



# EMBRACING A FUTURE WITH FREE RACING DATA

Published: March 11, 2019

This is the fourth in a series of foundational white papers published by the Thoroughbred Idea Foundation (TIF).

We believe the North American thoroughbred industry can reverse more than a decade of decline and stagnance through a series of actions which will inspire participation from horseplayers.

A positive environment for bettors yields a healthier sport for horse owners. Horseplayers and owners are the only optional financial participants in the industry – without them, racing would be a shell of itself, a club sport.

To date, TIF has published three white papers on the topics of pricing, transparency and product development. The recommendations associated with these topics include: shifting to a penny breakage

model, adopting Category 1 interference rules and enhancing steward communications with the public, and transforming racing's wagering product to allow for fixed-odds and exchange betting.

This fourth paper addresses racing's data, and specifically, serves as an appeal to Equibase, its board and management.

## THE IDEA

Comments from a 1995 conference were particularly prescient given the fast, data driven times over the next two-plus decades.

*"It may not be apparent to all of us sitting*

*here today, but there is a revolution going on. It's not a political one, or a military one, but a technological one. And like any revolution it will have an enormous impact on our lives and on our businesses.*

*"There are few certainties except one. And that is, when the revolution is over, racing will be changed forever. What is this revolution I'm talking about? It's the information revolution..."*

**"The information revolution is coming. It won't go away. We can't ignore it; it will sweep us up just the same. We can fight it, build a moat, man the battlements, and dig in for a long siege. Or we can join it and get busy with the hard work necessary to put it to our advantage."**<sup>1</sup>

What might surprise you is that those remarks came from then Equibase CEO Richard LeBer at the Jockey Club's 1995 Round Table.

LeBer, a former McKinsey consultant and Harvard Business School graduate, had a vision that could have led the sport down a much different path. He served as Equibase boss for nearly two years before moving on to a very successful career beyond racing.

*"[F]ormed in 1991 to provide the Thoroughbred racetracks of North America with a uniform, industry-owned database of racing information and statistics,"*<sup>2</sup> Equibase serves as a partnership between the Thoroughbred Racing Associations of North America (TRA) and The Jockey Club. Dividends accrued by Equibase are split between the two groups in a 67%/33% fashion, respectively.

A major concern at the time Equibase was created was that racing lacked control over its own data. Equibase changed that. **The service they have provided to the industry in the subsequent three decades has been extraordinary, a point that should not go underreported.**

But, times change. And this paper addresses what we believe is the need for a dramatic shift in the business of racing's data.

This is the era of analytics, particularly in sports, and racing should take every step possible to be a part of it. While the suggestions outlined in this paper may seem a significant departure from the status quo, we believe they are well within reason given the state of the industry.

The sooner racing and Equibase embrace this path forward, the better.

Data has been big business – for Equibase and its partners. We believe liberalizing access to the data – making it far more flexibly available, and at a significantly changed price point – will drive better participation. Enabling data to be both affordable and accessible in a developer-friendly format

will inspire bettors, including those not already investing in racing's markets. Access to basic, free past performances will inspire existing, long-time customers and mass market new entrants.

**The current cost of racing data, and the inflexible formats in which it is provided, serve as hurricane-force headwinds in today's wagering environment.**

Add to this potent mix the rise of legal sports betting. Data which can inform sports bettors' decisions is widely available, in many instances at little to no cost, and offered in formats designed for the tech-savvy.

*"For the cost of roughly 20 days of the Daily Racing Form, I can buy access to enough incredibly detailed, highly accurate, and developer-friendly information to inform me for an entire baseball season, and have enough left over for a few coffees at Starbucks,"* says long-time professional bettor Maury Wolff.

Garret Skiba, an analytics and financial professional, is also one of the most successful live-money contest players in thoroughbred racing history, with tournament cashes in excess of \$800,000. He suggests that comparing racing data in its current form to traditional sports is a futile exercise.

