

This paper uses recent advances in event study methods to estimate the effect of new MLB stadia on attendance over the last half century. Combining attendance data with price data covering the last 30 years, the new methodology allows us to estimate the revenue effects on 18 stadia opened since 1992.

The resulting analysis addresses problems with identification and estimation associated with TWFE models that are commonly used in event studies. The new estimator, suggested by Borusyak et al (2022), Gardner (2021) and Liu et al (2021), is calculated by running an OLS regression, with time and unit fixed effects, on never-treated or not-yet-treated units and then using the difference between the outcome of the target variable for treated observations and the counterfactual prediction to impute the ATT.

The main discovery using the new methodology is that there are significant anticipation effects associated with a new stadium, not previously identified in the literature. The estimated ATT from the opening of the new stadium up until ten years after opening are only in line with the estimates generated in previous studies if one considers the season prior to opening as the baseline—this paper shows that that may constitute a significant omission. With the anticipation effects, the impacts are much larger, and potentially extend much further. Indeed, for our preferred model (the imputation effects in Figure 3), the ATT is significant in the 12 seasons following opening.

One particular advantage of the league sports data used in this paper is that for staggered treatments, it is possible to generate a regression-free unconditional estimate of the control group by constructing a matched sample of home games in the same season for treated and never-treated teams. The results of this comparison (Figure 2) suggest the same pattern of treatment effects as the imputation method.

Annual ticket and car park price estimates available since 1991 make it possible to generate estimates of incremental revenues attributable to new stadia. Adopting a conservative approach, by only considering the first ten seasons after opening, the revenue effects are estimated to be considerably larger than the private investment in construction, but generally below the full cost.

https://drive.google.com/drive/folders/1kKsHCCKL_aJ2gg9ZthNLGT3y0ppgkDMo) compared to the figure of only $217 million as reported by Forbes (https://docs.google.com/spreadsheets/d/1TYZOWtIqtyX6BEEVG6uR-JKtt37G8jgl/edit#gid=1101140078). While there may be explanations for these differences, the lack of transparency is reason for caution.

This article is protected by copyright. All rights reserved.

One issue not addressed in this paper concerns renovations. Most stadia undergo renovations as they age, and these can be quite extensive. This is particularly the case for the control group of franchises in the paper (e.g. The Boston Red Sox and Chicago Cubs). It is possible that renovations to treated team stadia (after the new stadium is opened) could exaggerate the impact of the new stadium itself. This could also help to explain why the TWFE regression estimates are larger than the imputation estimates, since in the former case the estimated regression includes these observations, while the imputation method only uses not-yet-treated observations to impute the value of the counterfactual. Since significant renovations are unlikely in the first decade of a new stadium, this is further justification for restricting the revenue analysis only to that period. Future research might shed further light on this issue.

There are a number of other potential avenues for future research. The analysis could easily be extended to other professional sports leagues. Another important policy question in the stadium literature is the economic impact on cities of new facilities. Team owners and their supporters often claim that a new stadium will enhance economic development in the host city and there is a substantial economic literature that seeks to estimate economic impact (see e.g. Humphreys (2019) and Matheson (2019) for a recent review of the debate). The imputation approach described here would be well suited to an evaluation of stadium impacts on cities hosting major league franchises.

Acknowledgments

I thank Ron Smith, Jan van Ours and two anonymous referees for helpful comments.

Data Availability Statement

The data that support the findings of this study are openly available in Western Economic Association International Data and Code Repository Project openicpsr at https://www.openicpsr.org/openicpsr/project/185802/version/V2/view, reference number 185802.

This article is protected by copyright. All rights reserved.

Author Manuscript

References

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

Baker, A. C., Larcker, D. F., & Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates? Journal of Financial Economics, 144(2), 370-395.

Borusyak, K., 2021. "DID_IMPUTATION: Stata module to perform treatment effect estimation and pre-trend testing in event studies," Statistical Software Components S458957, Boston College Department of Economics, revised 05 Mar 2022.

Borusyak, K., Jaravel, X., 2017. Revisiting event study designs with an Application to the Estimation of the Marginal Propensity to Consume. Working Paper.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. "Revisiting event study designs: Robust and efficient estimation." (2022), https://arxiv.org/pdf/2108.12419.pdf.

Bradbury, John Charles and Coates, Dennis and Humphreys, Brad R., (2022) The Impact of Professional Sports Franchises and Venues on Local Economies: A Comprehensive Survey (January 31, 2022). Journal of Economic Surveys, forthcoming, Available at http://dx.doi.org/10.2139/ssrn.4022547

Bradbury, J.C. (2019). Determinants of revenue in sports leagues: An empirical assessment. Economic Inquiry, 57, 121-140.

Callaway, Brantly, and Pedro HC Sant’Anna. "Difference-in-differences with multiple time periods." Journal of Econometrics 225, no. 2 (2021): 200-230.

de Chaisemartin, Clément, and Xavier d'Haultfoeuille. "Two-way fixed effects estimators with heterogeneous treatment effects." American Economic Review 110, no. 9 (2020): 2964-96.

Clapp, C. M., and Hakes, J. K. “How long a honeymoon? The effect of new stadia on attendance in Major League Baseball.” *Journal of Sports Economics*, 6(3), 2005, 237-263.

Coates, D., and Humphreys, B. R. “Novelty effects of new facilities on attendance at professional sporting events.” *Contemporary Economic Policy*, 23(3), 2005, 436-455.

Depken, C. A. (2006). The impact of new stadia on professional baseball team finances. Public Finance and Management, 6, 436-474.

Gardner, John. "Two-stage differences in differences." arXiv preprint arXiv:2207.05943 (2021).

Garicano, Luis, Ignacio Palacios-Huerta, and Canice Prendergast. "Favoritism under social pressure." Review of Economics and Statistics 87, no. 2 (2005): 208-216.

Gitter, S. R., and Rhoads, T. A. “Stadium construction and minor league baseball attendance.” *Contemporary Economic Policy*, 32(1), 2014, 144-154.

Goodman-Bacon, Andrew. "Difference-in-differences with variation in treatment timing." Journal of Econometrics 225, no. 2 (2021): 254-277.

Hakes, J. K., & Clapp, C. M. (2006). The edifice complex: The economics of public subsidization of major league Baseball Facilities. International Journal of Sport Finance, 1(2), 77-95.

This article is protected by copyright. All rights reserved.

Author Manuscript

Humphreys, Brad R. "Should the construction of new professional sports facilities be subsidized." Journal of Policy Analysis and Management 38, no. 1 (2019): 264-270.

Imbens, Guido W., and Donald B. Rubin. Causal inference in statistics, social, and biomedical sciences. Cambridge University Press, (2015).

Leadley, J. C., and Zygmont, Z. X. “When is the honeymoon over? National Basketball Association attendance 1971-2000.” *Journal of Sports Economics*, 6(2), 2005, 203-221.

Leadley, J. C., and Zygmont, Z. X. “When is the honeymoon over? National Hockey League attendance, 1970-2003.” *Canadian Public Policy/Analyse de Politiques*, 2006, 213-232.

Liu, Licheng, Ye Wang, and Yiqing Xu. "A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data." arXiv preprint arXiv:2107.00856 (2021).

Long, J. G. *Public-private partnerships for major league sports facilities*. New York: Routledge. 2013

McEvoy, C.D., Nagel, M.S., DeSchriver, T.D., and Brown, M.T. (2005). Facility age and attendance in Major League Baseball. Sport Management Review, 8, 19-41.

Matheson, Victor. "Is there a case for subsidizing sports stadia?." Journal of Policy Analysis and Management 38, no. 1 (2019): 271-277.

