

# GEM FINDER

**‘Speed Above Replacement Jockey’ Metric Can Uncover Underappreciated Riders**

---

By PAUL VON HIPPEL

**O**VER THE PAST 30 years, sport after sport has revised the way it measures athletic performance—replacing older statistics with new metrics that more accurately identify which players can help a team win. Many teams, managers, and coaches that adopted new performance metrics early enjoyed an edge in identifying, recruiting, and deploying top contributors overlooked by their rivals.

The most famous example, introduced to a mass audience by the book and movie “Moneyball,” was the Oakland Athletics’ baseball team that in 2002-03 won nearly as many games as the New York Yankees, despite having paid its players only a third as much. The A’s spent money more efficiently than the Yankees by ignoring players who excelled based



ANNE M. EBERHARDT



**Jockey Joel Rosario rates well by any measure, topping a new performance-based rating system in 2020 and earning the Eclipse Award as outstanding rider for 2021**

on traditional statistics. Instead, the A's identified undervalued talent using newer measures that better predicted whether a player could help the team win. A similar revolution in statistical talent assessment has transformed basketball and is making inroads in both American football and soccer.

Although new performance measures come in many flavors, they all have one thing in common: They try to isolate the contributions of individual athletes to their team's success. Many older statistics failed to do that. An example of an older statistic is a baseball pitcher's win-loss record. This statistic fails to separate the pitcher's contribution from the contributions of his teammates and the quality of the opposition. A mediocre pitcher can win a respectable number of games if he plays for a strong team in a weak division. An outstanding pitcher might struggle to win if his teammates can't back him up with effective hitting and fielding.

By contrast, wins above replacement player (WAR or WARP) is a newer statistic that estimates how many more (or fewer) games a team wins with a certain player than it would win with an average player. A player can have a high WARP despite playing for a losing team, and WARP makes its contribution visible to other teams looking for talent. Similar WARP statistics have been developed for basketball and American football.



While purse earnings declined for jockey Rafael Bejarano in 2020, his impressive Speed Above Replacement Jockey rating suggests that the fall is more likely from decreased opportunity than diminished performance

### MONEYBALL FOR JOCKEYS?

Applying this approach to racing, I've developed a measure I call Speed Above Replacement Jockey (SARJ). Inspired by WARP, SARJ estimates how much faster (or slower) horses run under a certain jockey compared with how

they would run under an average jockey. Like WARP, SARJ puts contributions of individual jockeys under a brighter, more focused light. Some big-name jockeys, it turns out, are getting top mounts in rich races despite not adding much speed while other jockeys that add a lot of speed are

## Leading North American jockeys by earnings, 2020

RANK	JOCKEY NAME	STARTS	1ST	2ND	3RD	TOTAL \$	PER START \$	WIN %	TOP 3	TOP 3%
1	Irad Ortiz, Jr.	1,266	300	225	185	\$21,050,686	\$16,628	24%	710	56%
2	Joel Rosario	1,052	194	205	155	\$18,229,097	\$17,328	18%	554	53%
3	Luis Saez	1,398	268	224	182	\$16,511,332	\$11,811	19%	674	48%
4	Tyler Gaffalione	1,408	271	225	183	\$15,614,762	\$11,090	19%	679	48%
5	John R. Velazquez	662	101	104	94	\$15,250,515	\$23,037	15%	299	45%
6	Florent Geroux	861	160	137	92	\$14,778,813	\$17,165	19%	389	45%
7	Jose L. Ortiz	1,027	190	181	136	\$13,811,296	\$13,448	19%	507	49%
8	Ricardo Santana, Jr.	1,057	174	166	127	\$12,550,253	\$11,873	16%	467	44%
9	Javier Castellano	640	107	105	109	\$12,498,470	\$19,529	17%	321	50%
10	Manuel Franco	1,011	140	120	163	\$12,302,732	\$12,169	14%	423	42%

THIS PAGE: KEENELAND/COADY PHOTOGRAPHY; OPPOSITE PAGE, TOP: MICHAEL BURNS; BOTTOM: COADY PHOTOGRAPHY

riding ordinary horses at out-of-the-way tracks.

Before I explain SARJ, let's look at how we're rating jockeys now.

The previous page lists 2020's leading North American jockeys according to Equibase. By default, the list ranks jockeys according to the total prize money won by their mounts. But this is a poor proxy for jockey effectiveness. Total prize money depends strongly on a jockey's opportunities: how many races in which the jockey starts, whether those races have big purses, and whether the jockey gets to ride favorites. Total prize money is strongly correlated with the number of races in which a jockey starts. It is no coincidence that among the top 10 jockeys for prize money in 2020, seven also ranked among the top 15 for starts.

