

Pennsylvania “Skill Machines” and Lottery Revenue

IMPACT ANALYSIS METHODOLOGY AND RESULTS, 2022 UPDATE

Executive Summary

In the following report, we show detailed analyses that estimate more than \$650 million in Pennsylvania Lottery scratch sales have been lost to unregulated Games of Skill machines across the Commonwealth. To put it another way, the \$14.9 billion in scratch product sold between October 2017 and March 2022 could have been more than 4.4% higher, which would in turn have generated over \$200 million more for older Pennsylvanians¹ and local businesses² during that time.

To reach this conclusion, analysts used detailed data collected and maintained by the Pennsylvania Lottery and its primary vendor, Scientific Games, to build statistical models and machine learning algorithms. These tools were used to predict the presence of Skill machines at retailers, isolate the effects of those machines on scratch sales at both the retailer and sales district levels of granularity, and project future impacts based on current Games of Skill expansion trends.

Additionally, this data has been used to monitor and analyze the growth and expansion of Skill machines across the Lottery's retailer network—both “deep” expansion, where the number of machines at each retailer increases, and “wide” expansion, where the number of retailers with at least one machine increases.

All approaches return the same conclusion: the appearance and expansion of Games of Skill machines across the Pennsylvania retail landscape has had significant, negative impact on the revenue realized by the Pennsylvania Lottery.

Data Collection

Any impact analysis must begin with the collection, storage, and cleaning of data. This project relied on two main sources of data: the Field Sales team at the Pennsylvania Lottery and a data warehouse maintained by Scientific Games analysts.

Pennsylvania Lottery Data Collection

Beginning in 2017, the Pennsylvania Lottery District Sales Representatives (DSRs, hereafter) began collecting field data on Games of Skill machine presence in lottery retail locations. The nearly 200 DSRs regularly visited 9,000+ retailers in the lottery network and would report the number of machines they observed in each retail location, along with the date of visit.

Starting in October 2020, the DSRs also began collecting data on retailers with the potential to be recruited for lottery business, as well as the number of machines that were observed in these retail locations each week. These potential retailers are referred to as “leads retailers” going forward.

This data has been compiled weekly and shared with Scientific Games analysts, allowing development of a detailed picture of the geographic spread and network penetration of Games of Skill machines.

¹ Via profit return to the state

² Via retailer commissions and incentives

Growth in Statewide Number of Skill Machines

Figure 1 shows the weekly rise in Games of Skill machines found at lottery retailers from September 2017 through the close of March 2022. Apart from a short plateau in the first few months of COVID-19³, the footprint of these machines has grown steadily and swiftly for over four years with no signs of a slowing trend.

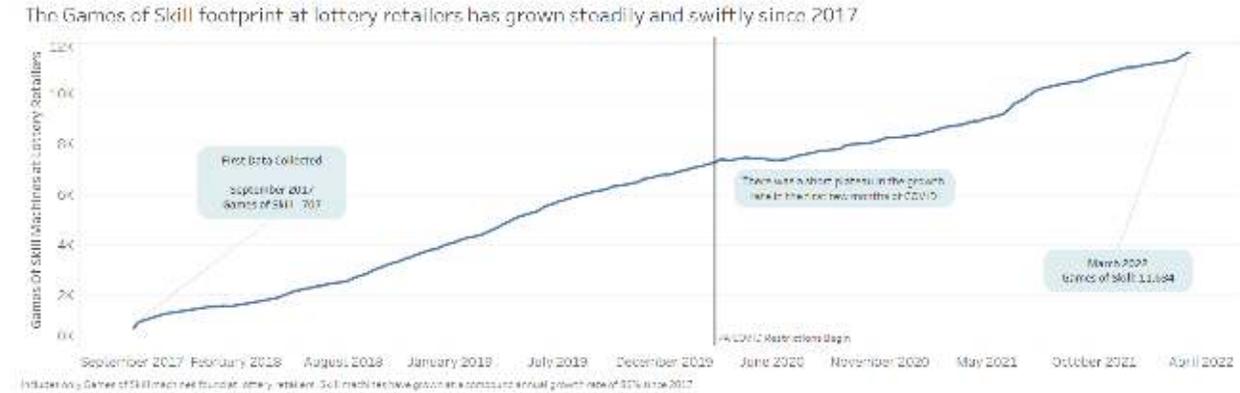


Figure 1

In Figure 2, we can see that the penetration of Games of Skill machines into the lottery retail environment in 2017 was centered mainly around Lycoming County, home of Williamsport, Pennsylvania and one of the first Games of Skill distributors in the Commonwealth. The numbers on the map below the county names signify the number of machines at lottery retailers in September of that year.

