

Can « big data » be useful to Public Health ?

1

Jean-Michel Costes

**Summer Interactive Symposium – Research 2.0,
Concordia University - Montréal**

June 10, 2015



Coopération France-Québec

Outline



Introduction :

- **The Legal Framework of e-Game in France**
- **The e-Game in France in 2014**
- **Prevention of Problematic e-Game (ePG)**

International studies on the use of gamblers data

Project of a predictive model of ePG based on the gamblers data, collected in France

- **« Massive data » on e-Game in France**
- **The ODJ-ARJEL Study**
- **Methodological, operational and political issues**

Conclusions

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Use of digital data, a sensitive topic in France

3

The bill on digital



The bill on intelligence services

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Gambling : legal framework before 2010

4

Gambling legal framework in France before 2010

	Casinos (Tables games, Slot machines) Poker	Lotteries	Sports betting	Horse racing
Offline	Authorization/ Concession	FDJ (Française des jeux)	FDJ (Française des jeux)	PMU (Pari mutuel urbain)
Online		FDJ (Française des jeux)		

	State monopolies
	Private operators
	Prohibited

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

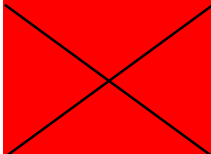
Projet de
modèle en
France


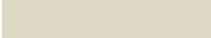

Conclusion

Gambling : legal framework since 2010

5

Gambling legal framework in France after 2010

	Casinos (Tables games, Slot machines) Poker	Lotteries	Sports betting	Horse racing
Offline	Authorization/ Concession	FDJ (Française des jeux)	FDJ (Française des jeux)	PMU (Pari mutuel urbain)
Online	 Prohibited	FDJ (Française des jeux)	Licensed operators	Licensed operators

 State monopolies
 Private operators
 Prohibited

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Online gambling in France

6

- National survey on online gambling 2012; OFDT/ODJ (N = 4,000 online gamblers)
(Tovar, Costes, and Eroukmanoff, 2013)
- National survey on gambling 2014; INPES/ODJ (N = 16,000 French, aged of 15-75 years old)
(Costes and al. 2015)
- In 2014, in France, online gambling prevalence is:
 - 4.1 % in general population aged from 15 to 75 (IC 95 % : 3.8% – 4.4%)
 - 7.3 % among gamblers

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Factors associated with the use of internet to gamble – univariate approach (1)

7

	% of gamblers who use internet to play	OR bruts	IC-95%
Socio-demographics			
Gender			
	Female	3.6	1
	Male	10.9	3.25 2.65 – 3.97
Age			
	15-24 years	9.7	1.16 0.86 – 1.54
	25-34 years	10.9	1.31 1.01 – 1.69
	35-44 years	8.5	1
	45-54 years	6.4	0.73 0.56 – 0.97
	55-75 years	3.6	0.40 0.31 – 0.53
Occupation			
	Employed	8.4	1
	Students	7.3	0.85 0.63 – 1.17
	Unemployed	9.4	1.13 0.82 – 1.56
	Other inactives	3.6	0.41 0.31 – 0.53
Socio-professional categorie			
	csp-	5.6	1
	Csp-m	9.2	1.72 1.40 – 2.12
	csp+	10.6	2.01 1.61 – 2.51
Diploma			
	< High school diploma	5.2	1
	High school diploma	8.5	1.71 1.35 – 2.16
	> High school diploma	11.2	2.32 1.89 – 2.85

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Factors associated with the use of internet to gamble – univariate approach (3)

8

	% of gamblers who use internet to play	OR bruts	IC-95%
Game behaviour			
Frequency			
	{1;24}	4.5	1
	{24;52}	1.58	1.21 – 2.07
	{52;104}	2.20	1.72 – 2.82
	{104;+}	3.94	3.13 – 4.96
Spending			
	< 250 €	4.5	1
	{250;500} €	2.06	1.59 – 2.68
	{500 ;1000} €	3.11	2.32 – 4.16
	> =1000 €	5.45	4.30 – 6.90
Nb of games practiced (7 classes)			
	1	4.6	1
	2	7.5	1.68 1.36 – 2.08
	3 and +	17.6	4.43 3.55 – 5.53
Practice (ref= 1; do not practice)			
	Draw Games	8.1	1.48 1.20 – 1.83
	Scratch Card Games	6.5	0.79 0.66 – 0.95
	Horseshoe Betting	13.9	2.33 1.83 – 2.96
	Sports bets	31.4	7.97 6.35 – 9.99
	Poker	46.0	15.00 11.53 – 19.51
	Casino Games	13.0	2.15 1.71 – 2.72
Problematic Game (ICJE >=5)			
	No	7.0	1
	Yes	23.0	3.98 2.45 – 6.45

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Factors associated with the use of internet to gamble - multivariate approach (2)

9

Logistic Regression

Variable to explain : online gambling

Explanatory variables:

Socio-demographics : sex, age, spc, diploma

Gambling behavior : frequency, spending, nb of games, type of game practiced, problematic gambling (CPGI >=5)

	OR ajusted	IC-95%		OR ajusted	IC-95%
Socio-demographics			Game behaviour		
Gender			Frequency		
	Female	1	< 52 times per year	1	
	Male	1.65 1.31 – 2.08	> =52 times per year	1.60	1.25 – 2.07
Age			Spending		
	15-44 years	1	< 500€ per year	1	
	45-75 years	0.78	> =500€ per year	1.56	1.18 – 2.06
Socio-professional categorie			Nb of games practiced (7 classes)		
	csp-	1	1 or 2	1	
	csp+	1.47 1.18 – 1.84	3 and +	0.92	0.70 – 1.20
Diploma			Practice (ref= 1; do not practice)		
	< High school diploma	1	Draw Games	1.86	1.43 – 2.41
> =High school diploma			Scratch Card Games	0.83	0.67 – 1.03
		2.14 1.68 – 2.73	Horsrace Betting	1.60	1.17 – 2.18
			Sports bets	4.00	2.94 – 5.45
			Poker	9.01	6.43 – 12.63
			Casino Games	1.06	0.77 – 1.47
			Problematic Game (ICJE >=5)		
			No	1	
			Yes	0.78	0.38 – 1.58

**Jeu en ligne :
« big data » et
santé publique**

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Outline



Introduction :

- **The Legal Framework of e-Game in France**
- **The e-Game in France in 2014**
- **Prevention of Problematic e-Game (ePG)**

International studies on the use of gamblers data

Project of a predictive model of ePG based on the gamblers data, collected in France

- **« Massive data » on e-Game in France**
- **The ODJ-ARJEL Study**
- **Methodological, operational and political issues**

Conclusions



**Jeu en ligne :
« big data » et
santé publique**

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Behavior marker and DSM criterion

11

Table 1: Summary of problem gambling criteria (DSM-IV) and likelihood of identification of problem gambling behaviour online

<i>DSM-IV Criterion</i>	<i>Likelihood of online identification</i>
Experiencing salience/preoccupation	Very good possibility
Experiencing tolerance	Reasonable possibility
Experiencing relapse	Slight possibility
Experiencing withdrawal symptoms	Unlikely
Escaping from reality	Unlikely
Chasing losses	Definitely
Concealing involvement	Unlikely
Engaging in unsociable behaviour	Unlikely
Ruining a relationship/opportunity	Unlikely
Other people providing a bail-out	Slight possibility

(Griffiths 2013)

**Jeu en ligne :
« big data » et
santé publique**

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Behavior Indicators

12

- ✓ Data « gambler account » : gender, postal code...
- ✓ Game data : type of game, bets, amount wagered, winnings...
- ✓ Financial data : deposits, withdrawals ...
- ✓ Moderators data : limits...

