Appendices

Appendix A1: Appendices to study 1.

The full industry report detailing our industry survey, inclusive of open ended comments and more descriptive tables is available online at https://foodsystems.colostate.edu/wp-content/uploads/2020/08/FF_BeerGrocery_8-26-20.pdf

Appendix A2: Appendices to study 2.

Data cleaning

Store visits are first normalized to account for growth in the SafeGraph panel of devices using the following formula:

$$NormVisits_{it} = Visits_{it} * \frac{ColoradoVisits_T}{ColoradoVisits_t} \qquad (A1)$$

where $Visits_{it}$ is the count of visits to all firms in market channel i in month t, $ColoradoVisits_t$ is the number of visits to all points-of-interests in Colorado in month t, and T is the final month of the panel.

The foot traffic data is extremely detailed but also somewhat noisy and requires cleaning. SafeGraph occasionally assigns a points-of-interest the incorrect NAICS code so we first validate the liquor stores in our data against a list of liquor stores provided by the state. Second, while opening and closures are certainly a useful indicator in their own merit, the zeros observed for not yet opened or shut down businesses could also be the result of temporary closures due to remodeling or data collection issues. Accordingly, we balanced the panel by eliminating store identifiers with no visit data for the first two or last two consecutive time periods (13% of observations), thereby eliminating store opening and closures. Consequently, our results should be interpreted as only applying to firms surviving for the entire study period and the exclusion of firms that close means our results are a conservative estimate of the effect of the policy. Third, a limited number of firms exhibit extreme, implausible variation between months that we could not attribute to seasonality. To filter outliers, we calculated the percentage change between periods[1] for each firm and dropped the firms with a percentage change outside of three standard deviations from the average per-period change. This results in a further loss of 10% of observations for liquor stores.

Table A.1 displays descriptive statistics for monthly liquor store foot traffic during each complete year of the study period and shows that, on average, foot traffic is approximately 20,000 higher in Colorado compared to Minnesota, likely due to our Colorado data containing almost 200 more firms. The difference in the level foot traffic is not a threat to our identification, provided that the trend and seasonality in the two states are comparable. The standard deviation appears similar prior to the policy change, suggesting that both time series experience similar annual variation. A concern is that the mean level of foot traffic in Colorado holds approximately constant in 2017 and 2018, the period before the policy change, but declines in Minnesota.

[1] The lowest recorded number of visits in any period is five, with lower visits counts appearing as NA (SafeGraph 2020). We assign all months with NA visits a value of one in order to calculate the percentages.

Table A.1. Aggregated monthly foot traffic by state in 2017, 2018, and 2019*

Year	Mean	Median	SD	Min	Max	Firm count
			Colorado			
2017	69,261	67,837	3,046	65,952	74,493	601
2018	69,601	68,992	2,823	61,940	72,466	601
2019	66,089	65,117	2,933	61,090	70,385	601
			Minnesota			
2017	49,404	48,202	3,319	45,736	57,282	408
2018	48,384	47,664	2,844	45,214	53,794	408
2019	47,018	46,016	3,826	41,323	54,192	408

*Visits are normalized to account for growth in the SafeGraph panel.

State space forecasting methodology

Following Hyndman et al. (2008), a general state space model can be expressed as

$$y_t = w(x_{t-1}) + r(x_{t-1})\epsilon_t \qquad (A2)$$

$$x_t = f(x_{t-1}) + g(x_{t-1})\epsilon_t \qquad (A3)$$

where y_t is the observed value at time t and $x_t = (l_t, b_t, s_{t-1}, \ldots, s_{t-m+1})$ is a state vector containing equations for l_t, b_t, and s_t, which denote the level, slope, and seasonal components at time t. ϵ_t, called innovations in the literature, is a normally distributed white noise process with variance σ^2. The first term of equation $(A2)$ captures the effect of past observations on y_t, whereas the first term in $(A3)$ describe how the state vector evolves over time.

We examine the decomposition of our time series in the pre-policy period, provided in figure A.1, to select an appropriate forecast model.