By 2022, however, Figure 3 shows the reach and penetration of these machines expanded considerably. In 2017, 53 of 67 counties had at least one machine at retail. In 2022, all 67 counties have at least one machine. In 2017, not a single county in Pennsylvania had more than 100 machines—there are now 36, over half, of the counties with more than 100 machines as of March 2022. This has been an unprecedented and extensive influx of unregulated competition into the Pennsylvania Lottery retail network.

³ This plateau could be the result of Skill companies' inability to recruit during lockdown restrictions; however, the Pennsylvania Lottery sales staff was also unable to visit retailers during this time, so it may just reflect the pause in in-person counting.

Games of Skill Machine Distribution: September 2017

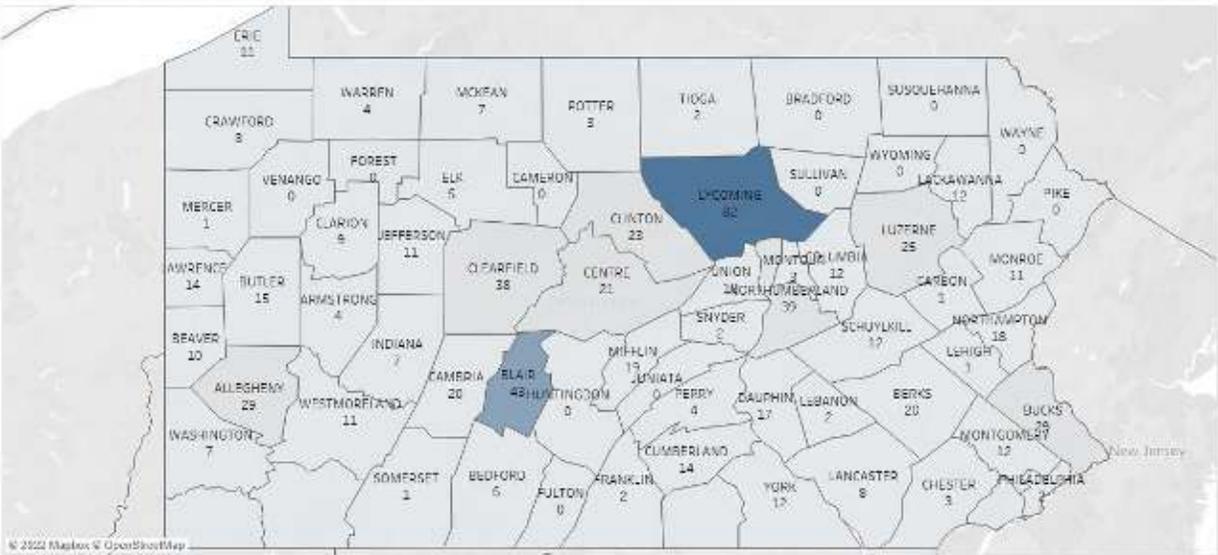


Figure 2

Game of Skill Machine Distribution: March 2022

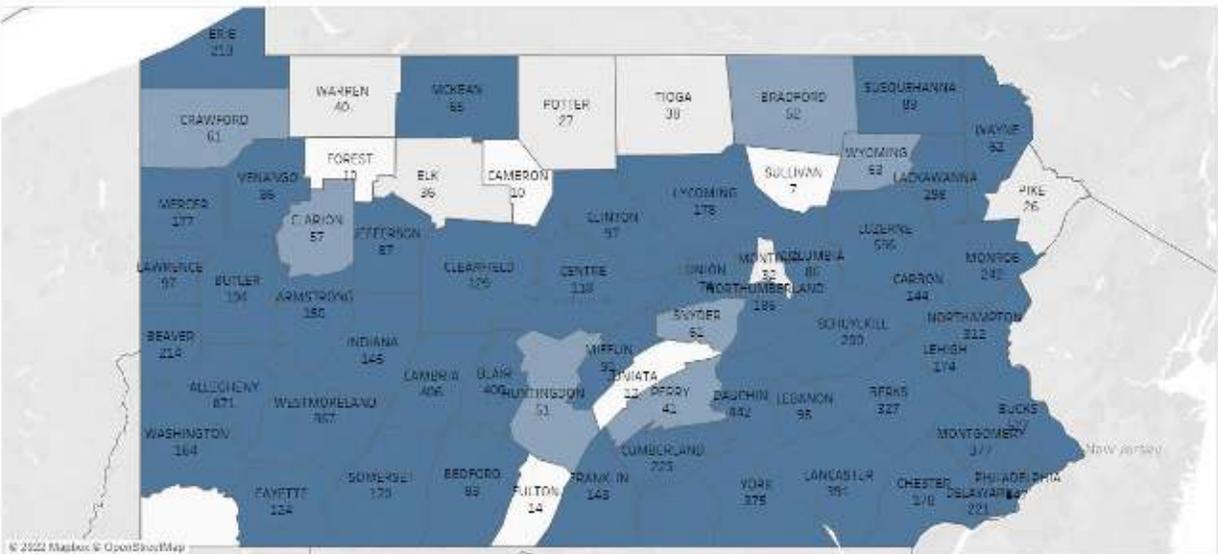


Figure 3⁴

Skill Machine Penetration into the Lottery Retail Network

As the number of Skill machines has grown, the number of lottery retailers with at least one machine on-premise has grown in parallel. As seen in Figure 4, Skill machines could be found in 3.7% of the retail network in September 2017. By March 2022, that figure had grown to 29.8%. In other words, there has been nearly an eight-fold increase in penetration of the lottery retail network in less than four years.⁵