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

The baseline - PG

13

Jeu en ligne :
« big data » et
santé publique

✓ Screening tool : CPGI ...

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

✓ An indirect indicator :

- Request of self-exclusion
- Evolution of the moderators' limits
- Interventions of operators

The Transparency Project

14

Division on Addiction, The Cambridge Health Alliance,
Harvard Medical School and Bwin

Links between PG and :

- **Level and frequency of bet, gambling intensity after opening an account** (*LaBrie and Shaffer 2011*)
- **Frequency, intensity, variability and evolution of bets in the 1st month** (*Braverman and Shaffer 2012*)
- **Level, frequency and accumulation of stakes, losses, number of days of gambling** (*Gray, LaPlante, and Shaffer 2012*)

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

The research program – « gambling machines » in the UK

15

PREDICTING PROBLEM GAMBLERS: Analysis of industry data (Wardle, Parke, and Excell 2014) (Wardle and al. 2014) (Excell and al., 2014)

19 potential predictive indicators have been identified :

- frequency of play
- duration of play
- net expenditure
- levels of play engagement (interactions between expenditure and time) number of activities undertaken
- chasing behaviour
- number and type of games played within a session of play debit card reload/cash and debit card switching
- debit card payment declines
- variability in staking behaviour
- use of repeat bets functions
- play of multiple machines simultaneously
- stake size
- game volatility
- cash-out behaviour.

**Jeu en ligne :
« big data » et
santé publique**

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

The research program – « gambling machines » in the UK

16

Survey on a sample of 4700 gamblers

- **Gambling data of the loyalty card**
- **Questionnaire (DSM and CPGI)**

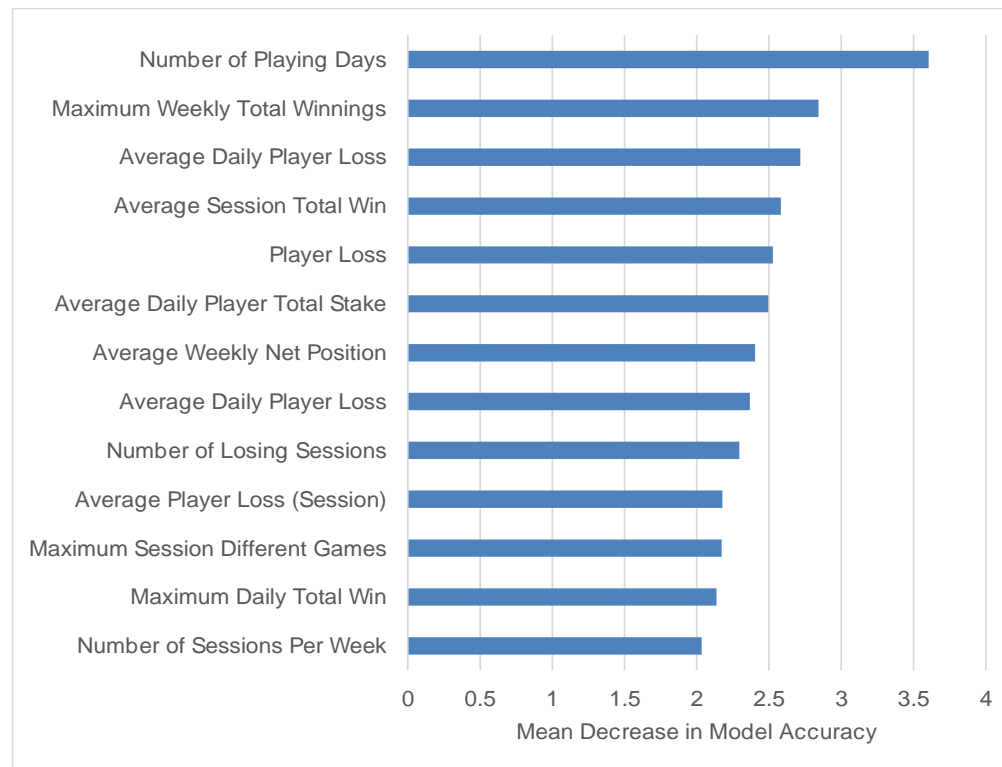
Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion



The research program – « gambling machines » in the UK

17

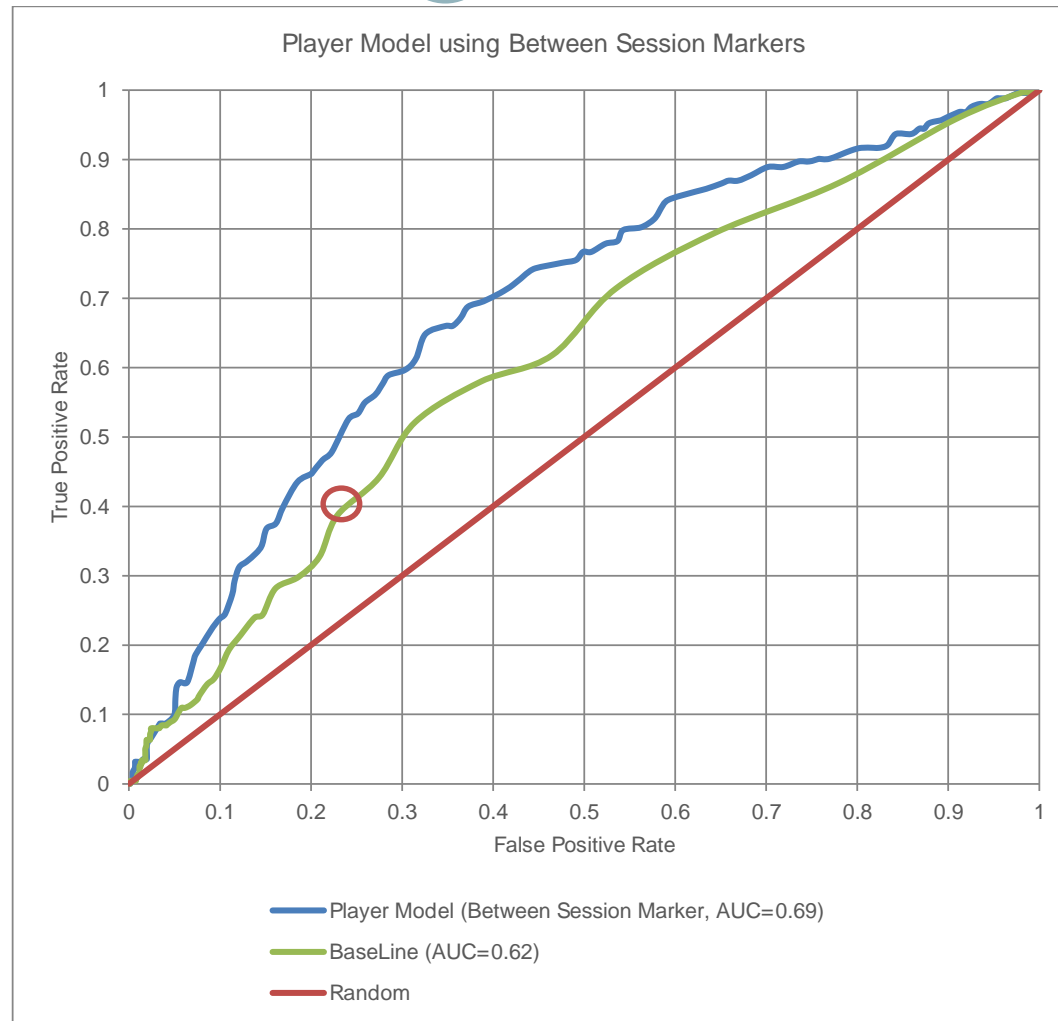
Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion



Outline



Introduction :

- The Legal Framework of e-Game in France
- The e-Game in France in 2014
- Prevention of Problematic e-Game (ePG)

International studies on the use of gamblers data

Project of a predictive model of ePG based on the gamblers data collected in France

- « **Massive data** » on eGame in France
- **The ODJ-ARJEL Study**
- **Methodological, operational and political issues**