Noll, R.G. (1974) “Attendance and price setting,” in Noll, Roger G. ed. Government and the Sports Business, 115-157, Washington, DC: The Brookings Institution.

Poitras, M. and Hadley, L. (2006). Do new major league ballparks pay for themselves? Journal of Business, 79, 2275-2299.

Quinn, K. G., Bursik, P. B., Borick, C. P., & Raethz, L. “Do new digs mean more wins? The relationship between a new venue and a professional sports team's competitive success.” *Journal of Sports Economics*, 4(3), 2003, 167-182.

Roth, Jonathan. "Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends." American Economic Review: Insights 4, no. 3 (2022): 305-22.

Soebbing, B. P., Mason, D. S., & Humphreys, B. R. “Novelty effects and sports facilities in smaller cities: Evidence from Canadian hockey arenas.” *Urban Studies*, *53*(8), 2016, 1674-1690.

Sun, Liyang, and Sarah Abraham. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." Journal of Econometrics 225, no. 2 (2021): 175-199.

Szymanski, Stefan. ECIN Replication Package for “Anticipating the Honeymoon: Event Study Estimation of New Stadium Effects in Major League Baseball Using the Imputation Method.” Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2023-04-10. https://doi.org/10.3886/E185802V2

Zygmont, Z. X., & Leadley, J. C. “When is the honeymoon over? Major League Baseball attendance 1970-2000.” *Journal of Sport Management*, 19(3), 2005, 278-299.

Author Manuscript

This article is protected by copyright. All rights reserved.

Author Manuscript

Tables A1, A2 and A3 here

This article is protected by copyright. All rights reserved.

Table 1: Attendance effects associated with the opening of a new stadium, relative to the last season before opening, in various professional sports leagues

Years since opening	Zygmont and Leadley (2005)		Leadley and Zygmont (2005)		Coates and Humphreys (2005)		Coates and Humphreys (2005)		Coates and Humphreys (2005)		Clapp and Hakes (2005)		Leadley and Zygmont (2006)		Gitter and Rhoads (2014)		Gitter and Rhoads (2014)		Gitter and Rhoads (2014)		Soebbing et al (2016)		Simple average	Average of coefficents with pvalues < .05
	MLB 1970-2000		NBA (1971-2000)		MLB (1969-2001)		NFL (1969-2001)		NBA (1969-2001)		MLB (1950-2002)		NHL (1970-3003)		MiLB A (1992-2006)		MiLB AA (1992-2006)		MiLB AAA (1992-2006)		Minor League hockey (1996-2007)			
0	32%		16%	***	28%	***	6%	**	9%	**	44%	***	22%	***	50%	***	22%	**	36%	***	46%	**	27%	28%
1	18%	***	19%	***	21%	***	8%	***	9%	***	27%	***	19%	***	47%	***	16%	***	36%	***	45%	**	22%	24%
2	19%	***	22%	***	19%	***	6%	***	8%	***	20%	***	18%	***	41%	***	19%	**	35%	***	44%	**	21%	23%
3	25%	***	25%	***	14%	**	5%	*	10%	***	19%	***	19%	***	40%	***	11%	***	27%	***	42%	***	19%	23%
4	23%	***	18%	***	15%	**	6%	***	10%	***	19%	***	21%	***	31%	***	8%	***	23%	***	37%	**	17%	19%
5	17%	***	13%	**	12%	**	4%	*	6%	**	15%	***	16%	***	21%	***	3%	***	9%		29%		12%	13%
6	25%	***	11%	*	11%	**	5%	***	5%	*	15%	***	11%	***	21%	***	3%	***	8%		21%		12%	13%
7	21%	***	8%	*	16%	***	2%		5%		10%	**	8%	**	17%	***	-2%	***	7%		15%		9%	12%
8	15%	*	8%		6%		3%		6%	**	10%	**	3%		14%	***	0%	***	6%		-1%		7%	7%
9	3%		8%		8%		1%		9%	***	3%		4%		12%	***	1%	**	5%		-1%		5%	7%
10	1%		2%								3%												2%	0%
11	-1%		6%																				2%	0%
12	-4%		4%																				0%	0%
13	0%		3%																				1%	0%
14	1%		5%																				3%	0%

*** p<0.01, ** p<0.05, * p<0.1

Author Manuscript

This article is protected by copyright. All rights reserved.

Table 2: New stadium construction in MLB

Treatment Group			**Control Group**	
Team	**New stadium opened**		**Team**	**Stadium opened**
Washington Nationals*	1976, 2008		Boston Red Sox	1912
Minnesota Twins	1982, 2010		Chicago Cubs	1916
Toronto Blue Jays	1989		Los Angeles Dodgers	1962
Chicago White Sox	1991		Los Angeles Angels	1966
Baltimore Orioles	1992		Oakland Athletics	1966
Cleveland Guardians	1994		Kansas City Royals	1973
Texas Rangers	1994		Colorado Rockies**	1995
Atlanta Braves	1997, 2017		Tampa Bay Rays**	1998
Seattle Mariners	1999		Arizona Diamondbacks**	1998
Detroit Tigers	2000			
Houston Astros	2000			
San Francisco Giants	2000			
Milwaukee Brewers	2001			
Pittsburgh Pirates	2001			
Cincinnati Reds	2003			
Philadelphia Phillies	2004			
San Diego Padres	2004			
St. Louis Cardinals	2006			
New York Yankees	2009			
New York Mets	2009			
Miami Marlins**	2012			

*Montreal Expos until 2004

**Expansion franchise in sample period

Author Manuscript

This article is protected by copyright. All rights reserved.

Table 3: Home and visiting teams

		Visiting team	
		Treated	*Never treated*
Home team	*Treated*	57860	22661
	Never treated	22589	7053

Author Manuscript

This article is protected by copyright. All rights reserved.

Table 4: The gap in years and seasons spanned between breaking ground on a new stadium and opening.

Franchise	Season opened	Broke ground	Difference (years)	Seasons spanned between breaking ground and opening
Toronto Blue Jays	1989	1/1/1986	3.25	3
Chicago White Sox	1991	5/1/1989	1.92	2
Baltimore Orioles	1992	6/1/1989	2.84	3
Cleveland Guardians	1994	1/1/1992	2.25	2
Texas Rangers	1994	4/1/1992	2.00	2
Atlanta Braves	1997	7/10/1993	3.73	4
Seattle Mariners	1999	3/1/1997	2.08	2
Houston Astros	2000	11/1/1997	2.42	2
San Francisco Giants	2000	12/1/1997	2.33	2
Detroit Tigers	2000	10/1/1997	2.50	3
Pittsburgh Pirates	2001	4/1/1999	2.00	2
Milwaukee Brewers	2001	11/1/1996	4.42	4
Cincinnati Reds	2003	8/1/2000	2.67	3
Philadelphia Phillies	2004	4/6/2001	2.99	3
San Diego Padres	2004	5/3/2000	3.92	3
St. Louis Cardinals	2006	1/17/2004	2.21	2
Washington Nationals	2008	5/4/2006	1.91	2
New York Mets	2009	11/13/2006	2.38	2
New York Yankees	2009	8/19/2006	2.62	3
Minnesota Twins	2010	8/1/2007	2.67	3
Miami Marlins	2012	7/1/2009	2.75	3
Atlanta Braves	2017	9/16/2014	2.54	3

Author Manuscript

This article is protected by copyright. All rights reserved.