⁴ Uses the same color scale as Figure 2 to show changes relative to 2017

⁵ In 2017, 1 in 27 lottery retailers had at least one Skill machine. In July 2021, 1 in 3.4 retailers have at least one machine on site.

The percentage of lottery retailers with at least one machine has grown from 3.7% to more than 29% in 4.5 years



Figure 4

This steady, swift growth has led to double-digit percentage increases in retailer penetration in nearly every Pennsylvania county, as shown in Figure 5.

The percentage of retailers with Skill machines has grown significantly in nearly every county:

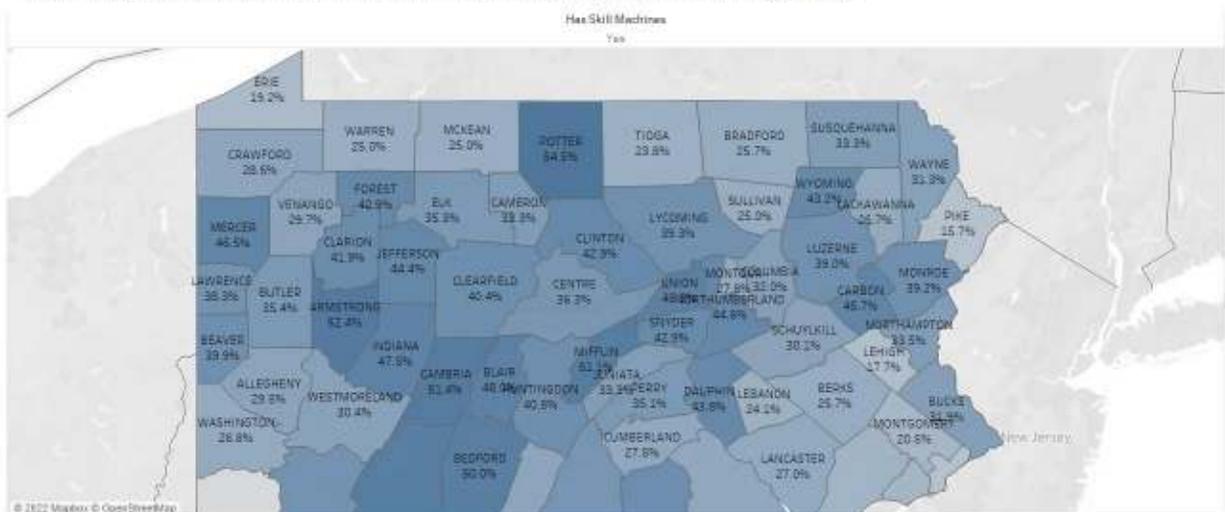


Figure 5

In addition to the widening penetration of the lottery retail network, there has recently been a trend towards deeper penetration as well.⁶ Until December 2020, the maximum number of Skill machines found at any one retailer held steady. Since then, however, there has been a sharp rise, with up to 35 machines located within a single retailer as of March 2022.

⁶ *Widening* penetration refers to the increasing number of retailers that have at least one machine. *Deepening* penetration refers to the increasing number of machines within each retailer.