Conclusions

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

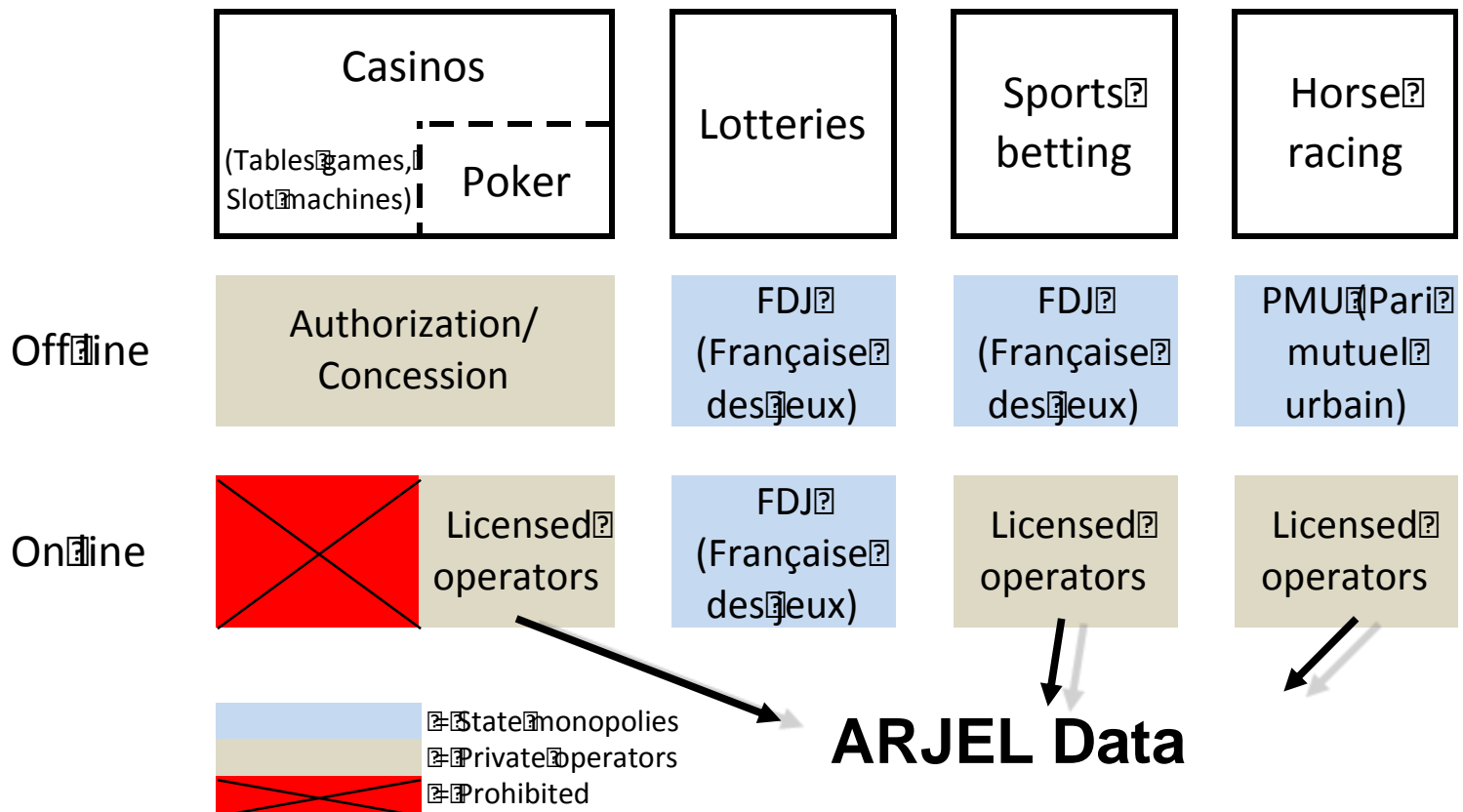
Projet de
modèle en
France

Conclusion

Gamblers data - ARJEL

19

Gambling legal framework in France after 2010



Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

ARJEL Data

20

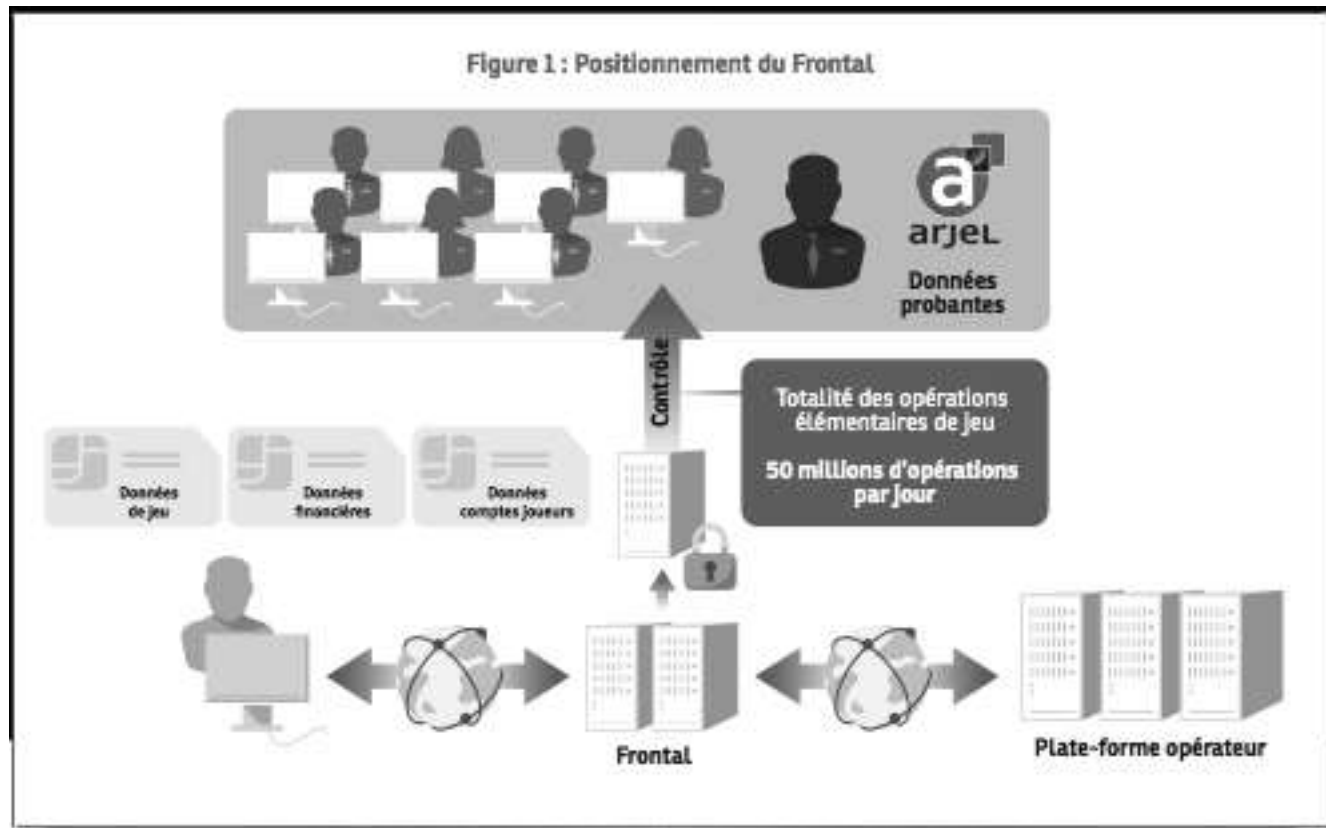
Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion



ARJEL Data

21

Jeu en ligne :
« big data » et
santé publique

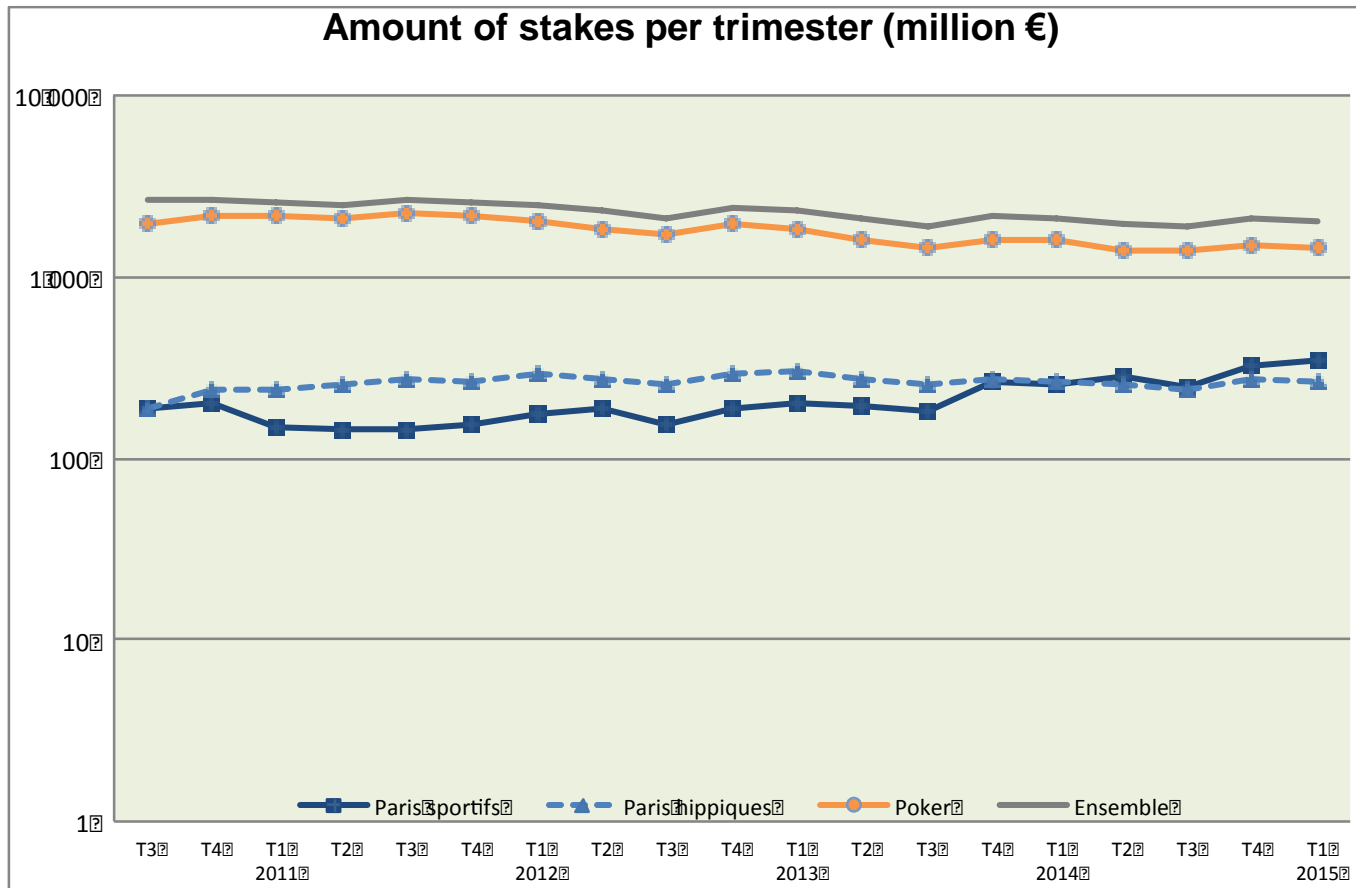
Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Amount of stakes per trimester (million €)



