Social Learning and Peer Effects in Consumption (Moretti, 2010)

Vasily Rusanov

May 5, 2019

Motivation

- How to identify and measure social learning from the aggregate data?
- Network externalities (fashion/herding) should generate similar patterns, how to separately identify learning?

Model

- Consumer i can watch a movie j on a week $t \in 1, ..., 8$.
  Gains utility $U_{ij} = \alpha_j^* + v_{ij}$, $v_{ij} \sim N\left(0, \frac{1}{d}\right)$
- Consumer has a prior $X_j'\beta$, such that $\alpha_j^* \sim N\left(X_j'\beta, \frac{1}{m_j}\right)$
- Gets a signal $s_{ij} = U_{ij} + \varepsilon_{ij}$ at time 0 (before the movie opens),
  $\varepsilon_{ij} \sim N\left(0, \frac{1}{k_j}\right)$
- Given the signal, expected utility is $E_1\left[U_{ij}|X_j'\beta, s_{ij}\right] = \omega_j X_j'\beta + (1 - \omega_j)\, s_{ij}$,
  $\omega_j = \frac{h_j}{h_j + k_j}, h_j = \frac{dm_j}{d + m_j}$
- Cost of seeing a movie is $q_{it} = q + u_{it}$
- Given all that, the probability that the consumer sees a movie is

$$\text{Prob}\left(E_1\left[U_{ij}|X_j'\beta, s_{ij}\right] > q_{i1}\right) = \Phi\left(\frac{(1 - \omega_j)\left(\alpha_j^* - X_j'\beta\right) + X_j'\beta - q}{\sigma_{j1}}\right)$$

With learning

- Consumers observe other consumers come to the theater, know their utility
- They use this information, being aware that there is selection in going
- $E_2\left[U_{ij}|X_j'\beta, s_{ij}, S_{ij2}\right] = \frac{h_j}{h_j+k_j+z_{i2}}X_j'\beta + \frac{k_j}{h_j+k_j+z_{i2}}s_{ij} + \frac{z_{i2}}{h_j+k_j+z_{i2}}S_{ij2}$
  - S_{ij2} is the forecast from learning from the peers (see Appendix 1)
  - Note: he does explicitly model learning, but all he really needs is that more estimates are available with time
- Key prediction of the model:
  $\frac{\partial P_t}{\partial t} > 0 \quad \text{if} \quad \alpha_j^* > X_j'\beta,$
  $\frac{\partial^2 P_t}{\partial t^2} < 0 \quad \text{if} \quad \alpha_j^* > X_j'\beta$
- No proof, just "it is possible to show".
  But how, if argmax $\mathcal{L}$ has no analytical expression?

(Implicit) assumptions

- Consumers get a signal before the opening that the producers/cinemas do not get
- Consumers learn nothing themselves from actually watching the movie (that's why they keep coming with the same probability)
- Consumers learn the exact utility that others get from seeing the movie

Data

- 4992 movies, observed for 8 weeks
  - Sales, reviews, ad spending, genre
- Key variable: number of screens, incl. the opening weekend
- (unbelievable number of typos in the summary stats description)
- Another key variable: weather on the opening weekend, provides an exogenous shock to quality

Measuring surprise

- Need a measure of "surpirse": difference between the outcome and the expectations
- Assume cinemas simply seek to predict the number of consumers (no profit or anything)
- Note 13 (p368) tries to make a tricky point that cinemas put higher weight on their prior, but the trick here is that for some reason Moretti assumes that cinemas predict α_j^* rather than U_{ij}. I don't see why any of this is necessary
- Regression: first-weekend sales on number of screens
  $\hat{e_j}$ is the surprise of j, mean-zero by construction
- (do screens have different numbers of seats?)

Estimated sales prediction error and observables

- Does the number of screens really reflect the expectations about the performance of the movie?
- If yes, then the prediction error (sales not explained by the number of screens) should be orthogonal to the observables
- Mincer-Zarnowitz test: no systematic correlation between residuals and observables
- Intuitive test: adding more regressors does not change the coefficient on $\log(screens)$ or the R^2

Data is consistent with the model: sales by level of "surprise"

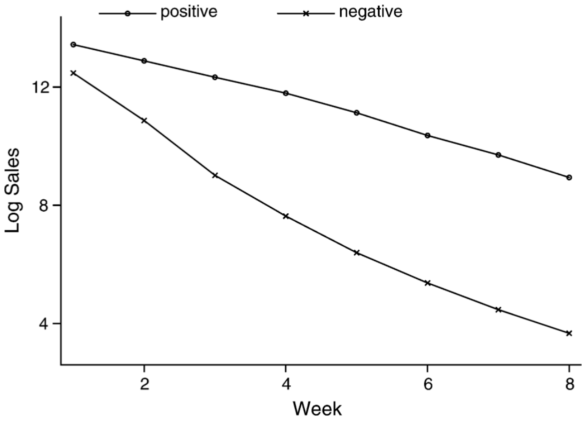

FIGURE 2

The decline in box-office sales over time by opening weekend surprise. The figure plots average log box-office sales by week for movies that experienced a positive first weekend surprise and for movies that experienced a negative first weekend surprise. Sales are in 2005 dollars. The sample includes 4992 movies

Empirical Evidence I

- The model predicts that movies that perform well in the opening will decay slower
- In the model, a good opening does not "cause" slower decay. It simply signals higher-than-expected quality
- Estimate $\ln y_{jt} = \beta_0 + \beta_1 t + \beta_2 \left(t \times S_j\right) + d_j + u_{jt}$

  y_{jt}: weekly sales
  S_j: measure of surprise (indicator, value, quanitle...)
  Instuition: good surprise means slower decline and vice versa

Results

TABLE 4

Decline in box-office sales by opening weekend surprise

	(1)	(2)	(3)	(4)
t	−0.926	−0.926	−1.258	
	(0.012)	(0.011)	(0.017)	
t × surprise		0.463		
		(0.016)		
t × positive surprise			0.616	
			(0.022)	
t × bottom surprise				−1.320
				(0.019)
t × middle surprise				−0.984
				(0.020)
t × top surprise				−0.474
				(0.017)
R^2	0.77	0.79	0.79	0.79

Notes: Standard errors are clustered by movie and displayed in parentheses. Each column reports estimates of variants of equation (15). The dependent variable is log weekly box-office sales. All models include movie fixed effects. Surprise refers to deviation from predicted first-weekend sales. By construction, surprise has mean 0. In Column 4, "bot-

Empirical Evidence II

- Advertising is a concern: what if producers spend more on movies that they are worried about?
  - Spending on advertising is significant, but does not change the main result
- Critic reviews is a concern: what if critics write about movies that are unexpectedly successful?
  - Inclusion of review does not change things significantly
- Other robustness checks: more controls in the "surprise" and sales equations
- Social learning is less important for movies with more precise signal (sequals)

Empirical Evidence III: Network effects

- What if all of it is just network effects (some movies become fashionable)?
- In that case, a negative shock to first week will make the sales decline significantly faster
- Estimate $\ln y_{jt} = \beta_0 + \beta_1 t + \beta_2 (t \times S_j) + d_j + u_{jt}$, instrumenting S_j with weather shocks
  - Problem: only observe national sales, not local. Uses 7 largest cities
  - finds that $\hat{\beta}_2^{2SLS} = 0$. But the first stage is very weak

Conclusion

- Nice model, clear identification strategy and the model-data connection
- The key result of the model comes without proofs, and I don't see how it was done
- With a week instrument, no wonder he did not find connection between weather and decline in sales over time
- Usually, the weather instrument is the deviation
  from the typical weather on a given day