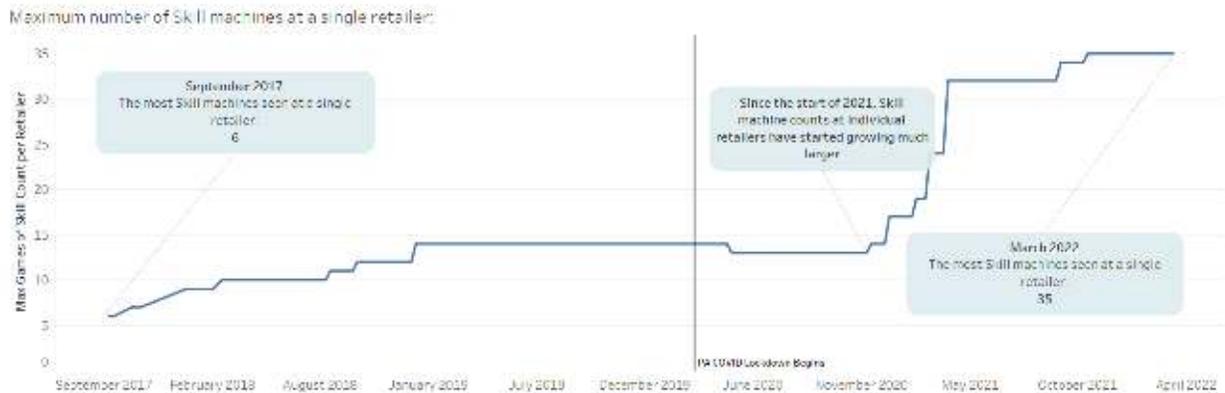


Figure 6

Since 2017, the Pennsylvania Lottery retail network has seen:

- 17 times more Skill machines found on-site across the state;
- Eight times the number of retailers with at least one Skill machine;
- Six times the maximum number of Skill machines found at any one retailer; and,
- The spread of Skill machines to every county in the Commonwealth.

This spread of unregulated gaming competition has dramatically altered the retail landscape in Pennsylvania in less than four years.

Scientific Games Data Warehouse

As the primary systems vendor for the Pennsylvania Lottery, Scientific Games maintains detailed data on sales performance, product offerings, retail network activity, and player behavior. By combining this data with the Games of Skill data collected by the Lottery sales team, Scientific Games analysts have been able to closely monitor any impacts to lottery market activity and performance.

Analysts have previously seen a direct, negative correlation between the density of Skill machines found at lottery retailers in a given geographic area and year-over-year lottery sales performance. In the charts seen in Figure 7, the percent change in sales is charted against the number of Skill machines per lottery retailer for each of the Pennsylvania Lottery's sales districts. The superimposed trend line shows a clear downward trend—as the number of Skill machines per retailer rises, the growth of district-level sales declines.

Relationship between Skill machine density and sales growth, by lottery sales district

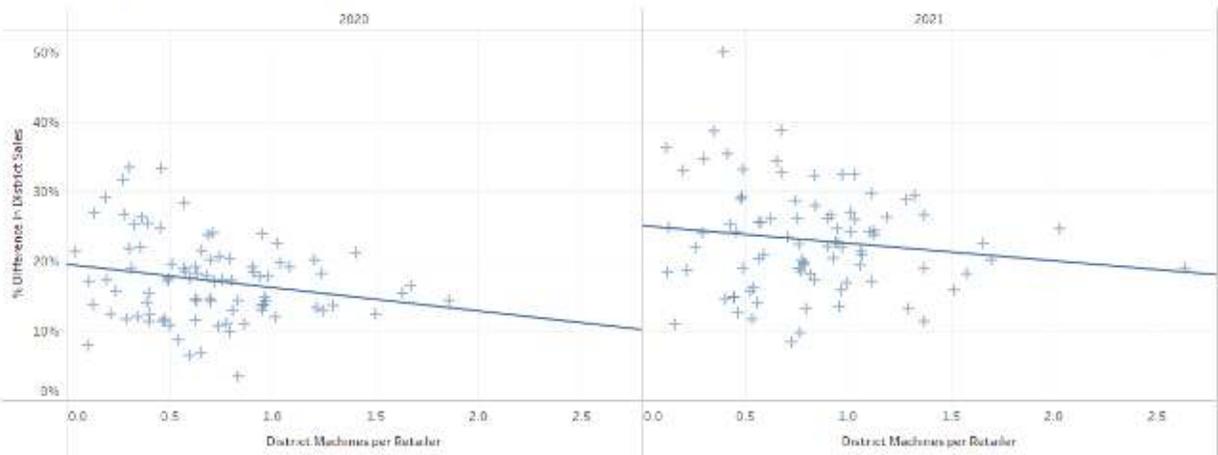


Figure 7: Each + symbol represents one of the lottery sales districts in Pennsylvania