Table 5. Revenue increases attributable to new stadia

Franchise	Stadium opened	Revenue attributable to increased attendance $m	Revenue attributable to increased prices for existing fans $m	Sum of estimated additional revenue $m	% revenue from increased attendance
Baltimore Orioles	1992	339	42	381	89%
Cleveland Guardians	1994	363	48	411	88%
Texas Rangers	1994	155	53	208	74%
Atlanta Braves	1997	185	80	265	70%
Seattle Mariners	1999	385	80	465	83%
Detroit Tigers	2000	108	25	132	81%
Houston Astros	2000	184	40	224	82%
San Francisco Giants	2000	403	21	423	95%
Milwaukee Brewers	2001	158	24	182	87%
Pittsburgh Pirates	2001	99	70	169	59%
Cincinnati Reds	2003	-81	63	-19	N/A
Philadelphia Phillies	2004	151	84	234	64%
San Diego Padres	2004	56	110	166	34%
St. Louis Cardinals	2006	234	15	249	94%
Washington Nationals	2008	81	-2	79	103%
New York Mets	2009	-49	74	25	N/A
New York Yankees	2009	310	412	722	43%
Minnesota Twins	2010	280	87	367	76%
Average		187	74	260	76%

Author Manuscript

This article is protected by copyright. All rights reserved.

Table 6: Comparison of additional revenues and cost

Franchise	Stadium opened	Sum of estimated additional revenue $m	Construction cost $m	Private funding contribution $m	Additional revenue as multiple of private funding
Baltimore Orioles	1992	381	106	9	42.36
Cleveland Guardians	1994	411	176	92	4.46
Texas Rangers	1994	208	147	12	17.36
Atlanta Braves	1997	265	235	235	1.13
Seattle Mariners	1999	465	517	126	3.69
Detroit Tigers	2000	132	310	177	0.75
Houston Astros	2000	224	265	85	2.64
San Francisco Giants	2000	423	324	309	1.37
Milwaukee Brewers	2001	182	392	102	1.79
Pittsburgh Pirates	2001	169	271	75	2.26
Cincinnati Reds	2003	-19	280	30	-0.62
Philadelphia Phillies	2004	234	458	229	1.02
San Diego Padres	2004	166	483	134	1.24
St. Louis Cardinals	2006	249	365	120	2.07
Washington Nationals	2008	79	524	81	0.98
New York Mets	2009	25	575	434	0.06
New York Yankees	2009	722	1308	1015	0.71
Minnesota Twins	2010	367	435	128	2.87
Average		260	398	188.5	1.38

Source of cost data: Long (2013). Chapter 2, pp 19-24.

Author Manuscript

This article is protected by copyright. All rights reserved.

Table 7: additional revenues by season imputed to 18 new stadia opened from 1992 to 2010.

0 = season stadium opened	Revenue attributable to increased attendance $m	Revenue attributable to increased prices for existing fans $m	Sum of estimated additional revenue $m
-3	110.2	0.0	110.2
-2	123.6	20.4	144.0
-1	199.6	54.2	253.8
0	431.7	253.7	685.4
1	309.7	219.0	528.6
2	263.2	156.4	419.5
3	258.9	150.7	409.5
4	256.7	121.3	378.1
5	255.0	119.1	374.1
6	244.1	84.6	328.7
7	274.8	74.5	349.3
8	250.8	41.9	292.8
9	216.1	32.5	248.6
10	165.1	-1.4	163.6
Total	**3359.7**	**1326.8**	**4686.4**

Author Manuscript

This article is protected by copyright. All rights reserved.

Table A1. ATT and pre-trend coefficients for six models summarized in Figure 6

	full sample imputation	full sample TWFE	matched sample imputation	matched sample TWFE	season data imputation	season data TWFE
T-10	-7	486	231	151	-1178	444
	(1167)	(2168)	(1405)	(1112)	(1112)	(1873)
T-9	-29	-575	-705	-765	-495	-776
	(1191)	(2141)	(1407)	(1082)	(1253)	(1887)
T-8	1790	1350	2070	1995	634	366
	(1345)	(2191)	(1588)	(1230)	(1217)	(1815)
T-7	1987	1543	1658	1682	1727	1541
	(1343)	(2103)	(1614)	(1170)	(1556)	(1864)
T-6	1712	1669	787	1129	1320	1566
	(1333)	(1854)	(1611)	(1274)	(1431)	(1514)
T-5	2249	2101	1448	1907	1593	1670
	(1417)	(1767)	(1709)	(1407)	(1292)	(1420)
T-4	3101*	2859	1799	2157	2236	2239
	(1588)	(2040)	(2100)	(1701)	(1385)	(1641)
T-3	2897***	3792*	2709***	3610**	3914***	4463**
	(655)	(2115)	(811)	(1751)	(613)	(1701)
T-2	2698***	3581*	2805***	3725**	3376***	4041**
	(693)	(2024)	(833)	(1662)	(661)	(1728)
T-1	4407***	5483***	3949***	5016***	4327***	5092***
	(723)	(1967)	(859)	(1401)	(652)	(1677)
T0	10689***	11933***	9845***	11226***	11154***	11941***
	(783)	(1895)	(922)	(1523)	(658)	(1557)
T1	8242***	9601***	8464***	9403***	8282***	9147***
	(965)	(2050)	(1101)	(1602)	(900)	(1702)
T2	6992***	8487***	7309***	8666***	7530***	8535***
	(1006)	(2081)	(1211)	(1770)	(960)	(1916)
T3	7279***	8314***	7300***	8152***	7206***	8043***
	(1010)	(2243)	(1212)	(2095)	(907)	(2051)
T4	6933***	8326***	6943***	8286***	7178***	8224***
	(978)	(2229)	(1094)	(2066)	(878)	(1898)
T5	6142***	7756***	5821***	7109***	5980***	7024***
	(1182)	(2018)	(1339)	(1937)	(1186)	(1806)
T6	6154***	7560***	6205***	7339***	6064***	7316***
	(1200)	(2286)	(1333)	(2112)	(1155)	(2089)
T7	5922***	7318***	4951***	6338***	5940***	6984***
	(1108)	(2033)	(1301)	(1991)	(1122)	(2010)
T8	5004***	6285***	3964***	5385***	5051***	5841***
	(1215)	(1995)	(1393)	(1863)	(1140)	(1905)
T9	4022***	5727***	3387**	5023***	4656***	6036***
	(1289)	(1698)	(1394)	(1604)	(1237)	(1635)
T10	3276***	5043***	2116	3942**	3287***	4653***
	(1257)	(1622)	(1379)	(1556)	(1234)	(1628)
T11	3433***	4924***	2373	3839**	3458***	4805***
	(1333)	(1517)	(1497)	(1566)	(1309)	(1496)
T12	2507*	3784**	1903	3234*	3660**	5029***
	(1525)	(1625)	(1646)	(1684)	(1601)	(1466)
T13	2930*	4363***	1488	3135*	2580*	3993**
	(1509)	(1367)	(1813)	(1762)	(1502)	(1501)
T14	889	2051	-1015	632	2164	3007***
	(1523)	(1496)	(1666)	(1594)	(1616)	(1348)
T15	2000	3497***	574	2425	2565*	3625**
	(1447)	(1135)	(1661)	(1571)	(1447)	(1340)
T16	2189	3668***	1104	2574*	3220*	4508***
	(1472)	(1010)	(1633)	(1346)	(1682)	(976)
T17	2152	3025**	1413	2582	2538	3184**
	(1954)	(1225)	(2025)	(1682)	(1666)	(1354)
T18	3376	4079***	1859	2897	3341*	3797**
	(2049)	(1398)	(2144)	(1945)	(1872)	(1103)
T19	1438	2394***	547	1863	2426	2995**
	(1771)	(935)	(1882)	(1620)	(1926)	(882)
>T19	-2773	-1763	-3387*	-2585	-821	-159
	(1714)	(1513)	(2037)	(1862)	(2214)	(2091)
Observations	104,256	107,064	43,092	44,109	1340	1375

This article is protected by copyright. All rights reserved.