The Equibase list can be re-sorted to rank riders by the prize money earned per start, or by the percentage of races where the jockey finished in the money. These statistics are better because they adjust for how many races a rider started. But they're still limited measures of a jockey's contribution.

Looking at the usual jockey rankings, we can tell which jockeys are riding a lot and getting good results. But we can't tell whether a jockey is getting better results than another jockey could if given the same opportunities. None of the common statistics separate the jockey's opportunities from what he or she makes of them.

SARJ compares jockeys by looking at instances where different jockeys have ridden the same horse. If horses perform better for Jockey 1, on average, than the same horses do for Jockey 2, then Jockey 1 will have a higher SARJ than Jockey 2.

That might sound a little limited, but it turns out that SARJ can compare jockeys even if they've never ridden the same horse. To take the above explanation further, we can directly compare Jockey 2 to Jockey 3 because they've both ridden Horse 2. And we can directly compare Jockeys 3, 4, and 5 because



**According to Speed Above Replacement Jockey ratings, connections that call on veteran jockey Huber Villa-Gomez (above) often see an improvement in their horses; as is the case with Eric Cancel (left) on the New York Racing Association circuit**

they've all ridden Horse 4. Although we can't directly compare Jockeys 4 and 5 to Jockey 2—because they've never ridden the same horse—we can compare them indirectly by comparing them all to Jockey 3, who has horses in common with each. For example, if Jockey 2 has a higher SARJ than Jockey 3, and Jockey 3 has a higher SARJ than Jockey 4, then Jockey 2 must have a higher

SARJ than Jockey 4. This approach allows each jockey to be compared with jockeys that he or she shares horses with, as well as to jockeys that those jockeys share horses with. And so on, and so on. In 2020, fully 98% of jockeys could be compared to each other by this approach.

To compare horses' speed under different jockeys, Equibase Speed Fig-

ures are among those used. They're designed to compare race performances across different distances, tracks, and conditions.

In terms of these speed figures, the difference between first and second is usually five points or less. The difference between second and third typically is four points or less. It follows that a jockey who adds four or five speed points, on average, can routinely improve a horse's placing.

As it turns out, though, less than 10% of active jockeys can add more than four speed points to their mounts.

In calculating SARJ ratings for 1,134 jockeys who raced in 2020, several adjustments for other factors that might affect the horse's speed figure were made. SARJ adjusts for the weight the horse carried so that apprentice jockeys don't get credit for competing with less weight or jockeys riding the high-weight in handicap races aren't penalized for the added burden.

Also, adjustments were made for the class of a race because horses generally run faster when there are faster horses around them. I didn't want to give a jockey undue credit for that, nor did I want to penalize jockeys who were racing strong horses against weak competition and not asking much from their mounts.

Finally, there also are accommodations by track. While the speed figures attempt to account for differences among tracks, I didn't want to lean too heavily on this variant and unduly reward or penalize riders.

## TOP 10 FROM SARJ ANGLE

Let's look again at the 10 jockeys who won the most prize money in 2020—the jockeys who appeared at the top of the Equibase list for 2020. The table at the top of this page estimates their SARJ scores, along with a 95% confidence interval—a wider interval gap suggests wider variance and lists the range.

SARJ confirms that these riders are generally excellent. Eight of the

## SARJ ratings of top 10 jockeys by prize money

NAME	ALL RACES	STAKES RACES	SARJ	95% CONFIDENCE INTERVAL
Joel Rosario	1,008	227	6.0	(4.0, 7.9)
Irad Ortiz Jr	1,239	194	5.1	(3.2, 7.1)
John R Velazquez	644	179	5.1	(3.2, 7.1)
Manuel Franco	980	158	4.8	(2.8, 6.7)
Jose L Ortiz	985	199	4.5	(2.5, 6.4)
Tyler Gaffalione	1,372	165	4.3	(2.3, 6.3)
Javier Castellano	620	163	4.2	(2.3, 6.2)
Luis Saez	1,359	169	4.2	(2.3, 6.2)
Ricardo Santana Jr	1,040	159	3.7	(1.7, 5.6)
Florent Geroux	836	169	3.2	(1.2, 5.1)