Of course, correlation does not equal causation; so in 2018, analysts began building more formal mathematical models of the relationship between Games of Skill machine presence and sales growth in an effort to understand the causal relationship, if any, between the two.

Regression Modeling

2018 District-Level Impact Model

The first model built by analysts was a “proof of impact” model. Analysts started with a set of data including basic variables that have predictive power for lottery district sales, then added in the count of Games of Skill machines found at lottery retailers in each county to isolate any potential effects on sales.

While this was a simple model⁷ using data over a limited timeframe,⁸ the results accounted for 92.9% of the variance seen in the data and showed a negative impact of \$3,384 per machine in each county over the studied period. This translates to a negative impact on scratch lottery sales of \$2,417 per month per Skill machine. With Games of Skill counts holding steady, the model estimated a sales loss of nearly \$100 million over the course of a year.

⁷ The initial county-level model used only three independent variables: total number of lottery retailers selling in each county, total number of Games of Skill machines found at those retailers, and county sales from the same period in 2015, before the influx of Games of Skill machines

⁸ A 42-day period in 2018

District-Level Impact Model

Variable	Estimate	Std Error	Significance
Retailer Count	\$10,560.00	\$1,923.00	***
Game of Skill Count	-\$3,384.00	\$830.40	***
Sales in 2015	\$0.92	\$0.04	***

Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 1: Initial Impact Model Results

Of course, as troubling as these results might have been, a proof-of-impact model is just that—an attempt to discover some level of impact and justify future research. With such a simple model, at such an aggregate level of granularity, it was entirely possible that the Games of Skill variable was picking up other impacts at retail and unable to isolate what could truly be attributed to machine presence. To refine results, analysts needed to build retailer-level models and track data for a much longer period of time.

Retailer-Level Impact Models

As opposed to a model focused on district-level effects, retailer-level granularity allows analysts to control for factors that have significant effect on sales at individual retailers. Data was included on:

- number of games actively selling;
- number of display facings available for games;
- the average age of games available for sale; and,
- the base trend of the store.

With this information, analysts could more effectively isolate the direct impacts of Skill machines on sales performance.

Beyond these core variables, a model with predictive power also needed a way to account for the diminishing impact of each additional machine at a single retailer, preventing the model from overestimating the effect on scratch sales. To achieve this, the model also included a quadratic version of the raw count of Games of Skill machines in each retailer. This quadratic variable acted as a proxy for the non-linear relationship between Skill machine counts and retailer performance. While this may seem fairly technical, the purpose behind it should be intuitive—adding a fifth machine to a location will likely have a lower impact than the first machine that was added. At some point, adding new machines has smaller negative effects, and analysts needed to be sure the model could find that point of diminishing returns and account for it so that effects were not overestimated.

Once built, the retailer-level impact model remained largely unchanged with only the data feeding it updating each week as new sales and machine counts arrived. The one notable exception is the addition of a COVID-19 control variable, with the period March 16, 2020 to April 30, 2021 designated as “During Covid-19”^[3].

Since the start of 2018, analysts have been able to compile results from 211 iterations of this model, building a detailed, multi-year picture of the negative effect these machines are having on scratch lottery sales.

^[3] COVID restrictions began in Pennsylvania on March 16 and the vaccine was widely available to residents after April 30, 2021.