ARJEL Data

22

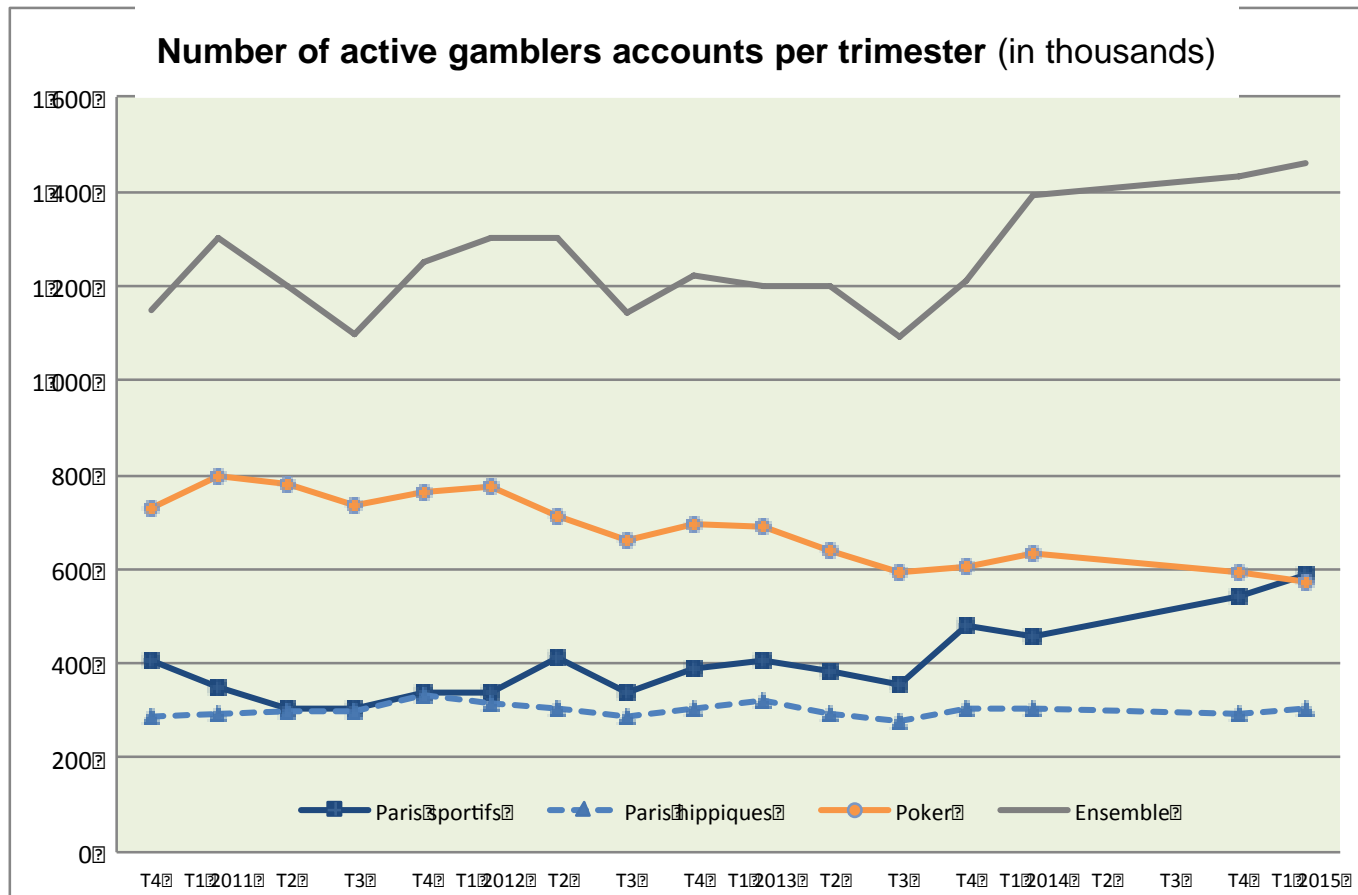
Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion



ARJEL Data

23

Jeu en ligne :
« big data » et
santé publique

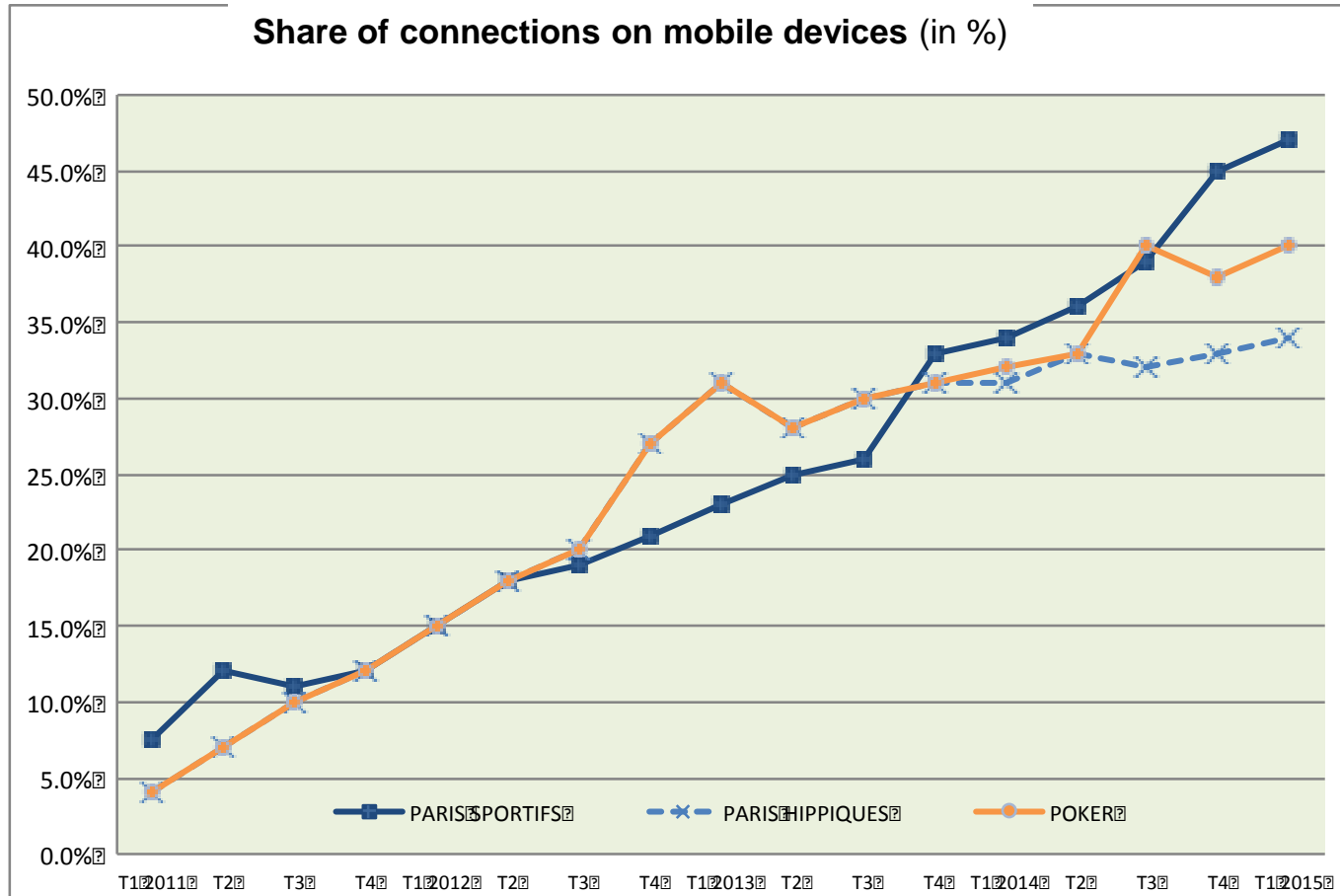
Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Share of connections on mobile devices (in %)