Table A2: Covariate estimates for six models summarized in Figure 6

	full sample imputation	full sample TWFE	matched sample imputation	matched sample TWFE	season data imputation	season data TWFE
Home team wpc to date	25033***	24984***	23549***	23258***		
	(1512)	(1344)	(1865)	(1562)		
Visiting team wpc to date	13731***	12267***	12916***	9981***		
	(998)	(610)	(1305)	(1001)		
wpc					28691***	27867***
					(2527)	(2411)
1st game of season	40029***	38648***	35911***	34296***		
	(2318)	(1697)	(2490)	(1829)		
Holiday	5251***	4347***	6043***	5290***		
	(648)	(541)	(774)	(629)		
March	-703	-3002**	1110	-1201		
	(2553)	(1478)	(3137)	(2245)		
April	-4991***	-5195***	-4670***	-4649***		
	(575)	(437)	(630)	(478)		
May	-4280	-4256***	-3921***	-3868***		
	(417)	(280)	(469)	(353)		
June	-1213.932***	-1419***	-1275***	-1365***		
	(191)	(137)	(245)	(214)		
August	-980***	-1037***	-923***	-909***		
	(220)	(212)	(256)	(194)		
September	-5502***	-4788***	-4956***	-4454***		
	(503)	(310)	(586)	(409)		
October	-5572***	-4399***	-4700***	-3666***		
	(820)	(587)	(1068)	(813)		
MSA per capita income (nominal)	0.0983	0.1407	0.0256	0.0649	0.0841	0.1506
	(0.1613)	(0.1151)	(0.1525)	(0.0880)	(0.1514)	(0.1110)
MSA population	0.0020**	0.0010**	0.0015*	0.0012**	0.0014**	0.0007*
	(0.0008)	(0.0005)	(0.0009)	(0.0005)	(0.0007)	(0.0004)
weekend	6967***	6563***	6675***	6387***		
	(382)	(294)	(462)	(347)		
night game	1513***	1171***	1925***	1501***		
	(411)	(352)	(415)	(354)		
inter-league game	1257**	1529***	1271	1788***		
	(570)	(271)	(873)	(491)		
temporary stadium	-6679***	-5284***	-3338	-2744		
	(1889)	(1504)	(2741)	(2379)		
wpc > .500 last season	3049***	3786***	2945***	3430***	1641***	2071***
	(454)	(401)	(482)	(394)	(344)	(320)
in postseason last season	1110***	969***	944***	332	2898***	3574***
	(173)	(134)	(332)	(219)	(419)	(540)
WS runnerup last season	1077***	860***	1053**	707*	2223**	1575
	(331)	(204)	(480)	(406)	(1093)	(1022)
WS winner last season	1960***	1884***	1822***	1703***	2497***	2844***
	(317)	(230)	(416)	(371)	(668)	(694)
new franchise year 2	18115***	16973***	18734***	17414***	19022***	17890***
	(4219)	(4055)	(4621)	(4508)	(4340)	(4270)
new franchise year 3	14689***	12975***	15284***	13444***	13329***	11425***
	(3111)	(3456)	(3248)	(3362)	(2998)	(3079)
new franchise year 4	11789***	10179***	13379***	11723***	9177***	7117***
	(3156)	(2827)	(3603)	(3419)	(2363)	(2071)
new franchise year 5	12072***	10823***	13002***	11488***	11790***	10130***
	(3042)	(2584)	(3575)	(3120)	(3391)	(3213)
new franchise year 6	8555***	8220***	9007***	8301***	6574**	6074**
	(2785)	(2562)	(2948)	(2635)	(3100)	(3073)
new franchise year 7	4791*	4888**	4728	4554*	5174**	5066**
	(2527)	(2330)	(3053)	(2566)	(2522)	(2223)
new franchise year 8	6569**	6522***	8230***	7706***	6646***	6263***
	(2801)	(2107)	(3161)	(2429)	(2587)	(2202)
Observations	104,256	107,064	43,092	44,109	1340	1375

This article is protected by copyright. All rights reserved.

Table A3: Robustness Tests

	full sample imputation (base model)	R1	R2	R3	R4
T-10	-7	315	470	-1004	-7
	(1167)	(1152)	(1177)	(1325)	(1169)
T-9	-29	-9	23	-637	30
	(1191)	(1219)	(1264)	(1326)	(1198)
T-8	1790	1753	1942	1083	1752
	(1345)	(1408)	(1461)	(1515)	(1343)
T-7	1987	1930	2153	1646	2039
	(1343)	(1359)	(1370)	(1421)	(1347)
T-6	1712	2046	1784	1333	1733
	(1333)	(1280)	(1347)	(1558)	(1320)
T-5	2249	2662*	2386	2158	2274
	(1417)	(1420)	(1524)	(1616)	(1411)
T-4	3101*	3248**	3334*	3540**	3056*
	(1588)	(1571)	(1776)	(1792)	(1591)
T-3	2897***	3064***	2493***	3496***	2892***
	(655)	(666)	(712)	(821)	(664)
T-2	2698***	2885***	2803***	3633***	2697***
	(693)	(712)	(730)	(890)	(702)
T-1	4407***	4748***	4856***	5583***	4299***
	(723)	(720)	(758)	(962)	(734)
T0	10689***	10578***	10681***	12110***	10584***
	(783)	(767)	(782)	(1079)	(775)
T1	8242***	8409***	8410***	9176***	7448***
	(965)	(996)	(1079)	(1300)	(937)
T2	6992***	7175***	7435***	7671***	6436***
	(1006)	(1052)	(1128)	(1373)	(1028)
T3	7279***	7488***	7568***	7720***	6518***
	(1010)	(1063)	(1105)	(1371)	(996)
T4	6933***	7432***	7273***	7693***	6288***
	(978)	(1004)	(1083)	(1349)	(979)
T5	6142***	6285***	5940***	7084***	5647***
	(1182)	(1145)	(1230)	(1535)	(1144)
T6	6154***	6757***	6531***	6081***	5585***
	(1200)	(1198)	(1336)	(1628)	(1090)
T7	5922***	6621***	5928***	6835***	5410***
	(1108)	(1111)	(1226)	(1423)	(1000)
T8	5004***	4998***	4918***	6832***	4566***
	(1215)	(1218)	(1333)	(1561)	(1171)
T9	4022***	3992***	3759***	5623***	3801***
	(1289)	(1286)	(1383)	(1749)	(1323)
T10	3276***	3365***	3275**	5541***	2846**
	(1257)	(1254)	(1378)	(1759)	(1244)

This article is protected by copyright. All rights reserved.

T11	3433***	3271**	2750*	5853***	2880**
	(1333)	(1348)	(1565)	(1767)	(1272)
T12	2507*	2313	2052	5184***	2030
	(1525)	(1529)	(1662)	(1929)	(1510)
T13	2930*	2732*	2342	5929***	2345
	(1509)	(1524)	(1713)	(1809)	(1503)
T14	889	645	423	3841**	644
	(1523)	(1507)	(1713)	(1944)	(1532)
T15	2000	1775	1644	4615**	1792
	(1447)	(1378)	(1537)	(1805)	(1445)
T16	2189	2065	1134	4023**	2207
	(1472)	(1449)	(1676)	(1646)	(1494)
T17	2152	2174	2168	5184**	2032
	(1954)	(1938)	(2155)	(2336)	(1915)
T18	3376*	3494*	3083	6445***	3086
	(2049)	(2110)	(2136)	(2208)	(2018)
T19	1438	1602	1239	4694**	1484
	(1771)	(1781)	(1837)	(1878)	(1794)
>T19	-2773	-2469	-2889	1731	-2395
	(1714)	(1832)	(1929)	(1682)	(1734)
obs	105701	90096	83228	105701	101010

Note: T0 is the season in which the stadium opened. ATTs include anticipation effects for three seasons prior to opening.