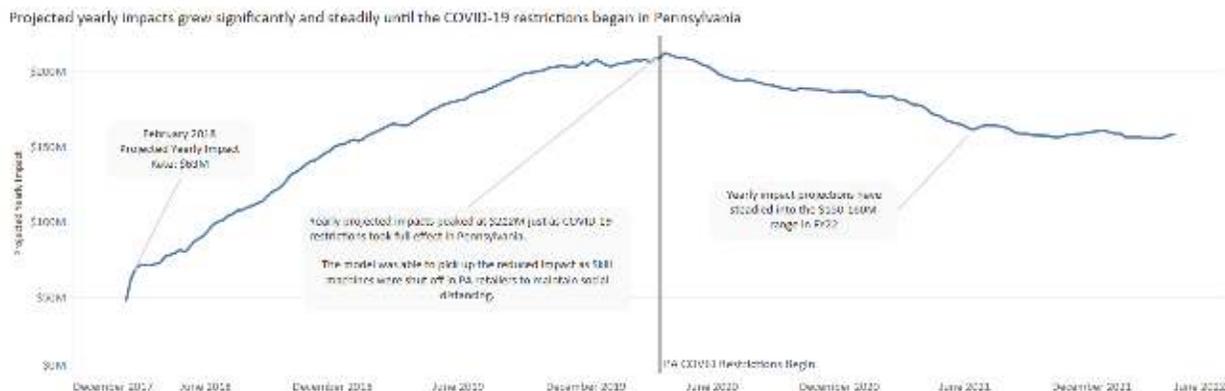


Figure 8

In Figure 8, we see the week-by-week projections the model created for estimated future impacts of Skill machines on scratch sales. Each projection gives estimated impact over the course of the next year, assuming counts and per machine impacts remain steady. As one can see, the model predicted approximately \$63 million in impact over the year following February 2018. That projection grows at about the same pace as the Skill machine counts increased, up to a peak impact estimate of \$212 million for the year following March 2020. Since March 2020, the yearly projections have declined⁹ during the COVID-19 pandemic, until starting to rise again in June 2021. Since July 2021, yearly impact estimates have steadied into the \$150-\$160 million range.

The short decline in impact makes sense. Skill games were not deemed an essential business, and field reports from Lottery sales staff indicate that many retailers had shut down the machines to comply with social distancing guidelines and discourage customers from lingering in their stores. The fact that the model was able to pick up on this change in retail landscape without being altered or given data on machine usage¹⁰ greatly increases confidence that the effects are being accurately measured.

With so many weeks of data and analysis completed, analysts need not rely solely on impact projections, but can also report on week-by-week actual impacts to see how the two trends compare. Figure 9 shows the actual week-to-week effects output by the retailer-level impact model. We see the same trend, with per week impacts rising steadily until COVID restrictions began, then declining until the vaccines became widely available, then beginning to rise again. This is unsurprising—after all, the look-ahead projections are based on the actual weekly impacts!

⁹ Though, at a slower pace than they increased in the years previous

¹⁰ Only the *presence* of machines is tracked by Lottery field staff, not usage

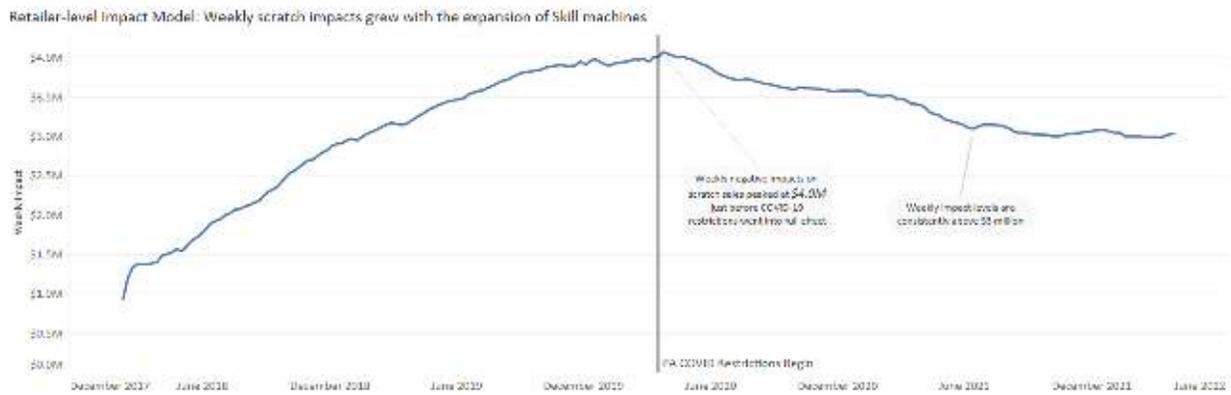


Figure 9

Using weekly actual impact numbers, analysts created 52-week running totals to see the long-term trend of Skill machines’ effects on scratch sales. Figure 10 shows that trend from the start of 2019 through July 2020. Over the past 52 weeks, the Pennsylvania Lottery has lost more than \$165 million in scratch sales to Games of Skill machine play.