*"The products out there for other sports, for free, make it so easy to take interest in it as a bettor. The most popular is fantasy football, and almost all of the data is completely free, downloadable, or 'scrapable.' To try and compare it to racing is impossible.*

*"The absolute ideal scenario going forward would be to open the raw data to the public and let them have at it to drive interest and innovation."*

Equibase – as a going interest of the racetracks and the Jockey Club – should eschew their interests to profit from data sales, absorbing the costs of data collection and distribution in favor of the wagering participation, and the trickle-down industry benefits, this would yield. In other words, the collection and distribution of racing data should be considered a marketing expense, used to attract and retain gamblers while also empowering owners to make educated decisions. It should be as free and open as possible. Equibase's success should not be measured in terms of data sales, but in the performance of racing's wagering markets.

Tom Reynolds spent more than three decades as a sales and marketing executive for Pepsi-Cola North America, and has significantly increased his interest in racing over time, both as a horseplayer and an owner through Eclipse Thoroughbred Partners. He is also a TIF board member.

*"The racing business survives by selling two experiences – one to owners and one to*

*horseplayers. The exorbitant cost of data in an era where data drives participation and decision-making is a barrier to both experiences," says Reynolds. "The status quo inhibits the growth that the sport needs. The path forward should be crystal clear."*

**In this paper, we recommend that Equibase should:**

- **provide free, raw data feeds for registered, non-commercial users;**
- **provide free, basic past performances on all North American tracks;**
- **provide responsive channels to regularly address errors and omissions in the data;**
- **partner with universities to study racing data, developing new and advanced metrics for the betterment of the sport.**

Make no mistake – the sale of premium products and commercial relationships with firms responsible for such products with proprietary analysis are important for a robust racing information market. They have a very viable place in the market, both for Equibase's own premium products and those outlets such as the Daily Racing Form, Brisnet, TimeformUS, Thorograph and others.

To frame these recommendations, we will examine the history of Equibase, some case studies beyond North America and present a vision for the future of racing, wagering and data on this continent.

## THE ROOTS OF EQUIBASE: "SERVING THE BEST INTERESTS OF THE INDUSTRY"

Then the Vice Chairman of The Jockey Club, William S. Farish offered the Club's annual update of activities at the 1992 Round Table on Matters Pertaining to Racing.

*"Before Equibase was formed, Thoroughbred racing stood out alone as just about the only major professional sport which was not responsible for its own records. These are the records on which so many of our decisions are based.*

*"They govern the eligibility of horses to run in our races. They are the basis for wagering decisions made by our fans; for buying and selling decisions made by our owners; and for mating decisions made by our breeders.*

*"The Thoroughbred industry has the responsibility and obligation to maintain control of those records, and make sure they are made widely available in whatever way suits the best interests of the industry."<sup>13</sup>*

Without question, this challenge has been well-met. Racing took control of its data and has done so for the best part of three decades. The market, however, has changed substantially.

In 1992, total handle was reported by The Jockey Club as \$9.632 billion. Adjusted for inflation in 2018, that amount is worth approximately \$17.421 billion.

Total handle in America in 2018 was just \$11.26 billion.

It is TIF's assertion that in the present day, the best interests of the industry would be to make the data more widely available than has ever occurred previously – in modern formats and at no cost to non-commercial interests. In other words, if you are using the data for your own private use to govern betting decisions or create betting programs – then have at it.

The history of Equibase has been well-reported in the transcripts of the Jockey Club's annual Round Table Conferences, outlining thoughts on how the organization should serve the industry.

A going concern at the time Equibase was formed was that the only database for racing information was controlled by the Daily Racing Form – a privately owned publication which had changed hands on numerous occasions. Now, with an industry-owned site for racing information, the future would be more secure, more sustainable.

Equibase became profitable, feeding dividends to its partners over time. Occasionally, these dividend payments were publicized. Racing information became a cash source for track partners and The Jockey Club.

In 2009, the racetrack partners of Equibase shared a dividend of more than \$3.5 million, according to then Jockey Club President and Chief Operating Officer Alan Marzelli, who offered, at that year's Round Table Conference, some unique insight to the internal perspective of Equibase's position.

*"It is not typical for technology companies to be owned by a tax exempt, non-profit parent.*

*"And it is also not typical for technology companies to remain in the hands of the same owner, as opposed to being sold every three to five years or eventually going public.*

*"But by operating these highly profitable companies for the long term and solely for the benefit of the industry rather than equity partners or shareholders, we are able to provide the capital – financial, technological and human – that is necessary to support many of the worthwhile initiatives that you will hear about throughout the rest of today's program."<sup>14</sup>*

The Jockey Club would have earned more than \$1 million in dividends from Equibase that year given the structure of the partnership. For perspective,

The Jockey Club reported in 2008 that total dividend distributions to the Equibase partners since 1998 was \$24.6 million.<sup>5</sup>

While the initiatives referenced by Marzelli are undoubtedly beneficial to the sport, funded in part by the share of dividends that reverted to The Jockey Club, TIF suggests that the current state of wagering in thoroughbred racing requires re-examination of the overall business model.