## Top SARJ ratings among jockeys who started at least 50 stakes races in 2020

NAME	ALL RACES	STAKES RACES	SARJ	95% CONFIDENCE INTERVAL
Rafael Bejarano	637	63	6.2	(4.3, 8.2)
Joel Rosario*	1,008	227	6.0	(4.0, 7.9)
Eric Cancel	806	56	5.3	(3.3, 7.3)
Junior Alvarado	666	116	5.2	(3.2, 7.2)
Irad Ortiz Jr*	1,239	194	5.1	(3.2, 7.1)
John R Velazquez*	644	179	5.1	(3.2, 7.1)
Mike E Smith	263	97	4.8	(2.9, 6.8)
Jose Lezcano	757	111	4.8	(2.8, 6.7)
Manuel Franco*	980	158	4.8	(2.8, 6.7)
Flavien Prat	812	172	4.7	(2.7, 6.6)
Kendrick Carmouche	775	91	4.5	(2.5, 6.5)
Jose L Ortiz*	985	199	4.5	(2.5, 6.4)
David Cabrera	1,082	51	4.4	(2.5, 6.4)
Tyler Gaffalione*	1,372	165	4.3	(2.3, 6.3)
Javier Castellano*	620	163	4.2	(2.3, 6.2)

\*These jockeys also appear among the top 10 for prize money.

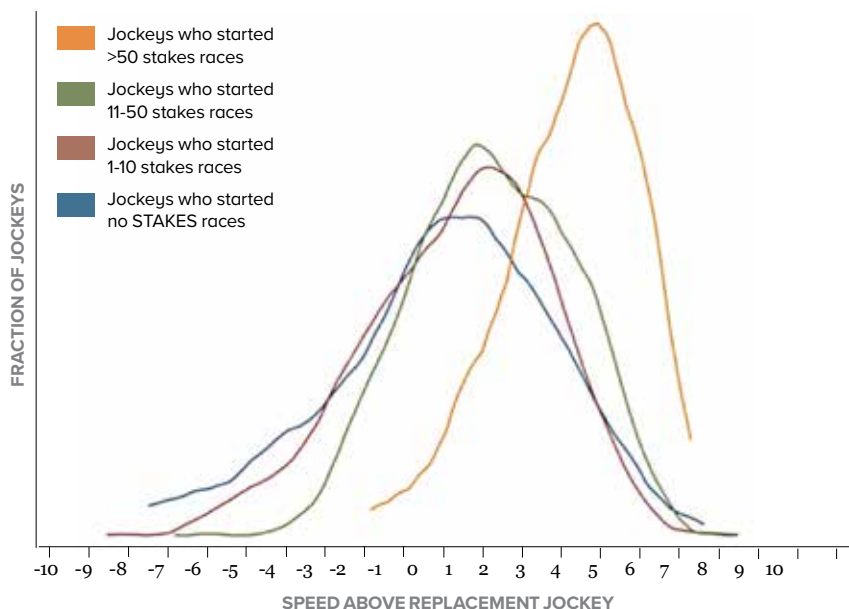
10 have SARJ estimates north of 4, meaning they could usually win on a horse that an average jockey would race to second. And one, Joel Rosario, has a SARJ rating of 6, meaning that in a mile race he would typically finish three lengths ahead of an average jockey riding the same horse. These jockeys deserve the opportunities they're getting. But they're not the only ones worthy of top horses.

Above is a list of the top 15 SARJ

ratings among jockeys who started at least 50 stakes races in 2020. Half were also among the top 10 earners in the previous table. All made good money, and, as frequent riders in stakes, are reasonably well known. But there definitely are some mild surprises on the list—jockeys who had great speed above replacement that didn't get the same opportunities.

At the top of the list is Rafael Bejarano. A multiple grade 1 winner,

### SARJ RATINGS, BY JOCKEY ACCOMPLISHMENT LEVEL



### Top SARJ ratings among jockeys who started at least 200 races in 2020

NAME	ALL RACES	STAKES RACES	SARJ	95% CONFIDENCE INTERVAL
Huber Villa-Gomez	379	7	6.6	(4.7, 8.6)
Rafael Bejarano	637	63	6.2	(4.3, 8.2)
Joel Rosario	1008	227	6.0	(4.0, 7.9)
Juan C Diaz	603	18	5.7	(3.7, 7.6)
Eric Cancel	806	56	5.3	(3.3, 7.3)
Junior Alvarado	666	116	5.2	(3.2, 7.2)
Irada Ortiz Jr	1239	194	5.1	(3.2, 7.1)
John R Velazquez	644	179	5.1	(3.2, 7.1)
Francisco Arrieta	696	36	5.1	(3.1, 7.0)
Hector I Berrios	581	24	5.0	(3.0, 6.9)

Bejarano has had an illustrious career that includes a pair of placings in the Belmont Stakes (G1) in 2005 and 2006 and five Breeders' Cup victories from 2005-2014. When looking only at Bejarano's earnings, it is easy to get the impression that his performance level is in decline. In 2020 he earned only a quarter as much prize money as in 2012.

