Project in applied econometrics
Report

Lucas Javaudin, Robin Le Huérou-Kérisel, Rémi Moreau

March 2018

Abstract

This project has aimed at reproducing Moretti's 2011 paper on social learning effects in movie sales with R. We also blabla. Main results:

Contents

1 Intuitions and detailed presentation of the model

1.1 Some intuitions

1.2 Presentation of the model

bonjour je m'appelle Rémi

2 Analysis and main results

2.1 Identification of the surprises

Surprises consist in the residuals of the regression of the log-number of sales in the first week on the log-number of screens available (opened by theaters). This definition of surprises holds because we suppose that theaters are profit-maximizing agents and make use of all the available information to predict the success of a movie. If this definition is correct, we should expect log-number of screens opened by theaters first week to be a good indicator of knowledge available on the movie quality before it is released. In the Table 1 we reproduce Moretti's regression of *log_sales_first_we* on *log_screens_first_week*. Each column is the result of the regression when we control with some variables (film genre, rating available, cost, distributor, weekday, month, week, year). The fact that adding control variables doesn't change the robustness of the regression proves Moretti's point which is that theaters take into account these factors when deciding their number of available screens.

Table 1: Regression of first-weekend sales on number of screens

	Dependent variable:						
	log_sales_first_we						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log_screens_first_week	0.893***	0.896***	0.883***	0.871***	0.803***	0.806***	0.813***
	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
R^2	0.907	0.909	0.910	0.912	0.932	0.936	0.938
Adjusted R^2	0.907	0.908	0.910	0.912	0.928	0.931	0.933
Note:							*p<0.1; **p<0.05; ***p<0.01

In fact, theaters perform more than what we could do using all the data available in the data set. Regressing first week-end sales on our control variables gives us a R^2 of .7, which is smaller than .9 performed by theaters only.

is this useful?

nope, you only need the significativity of the coefficient

Table 2: Regression of first-weekend sales on control variables

	Dependent variable:
	log_sales_first_we
Observations	4,992
R^2	0.699
Adjusted R^2	0.674
Note:	*p<0.1; **p<0.05; ***p<0.01

We have performed the same kind of regression on France data from 2004 to 2008 and find quite similar results (Table 3 for France data and Table 4 for Paris data only[1]).

FRANCE

PARIS

[1]Data available for Paris are richer of 600 movies than France.

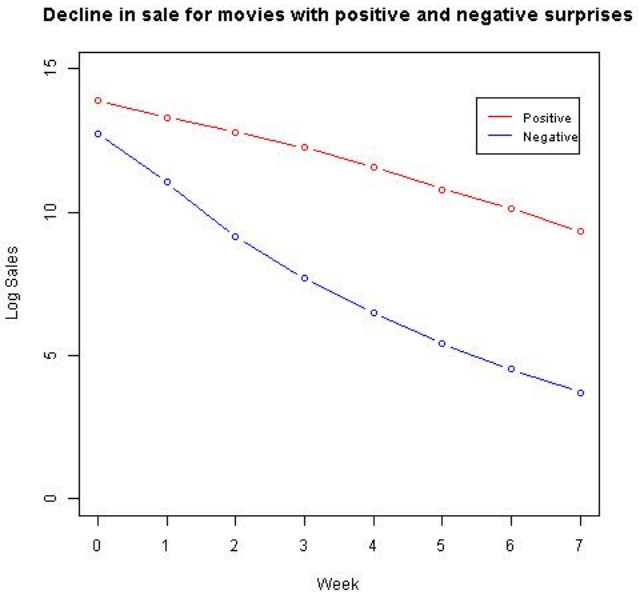
Table 3: Regression of first-week entries on number of screens for France