Are Equibase's track partners benefiting more from the cash dividends they receive than the revenue from increases in handle they might expect once data is more broadly accessible?

Is the state of the larger sports gambling industry, and the availability of data in other sports, suggestive of a need to engage present day customers differently than the approach of the last three decades?

Despite the generally pitiful performance of the horse racing wagering markets in North America over the last 15 years, should the cost of data remain as exorbitantly high as it is currently? Even while the price of wagering on racing is itself higher than the competition?

Could racing spark more participation from current customers and any participation from new customers by using data as a marketing hook?

## INSIGHT FROM ABROAD

Occasional efforts are made to trigger horseplayer participation in select wagering pools with free data. The multi-track Stronach 5, a low-takeout pick-five sequence has offered players free past performances since it the bet was re-launched several months ago. Brisnet traditionally offers similar when future pools for the Kentucky Derby and Kentucky Oaks come around each season. But the "bread-and-butter," organized data for this continent's roughly 40,000 races is otherwise vaulted.

Unlimited access to "basic" past performances via Equibase for a year runs roughly \$400.

For premium, value-added products, racing data is significantly more expensive. An annual subscription to access an unlimited selection of the Daily Racing Form's "Classic" past performances runs \$1,299. For the same unfettered access to STATS Race Lens, a product offered exclusively via Equibase, the annual cost is \$999.

Equibase packages their various products into a series of day passes – designed to give customers all of the data they might need in one bundle. The cost for the most comprehensive of these – a "gold" day pass – is listed at \$30/day, though the

published average retail value is noted at \$500.

Extrapolated over the course of a racing year, a gold day pass would run more than \$10,000 (the retail value would be an even more alarming \$182,000).

Data in North American racing is big business. But should it be?

As wagering on racing has experienced significant declines, how can the industry-led data monopoly justify current pricing levels and continue its restrictive access to said data?

*"There is an infatuation to pricing racing data, but that sort of model seems to ignore the fundamental business of the sport," says Gary Crispe, Chief Executive Officer of Racing And Sports Pty. Ltd. Crispe's Australian-based website offers comprehensive racing data, including free past performances, for more than 15 countries.*

**"Data and its derivatives should be used to drive betting,"** Crispe adds.

*"Outside of a few relationships with some big clients, nearly all of the data we supply is free to our site's visitors. We offer full form guides [past performances] for races in 17 jurisdictions around the world, which includes speed maps [pace projections] and a whole host of value-added services. They come in a variety of formats, some of which can be tailored to the site user's preferences.*

*"Our model is to offer it for free and we make money on the relationships we have – advertising, affiliate programs and marketing partnerships with different racing jurisdictions. We have some great relationships promoting racing in Hong Kong, Korea, Singapore. The Asian jurisdictions, in particular, recognize the importance of getting as much data as possible into the hands of bettors. That is a driving force behind their overall success."*

The Hong Kong Jockey Club (HKJC) offers free, printable past performances on its 88 live race days and 23 simulcast race days each season. The volume of free information on their website is nearly overwhelming, including veterinary records, full workout details and much more.

General comparisons between Hong Kong and the highly fragmented American market are normally quite unfair – Hong Kong's insular, fully-controlled marketplace enables this level of detail.

But when pari-mutuel commingling opened on its races, allowing customers from foreign jurisdictions to bet into its massive tote pools, the HKJC recognized that data in a style familiar to its foreign customers was a key driver to encourage participation.

*"We are investing in an effort to tailor make information in the form and factor which overseas customers are used to looking at,"* said the HKJC's

Richard Cheung, Executive Director of Customer and International Business Development, at the Asian Racing Conference in May 2018.