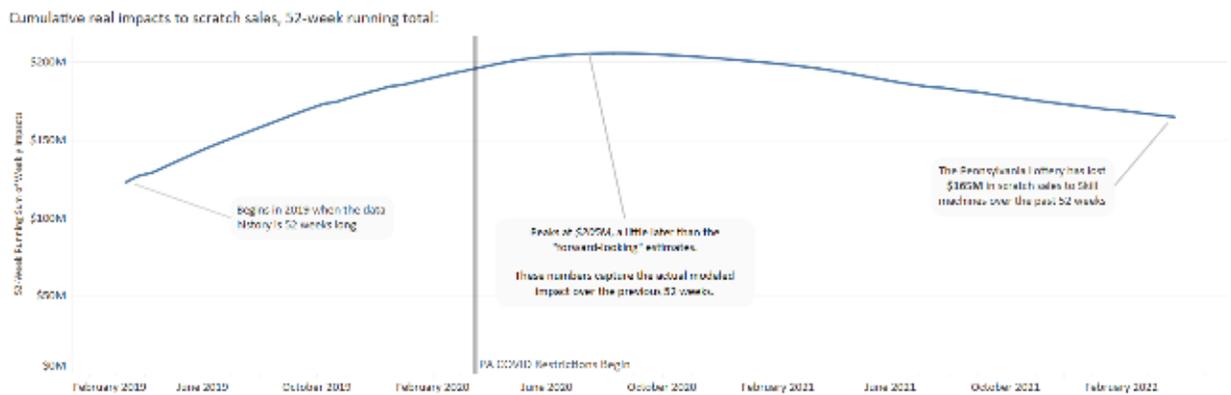


Figure 10

Cumulatively, the retailer-level impact model estimates that \$658.6 million in scratch sales have been lost due to the presence of Games of Skill machines at lottery retailers. This amounts to a 4.4% of the \$15.6 billion in scratch product that could have potentially been sold in Pennsylvania since September 2017¹¹.

Games of Skill Presence Prediction Model

Analysts also approached the question from the opposite direction—could the presence of Games of Skill machines be predicted based on the sales patterns present at each store?

¹¹ Actual sales were \$14.9 billion

Using a modern machine learning¹² technique called stochastic gradient boosting¹³, analysts developed a model that looked to identify which retailers in the network had at least one Skill machine on site and which did not. The algorithm only considered sales patterns at each store in question, overall market seasonality for scratch products, the sales area in which the store could be found, and the current product offering on sale.