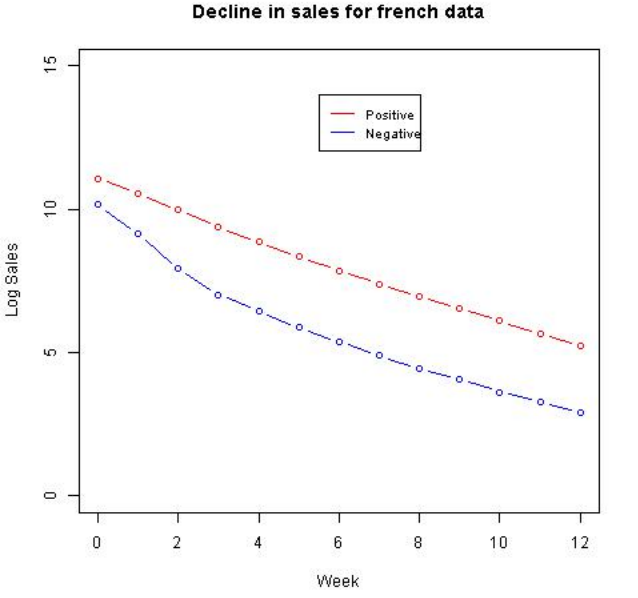
	Dependent variable:						
	log_entree_fr						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log_seance_fr	1.208***	1.237***	1.237***	1.279***	1.282***	1.287***	1.196***
	(0.009)	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)	(0.014)
Observations	2,046	2,046	2,046	2,046	2,046	2,046	2,046
R^2	0.893	0.899	0.900	0.917	0.924	0.925	0.943
Adjusted R^2	0.893	0.898	0.898	0.910	0.915	0.916	0.935
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 4: Regression of first-week entries on number of screens for Paris only

	Dependent variable:						
	log_entree_paris						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log_seance_paris	1.342***	1.336***	1.337***	1.281***	1.281***	1.284***	1.152***
	(0.010)	(0.011)	(0.011)	(0.014)	(0.014)	(0.014)	(0.014)
Observations	2,701	2,701	2,701	2,701	2,701	2,701	2,701
R^2	0.875	0.880	0.881	0.901	0.908	0.909	0.927
Adjusted R^2	0.875	0.879	0.880	0.892	0.897	0.898	0.918
Note:							*p<0.1; **p<0.05; ***p<0.01

Figure 1: We find the same graph as Moretti





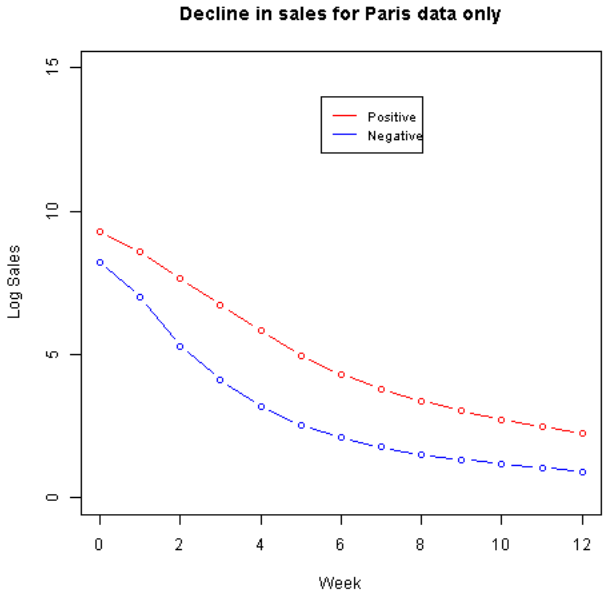


Table 5: Decline in box-office sales by opening week surprise

	Dependent variable:			
	log_sales			
	(1)	(2)	(3)	(4)
t	−0.952***	−0.952***	−1.289***	
	(0.007)	(0.006)	(0.009)	
t:surprise		0.475***		
		(0.009)		
t:positive_surprise			0.640***	
			(0.013)	
I(t ∗bottom_surprise)				−1.353***
				(0.011)
I(t ∗middle_surprise)				−1.011***
				(0.011)
I(t ∗top_surprise)				−0.491***
				(0.011)
Observations	39,936	39,936	39,936	39,936
R^2	0.772	0.788	0.787	0.790
Adjusted R^2	0.739	0.758	0.756	0.760
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 6: Precision of the prior

	Dependent variable:	
	log_sales	
	(1)	(2)
t	-1.291^{***}	-1.267^{***}
	(0.010)	(0.087)
t:positive_surprise	0.654^{***}	-0.061
	(0.013)	(0.121)
t:sequel	0.037	
	(0.038)	
t:positive_surpriseTRUE:sequel	-0.225^{***}	
	(0.053)	
t:var_surprise		-0.045
		(0.174)
t:positive_surpriseTRUE:var_surprise		1.416^{***}
		(0.243)
Observations	39,936	39,936
R^2	0.787	0.787
Adjusted R^2	0.756	0.757
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 7: Decline in box-office sales by opening week surprise