R1: Excludes 1990s expansion franchises (Miami, Colorado, Tampa Bay and Arizona)

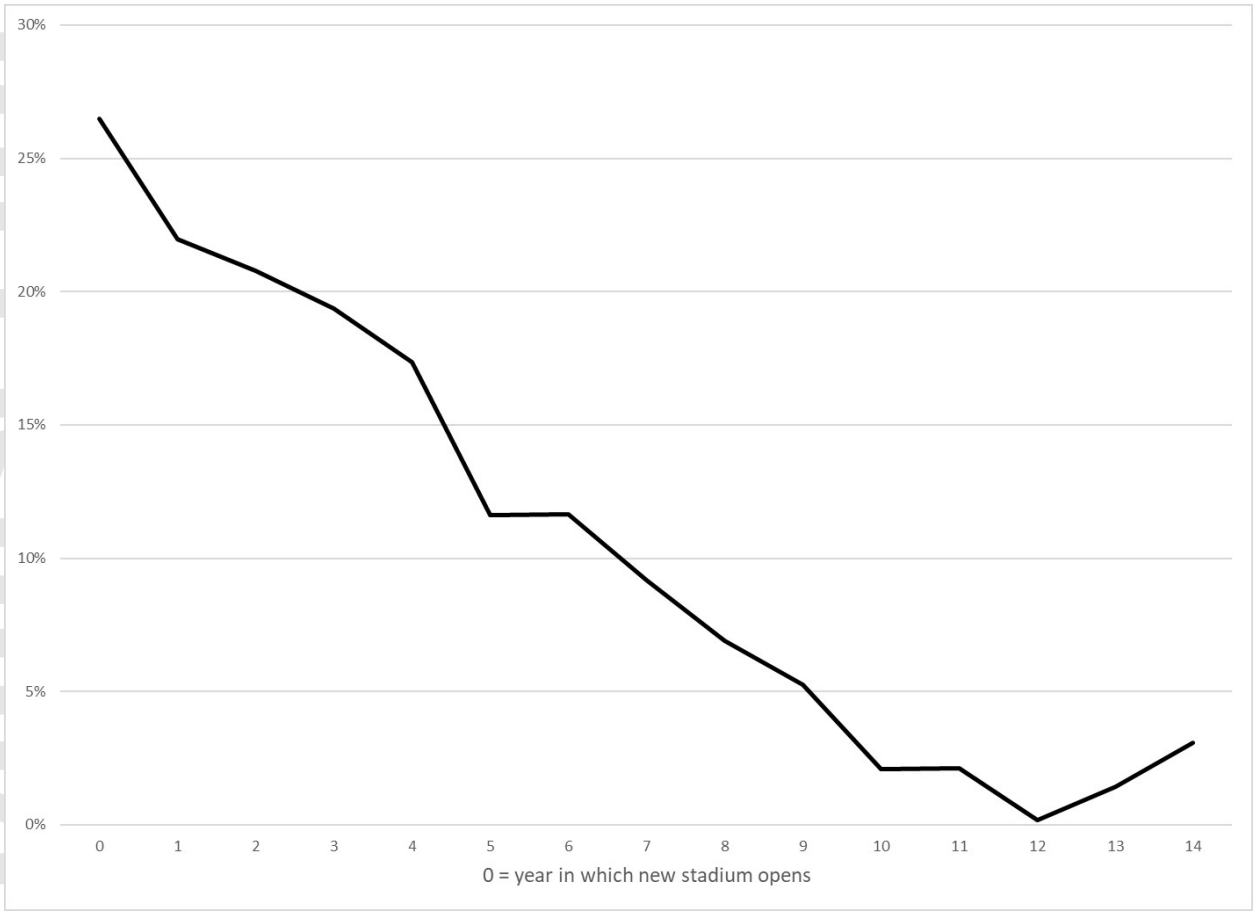
R2: As R1 and excludes Montreal/Washington franchise

R3: Excludes all stadiums built before 1992

R4: drops all cases where capacity utilization exceeded 100% (about 2% of all games).

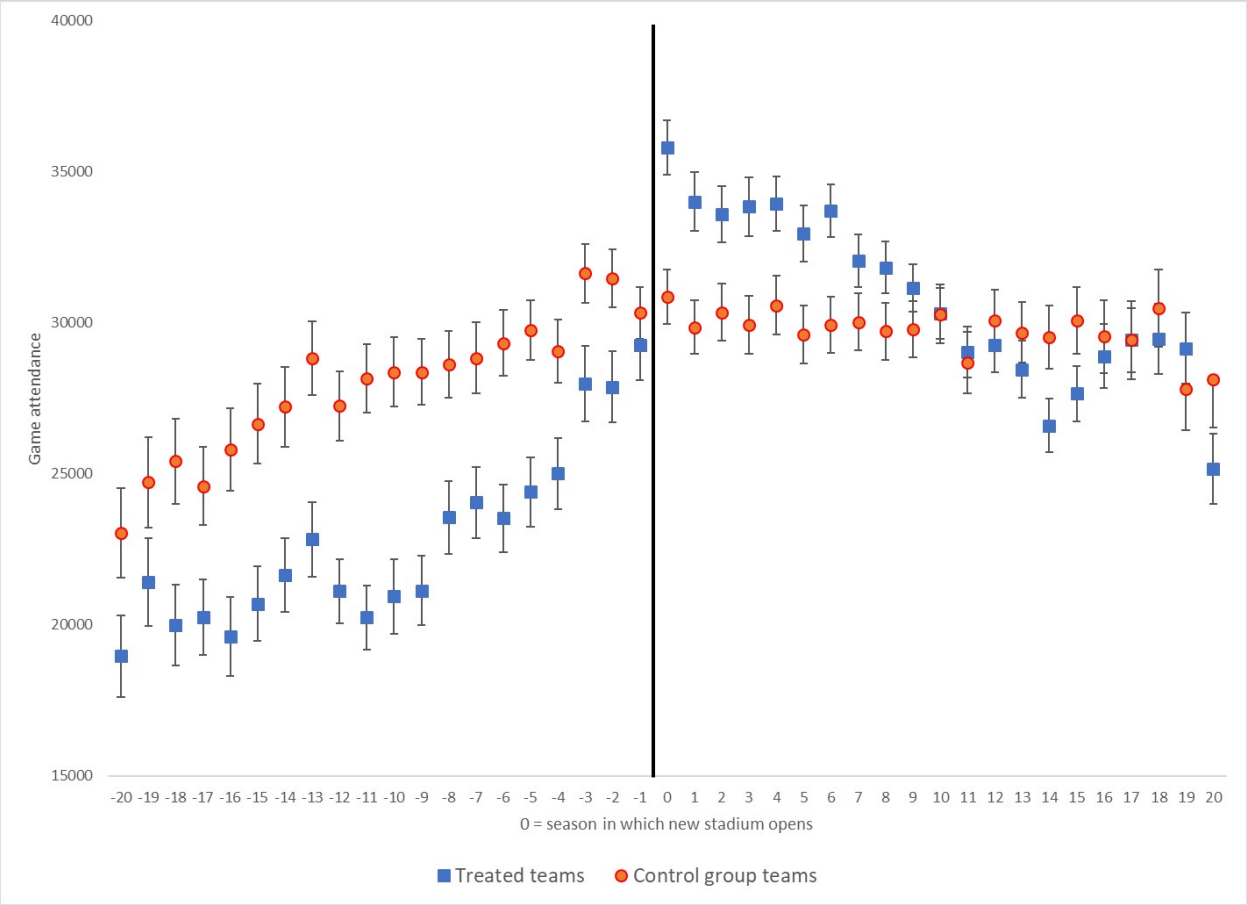
This article is protected by copyright. All rights reserved.