ARJEL Data

24

Jeu en ligne :
« big data » et
santé publique

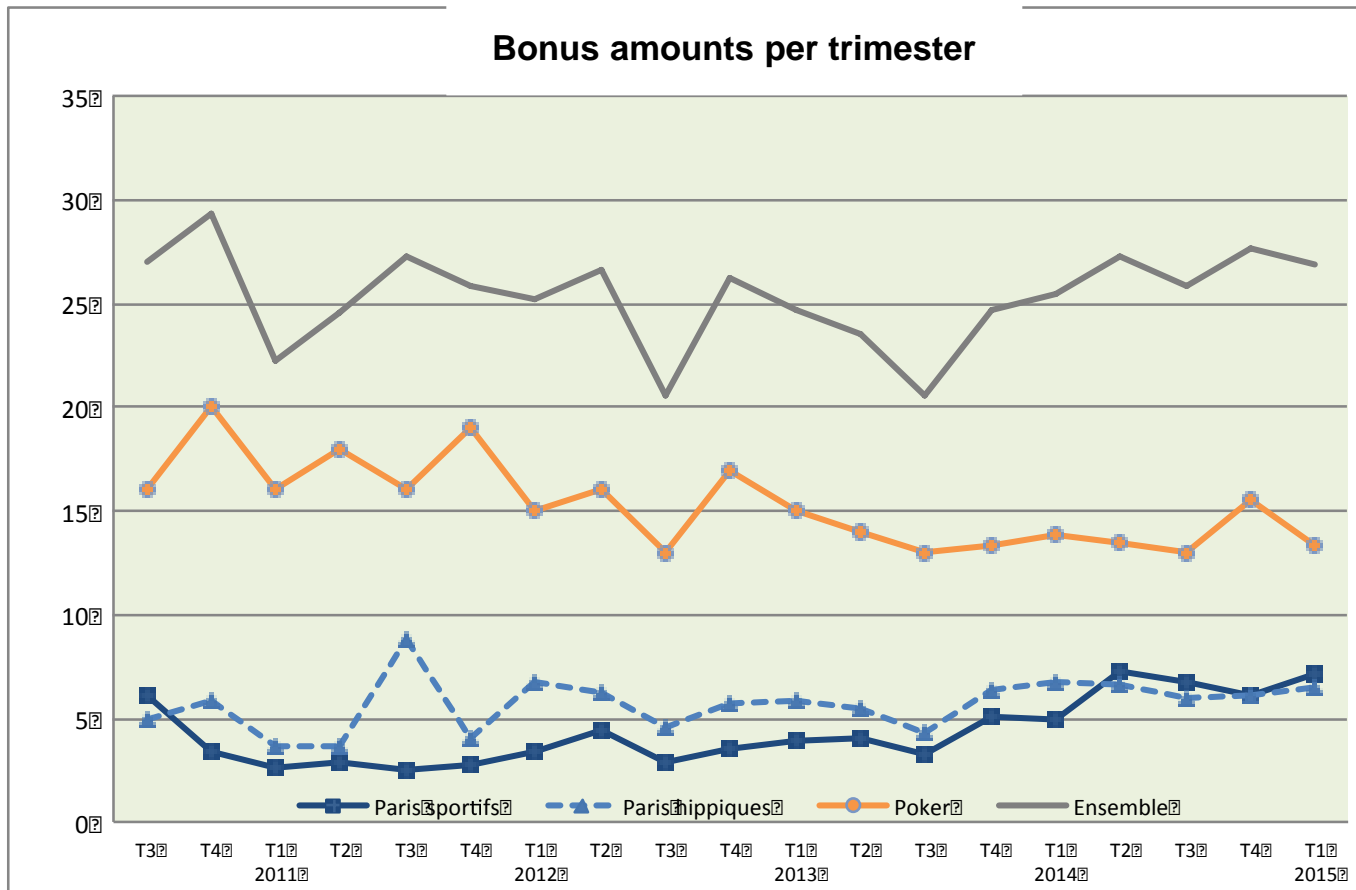
Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