	Dependent variable:			
	log_entree_fr			
	(1)	(2)	(3)	(4)
t	−0.526***	−0.526***	−0.571***	
	(0.002)	(0.002)	(0.003)	
t:surprise		0.076***		
		(0.004)		
t:positive_surprise			0.087***	
			(0.004)	
t:bottom_surpriseFALSE				−0.459***
				(0.004)
t:bottom_surprise				−0.574***
				(0.004)
t:middle_surprise				−0.088***
				(0.005)
Observations	26,598	26,598	26,598	26,598
R^2	0.851	0.853	0.853	0.854
Adjusted R^2	0.838	0.841	0.841	0.841
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 8: Precision of the prior

	Dependent variable:		
	log_entree_fr		
	(1)	(2)	(3)
t	−0.570***	−0.698***	−0.678***
	(0.003)	(0.013)	(0.004)
t:positive_surprise	0.105***	0.109***	0.009
	(0.005)	(0.018)	(0.006)
t:saga	−0.027		
	(0.016)		
t:positive_surpriseTRUE:saga	−0.145***		
	(0.019)		
t:var_surprise		0.370***	
		(0.035)	
t:positive_surpriseTRUE:var_surprise		−0.062	
		(0.050)	
t:art_essai			0.259***
			(0.006)
t:positive_surpriseTRUE:art_essai			0.066***
			(0.008)
Observations	26,598	26,546	26,598
R^2	0.855	0.854	0.880
Adjusted R^2	0.843	0.842	0.870
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 9: Decline in box-office sales by opening week surprise

	Dependent variable:			
	log_entree_paris			
	(1)	(2)	(3)	(4)
t	-0.583^{***}	-0.583^{***}	-0.564^{***}	
	(0.002)	(0.002)	(0.003)	
t:surprise		-0.032^{***}		
		(0.004)		
t:positive_surprise			-0.039^{***}	
			(0.005)	
t:bottom_surpriseFALSE				-0.594^{***}
				(0.004)
t:bottom_surprise				-0.541^{***}
				(0.004)
t:middle_surprise				-0.021^{***}
				(0.006)
Observations	35,113	35,113	35,113	35,113
R^2	0.810	0.810	0.810	0.811
Adjusted R^2	0.794	0.794	0.794	0.795
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 10: Precision of the prior

	Dependent variable:		
	log_entree_paris		
	(1)	(2)	(3)
t	−0.560***	−0.772***	−0.616***
	(0.003)	(0.017)	(0.005)
t:positive_surprise	−0.030***	−0.213***	−0.126***
	(0.005)	(0.024)	(0.007)
t:saga	−0.118***		
	(0.017)		
t:positive_surpriseTRUE:saga	−0.022		
	(0.020)		
t:var_surprise		0.576***	
		(0.045)	
t:positive_surpriseTRUE:var_surprise		0.480***	
		(0.065)	
t:art_essai			0.087***
			(0.006)
t:positive_surpriseTRUE:art_essai			0.156***
			(0.009)
Observations	35,113	35,074	35,113
R^2	0.811	0.814	0.819
Adjusted R^2	0.795	0.798	0.804
Note:			*p<0.1; **p<0.05; ***p<0.01

2.2 Precision of the prior

2.3 Size of the Social Network

Consumers with a larger social network receive more feedbacks from their peers and thus they are able to evaluate more precisely the quality of the movie.

Table 11: Precision of peers' signal

	Dependent variable:	
	log_entree_fr	
	(1)	(2)
t	-0.663^{***}	-0.451^{***}
	(0.007)	(0.005)
t×positive_surprise	0.061^{***}	0.076^{***}
	(0.010)	(0.006)
t×tout_public	0.115^{***}	
	(0.008)	
t×positive_surprise×tout_public	0.031^{***}	
	(0.011)	
t×seance_fr_first_week		-0.033^{***}
		(0.001)
t×positive_surprise×seance_fr_first_week		0.011^{***}
		(0.001)
Observations	26,598	26,598
R^2	0.856	0.867
Adjusted R^2	0.844	0.856
Note:	*p<0.1; **p<0.05; ***p<0.01	

2.4 Does Learning Decline Over Time?

Table 12: Convexity of the sales profile

	Dependent variable:
	log_entree_fr
t	-0.978^{***}
	(0.011)
t^2	0.034^{***}
	(0.001)
$t \times$positive_surprise	0.393^{***}
	(0.016)
$t^2 \times$positive_surprise	-0.026^{***}
	(0.001)
Observations	26,598
R^2	0.861
Adjusted R^2	0.850
Note:	*p<0.1; **p<0.05; ***p<0.01

3 Conclusion: some comments