Author Manuscript



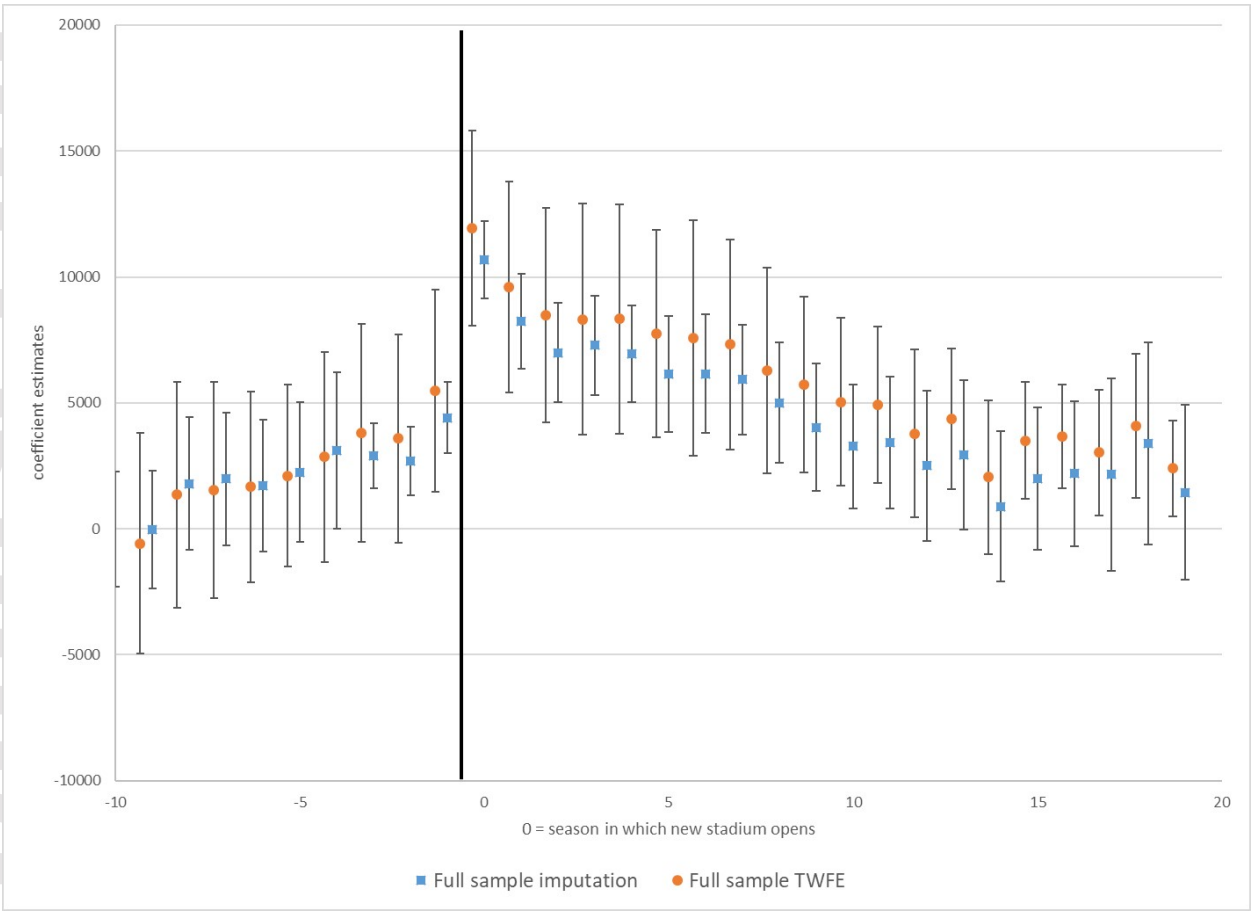
This article is protected by copyright. All rights reserved.

Author Manuscript



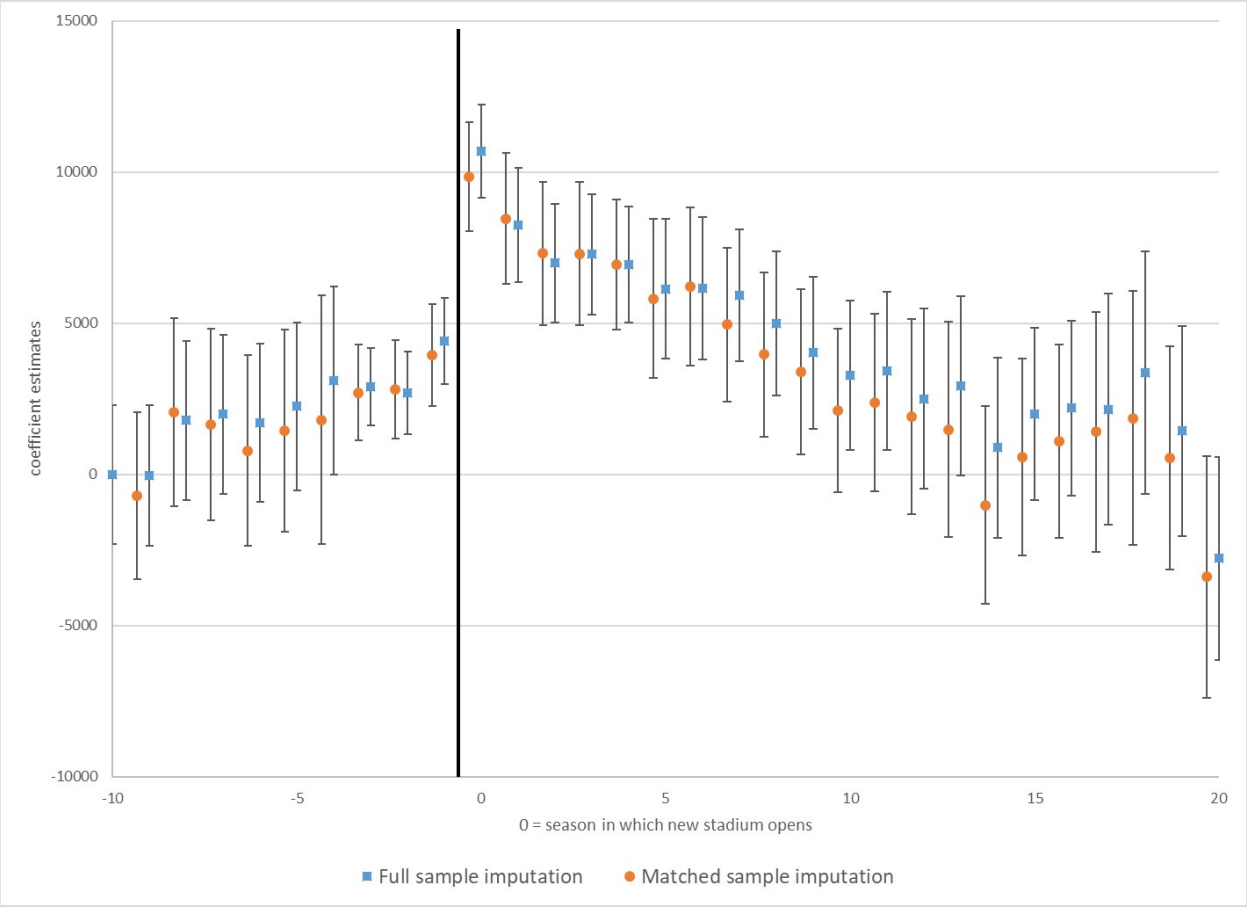
This article is protected by copyright. All rights reserved.