*"Our commingling business started in 2014 and by the end of 2015/16, we were at about US\$440 million in turnover... [by the end of] this season 2017/18, the business has already grown four to five times, to US\$2 billion."*<sup>6</sup>

Despite already offering past performances in their own style, the HKJC pays to have the form translated to a style familiar to American bettors. Those are offered free online, via a partnership with the Daily Racing Form. The HKJC does the same form for the Racing Post in the U.K., and provides Chinese content to major Chinese-language outlets in Canada and Australia.



For Australia's Gary Crispe, though, the processing power present in today's market suggests organizations such as RacingAndSports.com need to continually evolve in order to offer its customers a free service that keeps them coming back.

*"We have built up so much data that we are now looking into the next steps, what's next for our business and how can we stay ahead of the curve and offer what the next generation really wants. We are applying machine learning to all of this – take years of history and all of the form data offered by a single race and run 100,000 simulations of it. Besides those results, you end up with some really accurate predictive pricing models that today's bettors are looking for."*

## RECOMMENDATIONS

Dramatic action should be considered to spark interest in racing wagering.

Performance over the last several years has

generally equated to treading water, while 2018 saw a small, but pleasing, spike, likely due to favorable federal tax law changes.

The early returns through the first two months of 2019, however, are not encouraging. Comparing the first two months of 2018 to those of 2019, handle is down 4% according to researcher Chris Rossi, while total races run was up by more than 1%.<sup>7</sup>

The longer-term figures, addressed earlier, present far greater challenges. The legalization of sports betting could be a benefit to racing provided actions are taken to present the sport's wagering opportunities in a fashion familiar to fixed-odds sports bettors. Much is to be done in this arena, as outlined in the TIF's white paper on the topic – Racing and 'Legal' Sports Betting.

TIF has derived several recommendations we believe would benefit the industry through increased wagering.

**First, Equibase should offer free data feeds to individual users who register to receive them, provided such use is for their own non-commercial purposes.**

Admittedly, this is a substantial shift from the current standard and the one that excites us the most when pondering the possibilities.

Horse racing needs customers who want to delve into its voluminous data, creating their own programs and metrics. The outcome of such an investment is almost impossible to predict, because this dramatic liberalization of racing data would be a first.

More than 3,500 people attended the MIT Sloan Sports Analytics Conference this month in Boston. Topics on the agenda included analytics development in all the major sports one would expect. And then, there were sessions on poker, curling and even Esports analytics.

University students presented compelling research papers and posters with their findings,<sup>8</sup> in many cases, with data sourced through free or inexpensive channels. Teams and leagues from around the world descend on Boston to source talent for their ever-growing analytics departments. But racing goes, essentially, unrepresented.

Offering free past performances serves as a token to racing's current customers. Taking a giant leap forward, free data feeds to talented researchers and programmers could yield a treasure trove of information for the sport and participation from new actors attracted to racing's robust data.

Chair of the Rhodes College economics department, Marshall Gramm, is also the co-founder of Ten Strike Racing, a syndicate responsible for nearly 1,000 starts and more than 200 wins with earnings in excess of \$6.5 million in just six years.

Gramm is a highly accomplished horseplayer, most recently recording a fifth-placing in the 2018 Breeders' Cup Betting Challenge and ninth in the 2019 National Horseplayers' Championship.

As a researcher and economist, the Equibase status quo bewilders Gramm.

*"Despite being one of the most data intensive sports imaginable, the lack of publicly available, well-formatted and affordable data is probably one of American racing's greatest head-scratchers," says Gramm.*

*"Easy access to historical data fueled the analytics revolution in baseball, resulting in thousands of articles and books, and it's going stronger than ever before. In horse racing, a sport that has a plethora of data and many willing researchers in the betting, breeding and horse ownership population, very little is actually happening.*

*"What has been a tremendous oversight to this point can be corrected. There is almost unmeasurable value in getting the raw data into the hands of the public."*

**Second, Equibase should offer free, basic past performances on all North American tracks.**

The "Basic PP" product Equibase currently offers for \$1.95 per race card, or roughly \$400 per year for unlimited access – should be free.

Reverting again to the words of then Vice Chairman Farish in 1992. "The Thoroughbred industry has the responsibility and obligation to maintain control of those records, and make sure they are made widely available in whatever way suits the best interests of the industry."

Racing's wagering markets have contracted substantially. It is in the best interest of the industry to make basic form data, the most fundamental information used by bettors to trigger wagering decisions, freely available.