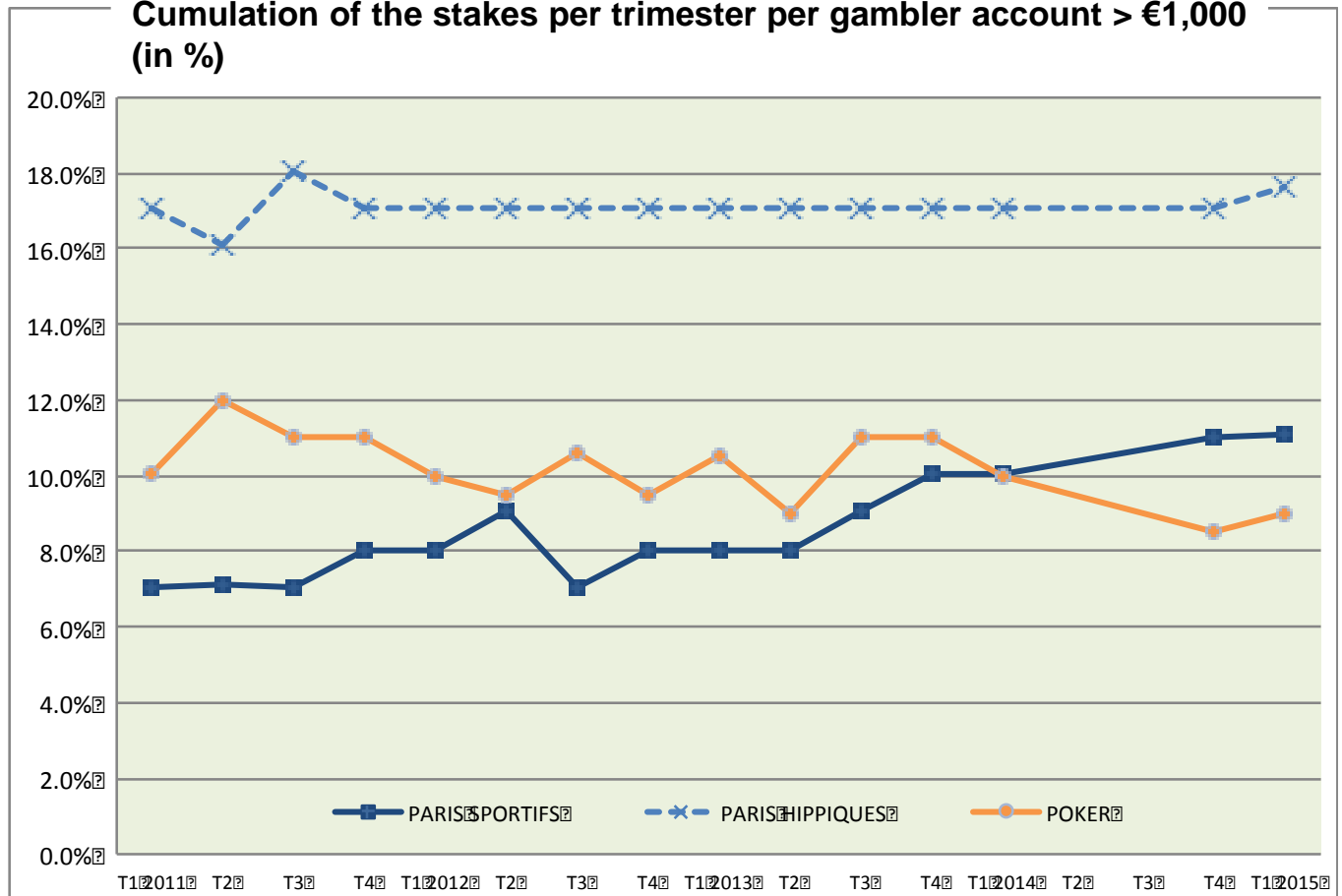
Bonus amounts per trimester



ARJEL Data

25

Cumulation of the stakes per trimester per gambler account > €1,000 (in %)



Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

ARJEL Data

26

Jeu en ligne :
« big data » et
santé publique

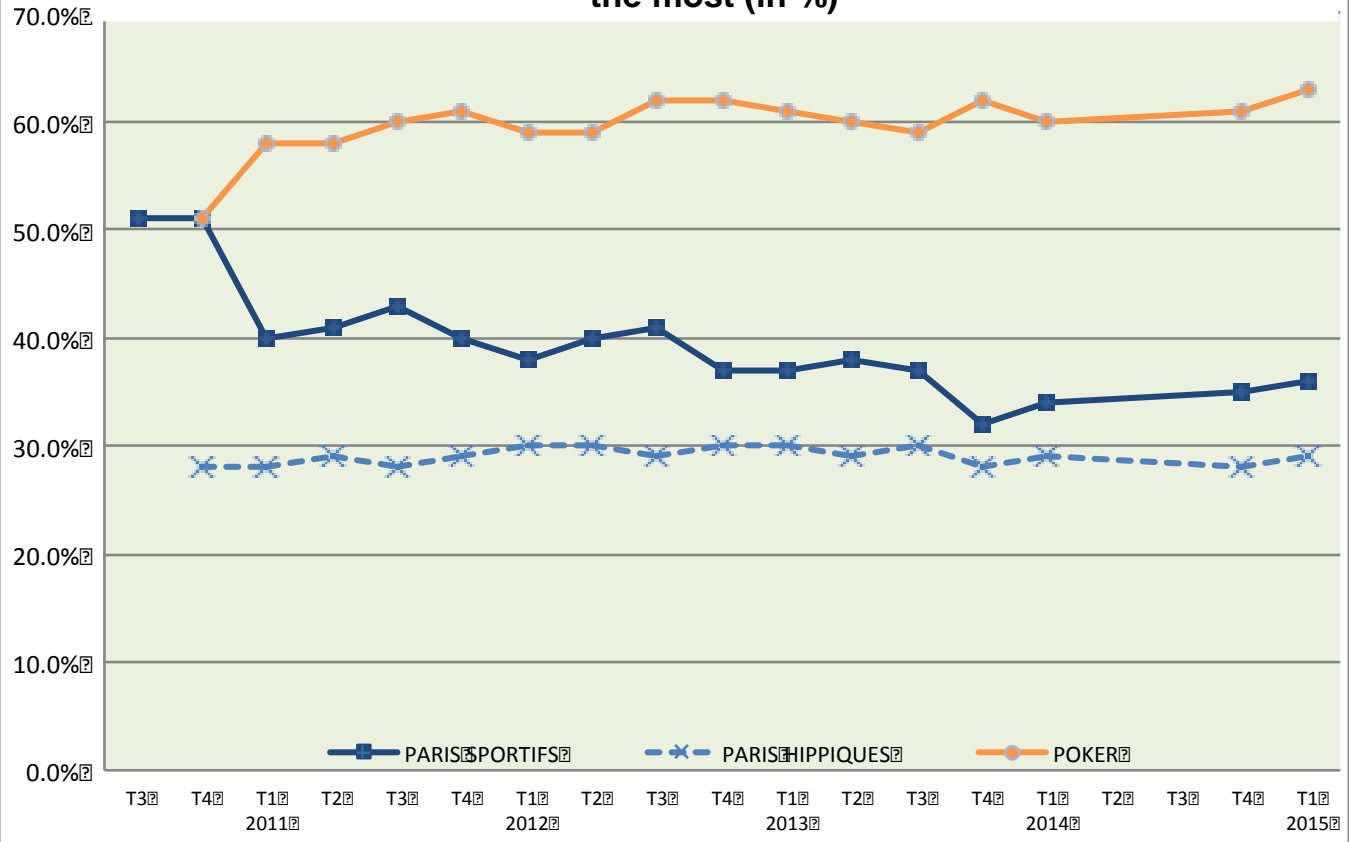
Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Cumulation of the stakes of gamblers accounts of the 1st centile who bet the most (in %)



The ODJ-ARJEL Study

27

- **Goal**

To elaborate a predictive model of problematic gambling based on regulation data.