Author Manuscript



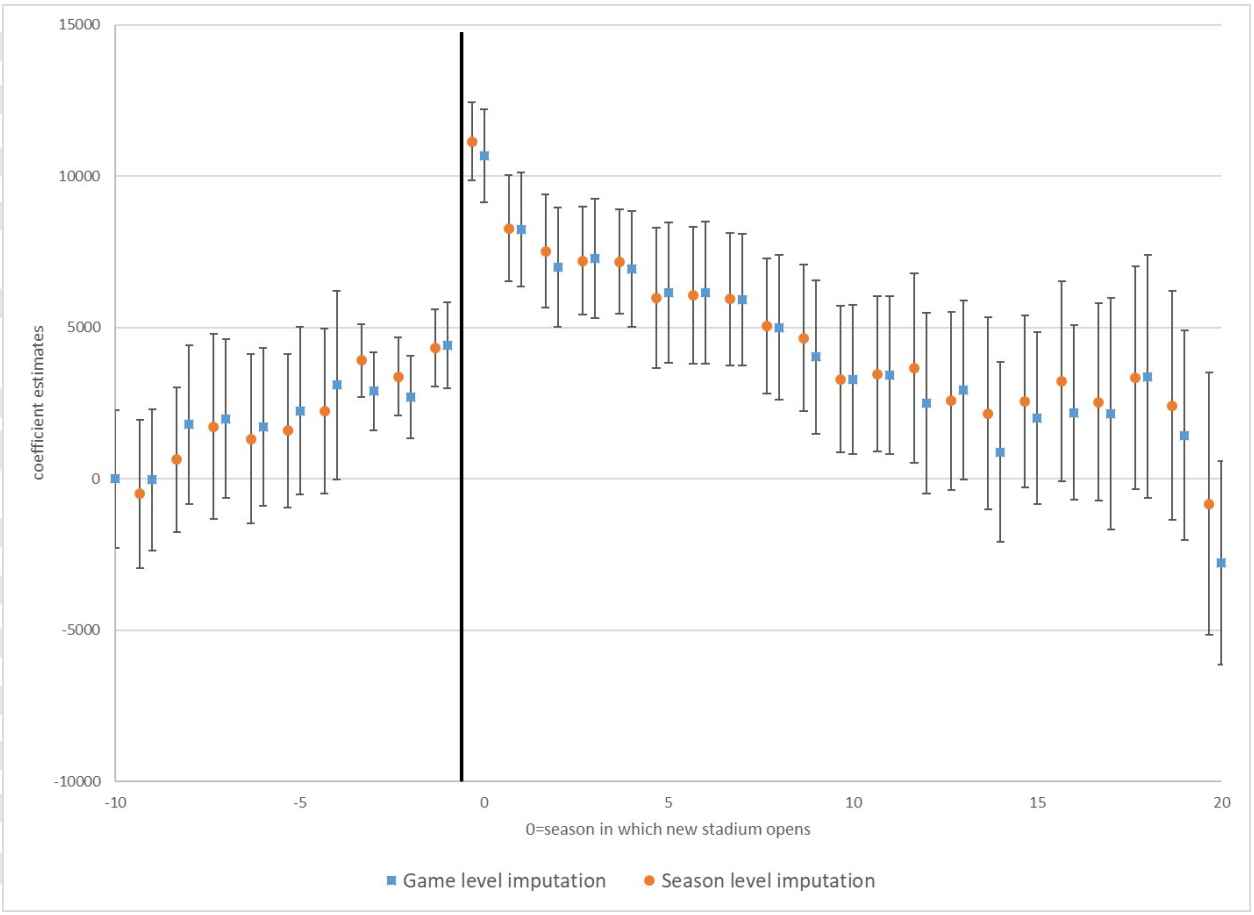
This article is protected by copyright. All rights reserved.

Author Manuscript



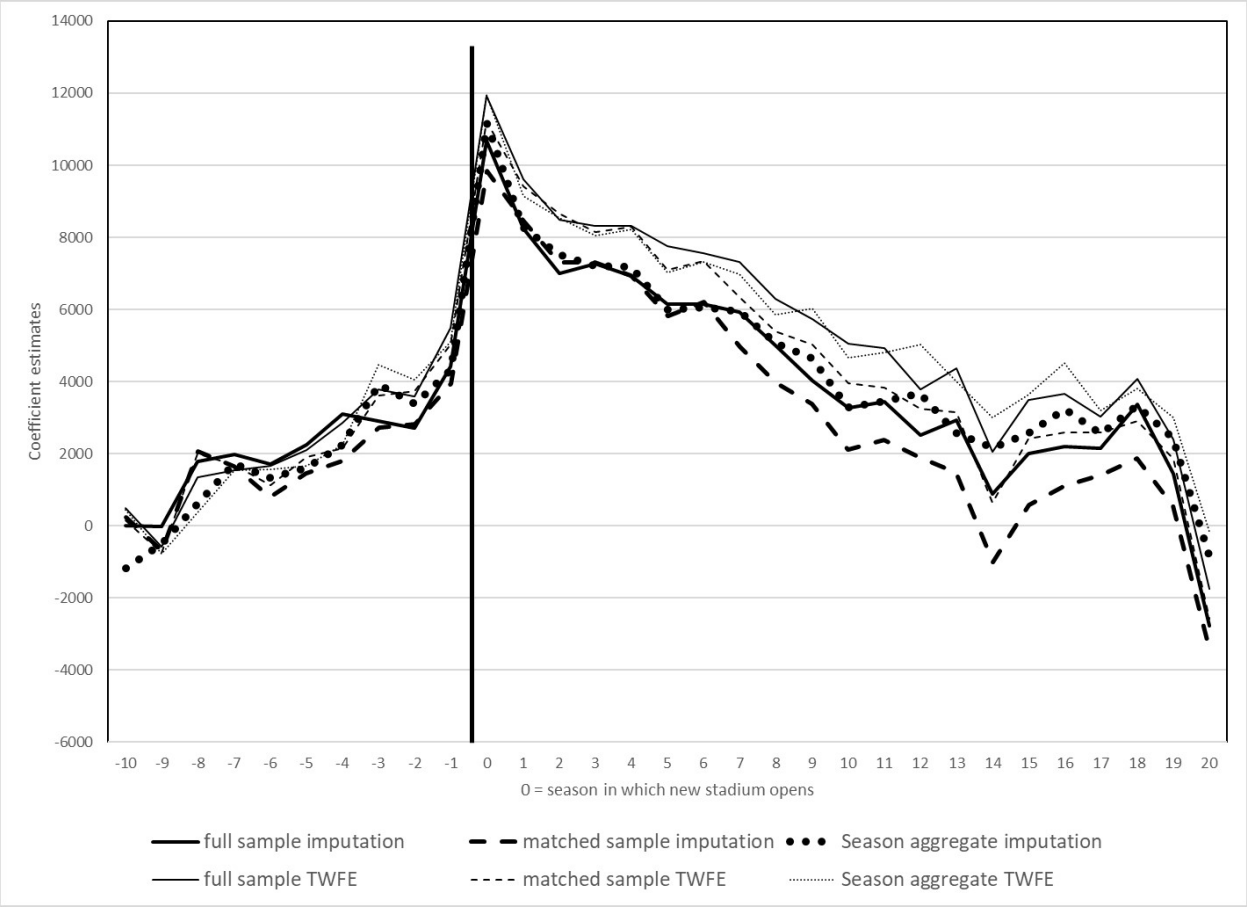
This article is protected by copyright. All rights reserved.