Yet according to SARJ, Bejarano isn't over the hill. To the contrary, he is still as effective as Rosario.

Some might also be surprised to

see that Eric Cancel, who started only 56 stakes races, had about the same SARJ rating as Irada Ortiz Jr., John Velazquez, or Mike Smith. Or that Kendrick Carmouche, who started in 91 stakes races, was about as effective as Flavien Prat or Jose Ortiz, who started in twice as many.

As one would expect, most top riders find their way to the top races. On average, jockeys who started more than 50 stakes races in 2020 had SARJ scores 2 points higher than jockeys who started 11-50 stakes. Likewise,

jockeys who started 10-50 stakes races had average SARJ scores 1 point higher than jockeys who started fewer than 10 stakes races. And, jockeys who started in 1-10 stakes races had average SARJ scores about a half-point higher than jockeys who started in no stakes at all.

But those are average differences. As the graph at left shows, there was substantial overlap among the distribution of SARJ scores for jockeys who started a lot of stakes races and jockeys who started in few or none. A small number of jockeys who started more than 50 stakes races were actually no better, according to SARJ, than an average jockey. Among jockeys who started fewer than 50 stakes races, there were a few who had SARJ ratings as good as the top prize money winners.

### DIAMONDS IN THE ROUGH

One of the strengths of SARJ is it can look at riders at all class levels. The 2020 ratings suggest that some of the most effective jockeys are not riding in top races. Who were these diamonds in the rough?

The table at left gives the top SARJ figures among jockeys who started 200 or more races in 2020, whether any of them were stakes races or not. Some are jockeys whom we saw in previous tables, but some are less familiar names.

At the top of the list is veteran Huber Villa-Gomez. Although Villa-Gomez has not won as much prize money as Bejarano, he has enjoyed his share of success. But as with Bejarano, his earnings suggest a career on a downward trajectory. By 2020 he was earning only a third of what he had earned in 2011, competing at tracks such as Tampa Bay Downs, Presque Isle Downs, and Delaware Park. Yet his SARJ figure suggests Villa-Gomez is still one of the top riders in the country—just as effective in 2020 as Bejarano and Rosario.



Just as new statistics in baseball and other sports have helped point to undervalued players, SARJ is capable of finding undervalued jockeys; the ratings suggest Francisco Arrieta is among North America's top riders



SARJ ratings suggest some talented women riders, such as Gulfstream Park-based Fanny Olsson, would be more productive if given additional opportunities

Based on SARJ ratings, other lesser-known jockeys who deserve a second look include Francisco Arrieta (currently competing at Oaklawn Park), Juan C. Diaz (mainly races in Puerto Rico), and Hector I. Berrios (who favors Gulfstream Park)

Beyond these riders, is it possible that jockeys whom by choice or circumstance are riding part-time also deserve further opportunity? This last table gives the top five SARJ figures among jockeys who started 50-200 races in 2020. All these jockeys had SARJ figures north of 5.

Among this group, three are women. Lauralea Glaser, who raced in the 2015 Prix Longines Future Racing Stars at Chantilly, has had starts and stops in her career while working regularly as an exercise rider at Laurel Park. Nakia Ramirez, active since 2013, has done

### *Top SARJ figures among jockeys who started 50-200 races in 2020*

NAME	ALL RACES	STAKES RACES	SARJ	95% CONFIDENCE INTERVAL
Lauralea Glaser	72	6	7.4	(5.5, 9.4)
Nakia Ramirez	85	13	7.4	(5.5, 9.4)
Jorge J Urdaneta	96	0	6.0	(4.0, 7.9)
Fanny Olsson	125	3	5.6	(3.6, 7.5)
A A Marin	71	0	5.2	(3.3, 7.2)
Hector I Berrios	581	24	5.0	(3.0, 6.9)

most of her racing near her native Idaho. Swedish jockey Fanny Olsson, who won more than 100 races in Scandinavia, started riding at Gulfstream Park in late 2018 in an effort to break onto the American circuit.

With rare exceptions, it has been difficult for talented women to get

a toehold in Thoroughbred racing. Traditional stats might have assisted in prolonging any biases because if female riders aren't getting quality horses to ride, their wins and earnings are going to lag. Through the lens of SARJ, clearly there are some top female riders who deserve more opportunity. **BH**

LEFT: COADY PHOTOGRAPHY; RIGHT: COGLIANESE PHOTOS/LAUREN KING