- **General approach**

The CPGI score is the variable to explain; the indicators of activity, the explanatory variables of the model

- **The reference screening tool: CPGI**
- **Inclusion criterion in the study**

Joueur

Compte Joueur ...

Mois

Jour

Critères d'inclusion des individus dans l'étude et effectifs des joueurs correspondant

	compte ouvert +12M	compte ouvert de 12 à 30 mois	compte ouvert de plus de 30 mois
Compte actif/12DM	440000	169000	31000
actif/30DJ			
inactif/30DJ	417000	181000	0

Un compte actif sur une période est un compte sur lequel on observe au moins une activité de jeu (mise) sur cette période. On considère qu'un individu a un compte actif au cours d'une période (12DM ou 30DJ) si au moins un compte a été actif au cours de cette période.

En colonne Pour la durée d'ouverture, on prend en compte la durée du plus ancien compte ouvert (non fermé).

Inclus dans l'échantillon

Inclus dans le modèle

Jeu en ligne : « big data » et santé publique

Introduction

Etudes internationales

Projet de modèle en France

Conclusion



The ODJ-ARJ

28

- **Goal**

To elaborate a predictive model of problematic gambling based on regulation data.

- **General approach**

The CPGI score is the variable to explain; the indicators of activity, the explanatory variable model

- **The reference screening tool CPGI**
- **Inclusion criterion in the model**
- **Choosing variables of activity entering the model**

Regulation data selected for the study

For Player	Type of data
Player number	
Age	
Sex	
For Account	Type of data
Player number	
Account number	
Opening date	
Validation date	
Bets limit	History since opening (amount date)
Deposits limit	History since opening (amount date)
Withdrawal limit	History since opening (amount date)
Self-exclusion	Number and history since opening
For Months	Type of data
Player number	
Account number	
Month number	
Winings	Cumulation per type of game
Losses	Cumulation per type of game
Simple horserace betting	Stakes cumulation
Complex horserace betting	Stakes cumulation
Complex sports bets	Stakes cumulation
Sports live betting	Stakes cumulation
Sports bet – type of sports	Stakes cumulation
Chasing 1	Number of "3 consecutive deposits on a rolling period of 12h
Chasing 2	Number of "deposit in the hour following the stake
For Day	Type of data
(definition: from noon to noon)	
Player number	
Account number	
Month number	
Day number	
Stakes	Number and history since opening
Stakes	Cumulation per type of game
Buy-in	Number and history since opening
Buy-in	Cumulation
Deposits limit	Nb and amounts cumulation
Withdrawals	Nb and amounts cumulation
Bonus	Nb and amounts cumulation per type of game
Color code:	Poker
specific variables	Horserace betting
for each type	Sports Bets
of game	Bets

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion



The ODJ-ARJEL Study

29

- **Goal**

To elaborate a predictive model of problematic gambling based in regulation data.

- **General approach**

The CPGI score is the variable to explain; the indicators of activity, the explanatory variables of the model

- **The reference screening tool: CPGI**
- **Inclusion criterion in the study**
- **Choosing variables of activity entering the model**
- **Recruitment of gamblers to complete CPGI.**

Sample of 10,000 gamblers

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Issues of the study

30

Methodological Issues :

- **Thee referential, CPGI, risk threshold**
- **Multi-activity**

Operational Issues :

- **Survey gamblers**
- **Choice of statistical tools**

Ethical Issues :

- **Use of personal data**

Political Issues :

- **Will, legal developments**

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Conclusion

31

The
us t

•Fac

•Re

•Ta

“
f
p
a
r

JE VEUX SAVOIR CE QUE PENSE
LE CONSEIL CONSTITUTIONNEL
DE LA LOI SUR LE RENSEIGNEMENT

ON PEUT POSER
DES MICROS
DANS LEURS LOUPOUX
SI VOUS VOULEZ...



es
se

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

odj
Observatoire
des jeux

References

32

Braverman, J., et H. J. Shaffer. 2012. « How Do Gamblers Start Gambling: Identifying Behavioural Markers for High-Risk Internet Gambling ». *The European Journal of Public Health* 22 (2): 273-78. doi:10.1093/eurpub/ckp232.

Costes, Jean-Michel, Vincent Eroukmanoff, Jean-Baptiste Richard, et Marie-Line Tovar. 2015. « Les jeux d'argent et de hasard en France en 2014 ». ODJ, Les notes de l'observatoire des jeux, , no 4: 9.

Excell, David, Georgiu Bobashev, Daniel Gonzales-Ordenez, Heather Wardle, Tom Whitehead, Robert J. Morris, et Paul Ruddle. 2014. « report 3 - predicting problem gambling - final.pdf ». Responsible Gambling Trust.

Gray, Heather M., Debi A. LaPlante, et Howard J. Shaffer. 2012. « Behavioral Characteristics of Internet Gamblers Who Trigger Corporate Responsible Gambling Interventions. » *Psychology of Addictive Behaviors* 26 (3): 527-35. doi:10.1037/a0028545.

Griffiths, Mark D. 2013. « Behavioral tracking in gambling: new empirical data and the development of a new social responsibility tool ».

Griffiths, M. D., et M. W. Whitty. 2010. « Online behavioural tracking in Internet gambling research: Ethical and methodological issues ». *International Journal of Internet Research Ethics* 3 (12): 104-17.

LaBrie, Richard, et Howard J. Shaffer. 2011. « Identifying Behavioral Markers of Disordered Internet Sports Gambling ». *Addiction Research & Theory* 19 (1): 56-65. doi:10.3109/16066359.2010.512106.

Philander, Kahlil S. 2014. « Identifying High-Risk Online Gamblers: A Comparison of Data Mining Procedures ». *International Gambling Studies* 14 (1): 53-63. doi:10.1080/14459795.2013.841721.

Tovar, Marie-Line, Jean-Michel Costes, et Vincent Eroukmanoff. 2013. « Les jeux d'argent et de hasard sur Internet en France en 2012 ». OFDT, Tendances, , no 85: 6.

Wardle, Heather, David Excell, Eleanor Ireland, Nevena Ilic, et Stephen Sharman. 2014. « report 2 identifying problem gambling - findings from a survey of loyalty card customers ». Responsible Gambling Trust.

Wardle, Heather, Jonathan Parke, et Excell. 2014. « report 1 theoretical markers of harm for machine play in a bookmakers - a rapid scoping review ». Responsible Gambling Trust.

Jeu en ligne :
« big data » et
santé publique

Introduction

Etudes
internationales

Projet de
modèle en
France

Conclusion

Thank you for your attention

To find out more:

<http://www.economie.gouv.fr/observatoire-des-jeux>