Author Manuscript



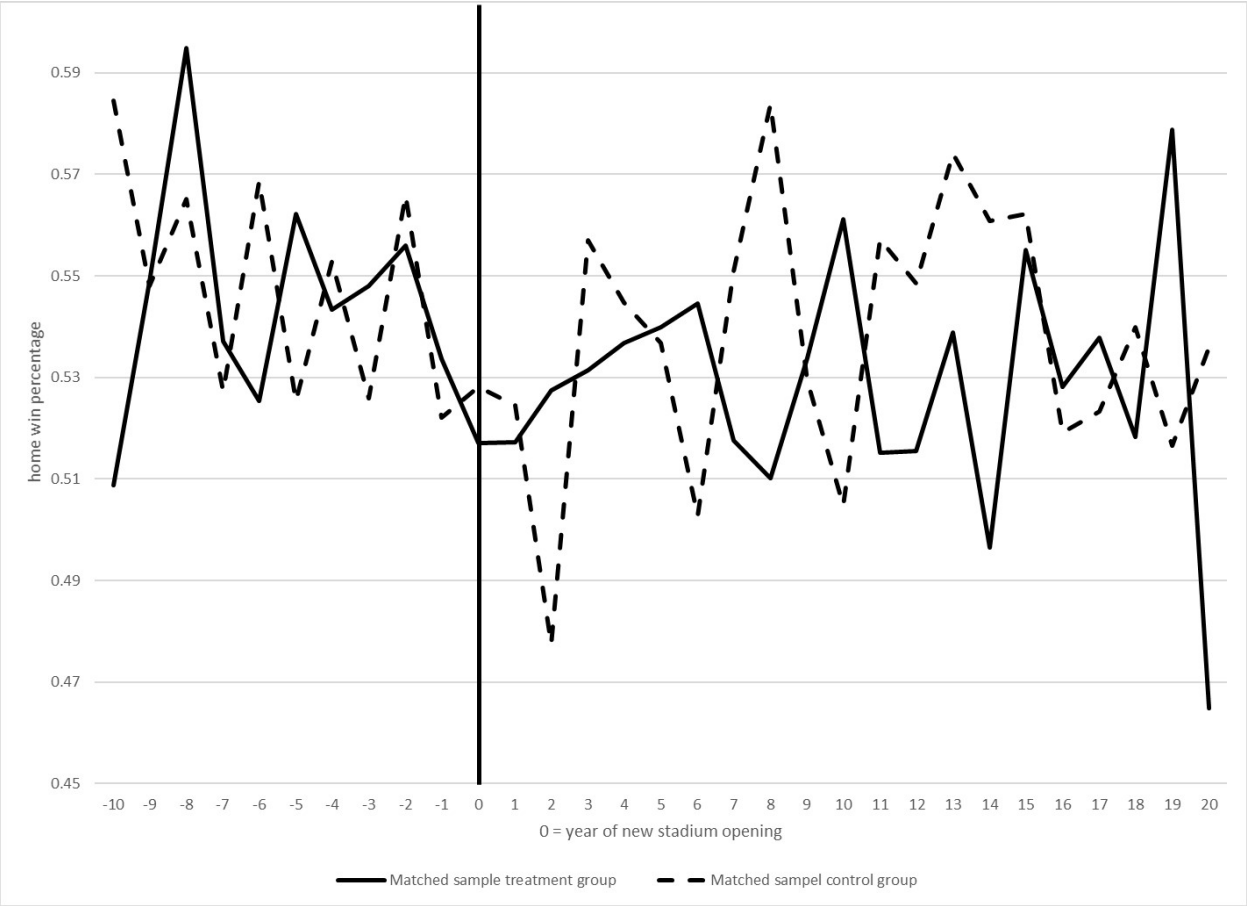
This article is protected by copyright. All rights reserved.

Author Manuscript



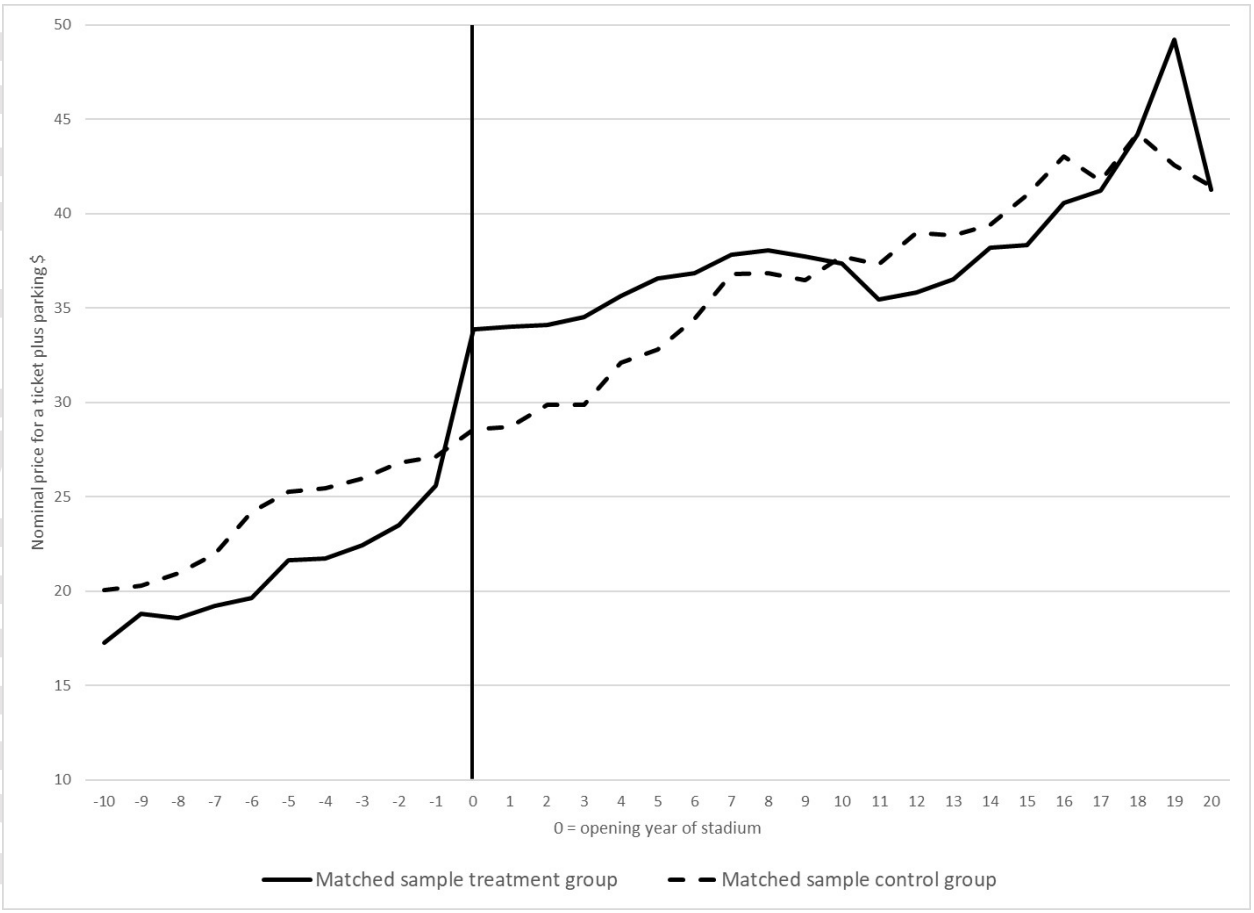
This article is protected by copyright. All rights reserved.

Author Manuscript



This article is protected by copyright. All rights reserved.

Author Manuscript



This article is protected by copyright. All rights reserved.