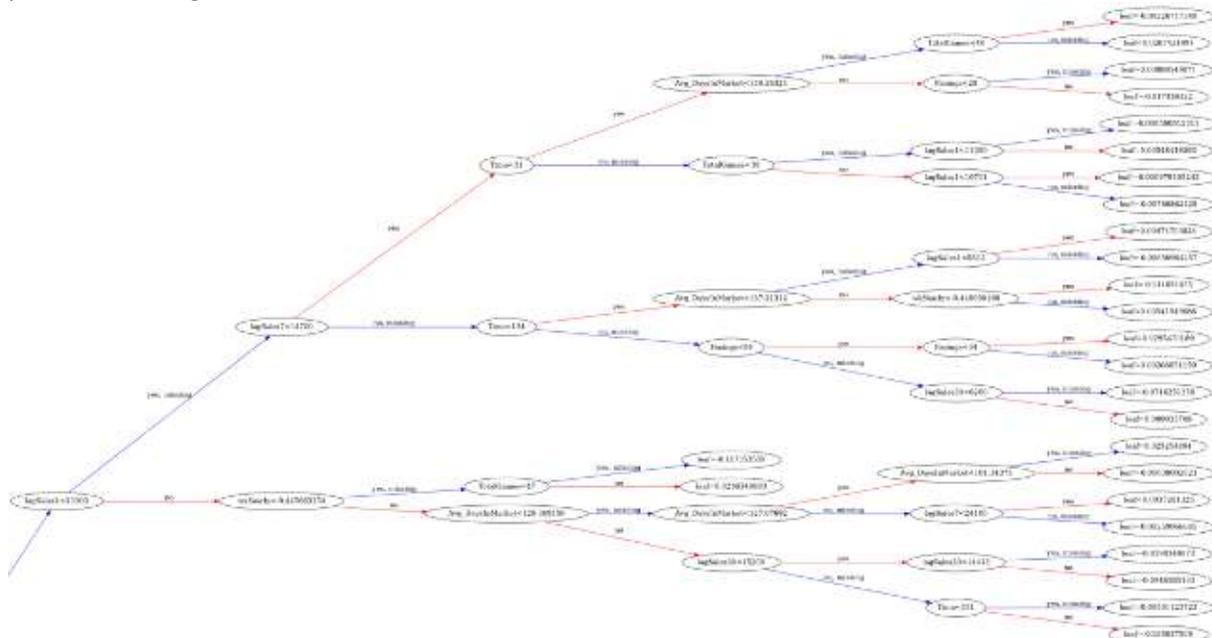


Figure 11: A graphic illustration of a small subset of the decisions made by the machine learning algorithm as it tries to predict the presence of Skill machines at individual lottery retailers

Even with this limited set of input features, the Skill presence prediction algorithm was able to detect with 88% accuracy whether or not there were Games of Skill machines in each retailer on any given week. Further, 89% of the time the algorithm thought there was a machine, there actually was¹⁴, and 88% of the time it thought there was not a machine, there was no machine on premise that given week¹⁵.

The Skill presence prediction algorithm shows that a significant, easily detectable impact on sales and sales trends occurs from the moment Skill machines get added to a store. While not surprising given the aforementioned research, these finds give additional weight to the impact effects previously discussed. From day one post-install, Games of Skill machines make a clear, negative impact on scratch lottery sales at Pennsylvania retailers.

¹² Machine learning is defined as, “the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data.”

¹³ While the theoretical specifics of stochastic gradient boosting are beyond the scope of this report, “boosting” is a technique within that discipline that allows “shallow,” or weak, learning methods to improve upon themselves until they have high predictive power.

¹⁴ Known as ‘sensitivity’

¹⁵ Known as ‘specificity’

Potential Leads Retailers

Games of Skill machines can also impact the Pennsylvania Lottery indirectly. A potential lottery retailer may opt for, or even decline, lottery products being sold in their store due to the presence of Skill machines. Beginning in October 2020, the Pennsylvania Lottery sales team began collecting data on the number of Skill machines seen in non-lottery retailers they attempted to recruit. The sales team has identified 531 potential retailers with a total of 1,541 Skill machines observe¹⁶ at these locations.

Using the data collected and estimating retailer sales based off current lottery retailers in the same district with the same lottery business type and number of machines present, we can estimate the potential lost sales if these retailers were successfully recruited by the lottery. If all 531 retailers were recruited, we estimate \$105 million of total lottery sales could have been realized between October 2020 and March 2022¹⁷.

Compared with peer lotteries, Pennsylvania has significantly lower social establishment penetration. The Maryland Lottery has been able to recruit 19.8% of its statewide bars and taverns to sell Keno and other lottery products. In Pennsylvania, that number sits just above 7%. Why is the penetration number so much lower in Pennsylvania? One reason is the difficulty of recruiting locations with active Skill machines. Of these 531 “leads retailers” with confirmed Skill machines on site, only 23 were successfully recruited—a success rate of only 4.3%.

It seems clear that the presence of Skill machines produces a chilling effect on the Lottery’s ability to successfully recruit new retailers, which has negative effects on future sales. If the Lottery had been able to achieve 20% recruitment success at locations with Skill machines, we estimate an additional \$16.5 million in sales could have been realized.

Conclusions

As we can clearly see, all approaches to understanding the effect of Games of Skill machine presence on Lottery revenue point to the same conclusion: the appearance and expansion of Games of Skill machines across the Pennsylvania retail landscape has had significant, negative impact on the revenue realized by the Pennsylvania Lottery.

Skill machine penetration of the retail network has grown from less than 5% in 2017 to nearly 30% today. In addition to this “wide” expansion, Skill machines have also expanded more deeply into the network, increasing the average and maximum number of machines found in individual retailers.

Even though scratch sales have consistently grown during this timeframe, our analysis finds that the total revenue realized by the Commonwealth has been 4.4% lower than it could have been, costing the Pennsylvania Lottery more than \$650 million in top-line revenue since tracking began in late 2017.

Finally, the presence of Skill machines also reduces Lottery recruitment success, further dampening sales, and significantly lowering returns to the Commonwealth.

¹⁶ An average of 2.9 Skill machines per retailer, just under the average of 3.7 machines found at lottery retailers

¹⁷ This number includes an estimated \$120M in gross potential lottery sales, with a \$15M reduction in that figure due to the presence of Skill machines at each retailer.