Charging customers for this information before a bet is made is a barrier to participation. Attempting to incentivize a player to get a past performance purchase refunded, provided they've made a bet on a day's races, is a convoluted approach to encouraging participation, and surely not a guarantee the customer will participate.

While a myriad of groups (DRF, BRIS, Timeform, "Sheets," etc) offer advanced handicapping products, and should as the market dictates, the basic information collected by racing's official industry-owned information body should be available and presented to potential customers without charge. The cost associated with the collection, retention, security and recall of this data should be absorbed as marketing expenses.

**Third, Equibase should create responsive**

**channels for users to submit feedback in an effort to refine the data.**

With incredible amounts of data, much of it manually compiled or with significant human interaction required to compile it, the opportunity for mistakes, omissions and oversights is substantial. Far more liberal access to the data will uncover even greater errors. A formal channel to collect and correct those will serve the industry better than the status quo.

**Fourth, we believe racing must directly connect with the significant researching power of higher education to study and improve racing's data.**

Three major research universities – Arizona, Kentucky and Louisville – already enjoy tremendous relationships with the racing industry. Collaborating with them, and any number of others who would surely be attracted to the value of racing's data, should be a priority. New, forward-thinking insights can be gleaned by mass-market access to the voluminous data available.

Racing's data also comes with a massive value-add for researchers – a connected financial market tied to every race in the form of pari-mutuel pool data, something other sports cannot offer. The possibilities are exciting.

## EMBRACING A NEW WAY FORWARD

There is an opportunity to change the course of racing, to engage the present-day customer and attract new customers with access to free data.

Equibase has offered some tremendous benefits to industry stakeholders. There is plenty of free information available on its website currently. But we suggest it needs to go much farther, transitioning site visitors from passive consumers of fixed-format information to active, engaged participants in flexibly-available data.

There are far too many variables in play for TIF to present accurate financial projections in this paper. That said, we would challenge the Equibase board and management team to project how their finances would change in a world where these recommendations were enacted – a first step.

Equibase is an incredibly valuable asset for the racing industry, a legacy to the decision making of its founders. With a firm foundation built three decades ago, we believe now is the time to take a giant leap ahead, to go the distance and transform, offering the data to a public customer far different than the ones that engaged the sport in 1990. The

current and potential racing wagering customer possesses programming and processing power the likes of which could never have been envisioned in those days. Enabling public access to racing's vast data sets would signal a new era for the sport.

Freely available data will grow handle, increasing racing's competitiveness for new customers and enhancing the retention of current customers. This is a direct benefit to the tracks, who currently enjoy the spoils of data sales. Growing handle several percentage points replaces the dividends tracks received via the data business, growing it by several more covers Equibase's costs.

Equibase could take racing ahead into a new

era. We believe this will spark participation and engagement. We want racing to have a seat at the analytics table, joining the many mass-market and even niche sports and games which have and continue to benefit from open and available data. How exciting it would be to unfurl the collective intelligence of tech-savvy bettors, fans, researchers and academics on a sport that desperately needs growth.

Combining this data access with more efficient pricing within wagering markets is a recipe for significantly increased bettor participation – a new way forward which all of the industry should support. **TIF**

## REFERENCES

1. [http://www.jockeyclub.com/roundtable/pdf/roundtable\\_95.pdf](http://www.jockeyclub.com/roundtable/pdf/roundtable_95.pdf)
2. <http://www.equibase.com/about/overview.cfm>
3. [http://www.jockeyclub.com/roundtable/pdf/roundtable\\_92.pdf](http://www.jockeyclub.com/roundtable/pdf/roundtable_92.pdf)
4. <http://www.jockeyclub.com/default.asp?section=RT&year=2009&area=2>
5. <http://jockeyclub.com/default.asp?section=Resources&area=10&archives=show&story=355>
6. [https://www.arcseoul2018.com/media/media\\_05\\_view.asp?videoCate=may16\\_1](https://www.arcseoul2018.com/media/media_05_view.asp?videoCate=may16_1)
7. [https://twitter.com/o\\_crunk/status/1101488421841842178](https://twitter.com/o_crunk/status/1101488421841842178)
8. <http://www.sloansportsconference.com/activities/research-papers/2019-research-paper-finalists-posters/>