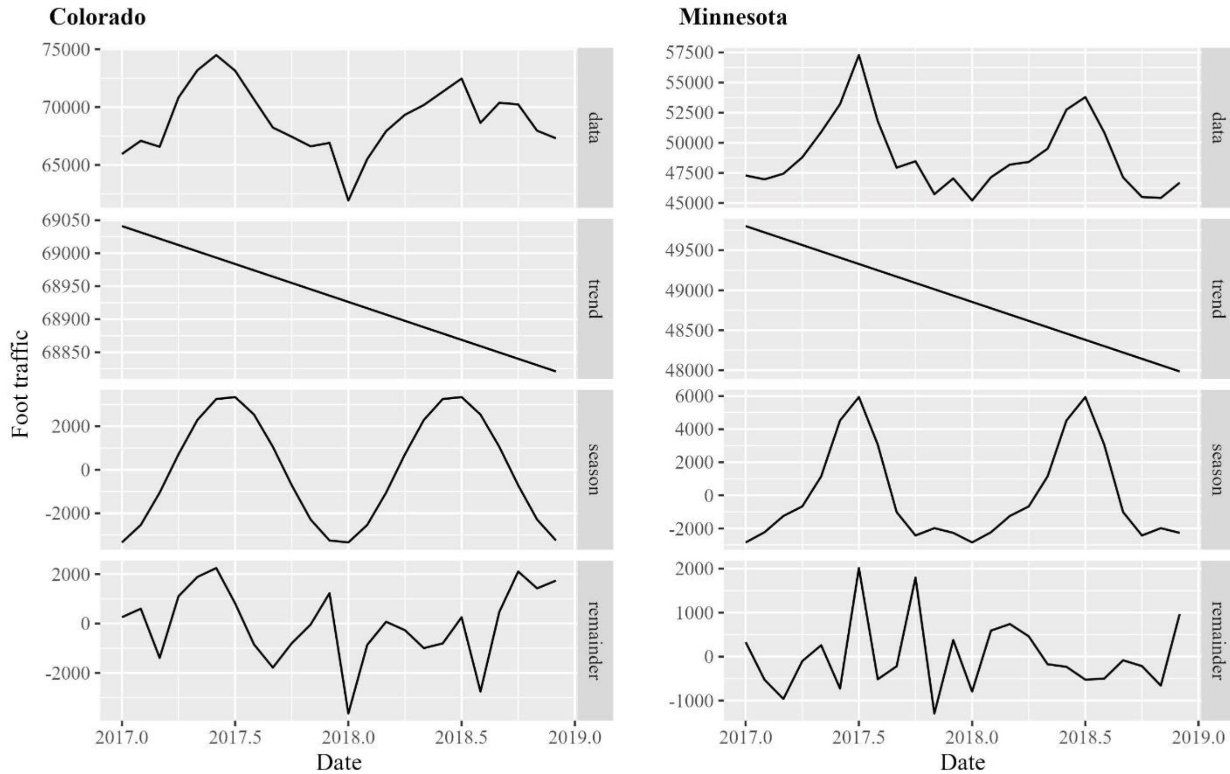

Figure A.1. Decomposition of Colorado and Minnesota time series pre-policy change. We decompose the time series, shown in the top row, into trend and seasonal components using a linear regression model with trend and fourier terms. The optimal number of fourier terms (1 for Colorado and 3 for Minnesota) was determined based on the AICc.

As expected, the decomposition reveals that both time series have a clear seasonal pattern as well as a negative trend, with the trend more pronounced in Minnesota. Note that a fourier transformation assumes seasonality to be fixed. Based on the decomposition, we determine that a model with an additive seasonal component and possibly a trend component is appropriate. Hyndman et al. (2008) suggests that a seasonal model without a trend component may improve forecasts when dealing with shorter time series owing to the simple state space structure of the data. Due to the limited number of observations, we compare the possible models using the AICc (Table A.2) and opt to omit the trend component.

Table A.2. AICc for state space models with and without trend

Model	AICc	
	Colorado	Minnesota
With trend	563.90	532.53
Without trend	511.79	494.32

Consistent with a seasonal model with multiplicative errors, our state space equations take the following form:

$$\mu_t = l_{t-1} + s_{t-m} \qquad (A4)$$

$$l_t = l_{t-1} + \alpha(l_{t-1} + s_{t-m})\epsilon_t \qquad (A5)$$

$$s_t = s_{t-m} + \gamma(l_{t-1} + s_{t-m})\epsilon_t \qquad (A6)$$

Where $\mu_t = \hat{y}_t$, α and γ are estimated smoothing parameters, m is an index for the months in a year, $y_t - \hat{y}_t = \epsilon_t$, and $\epsilon_t \sim NID(0, \sigma^2)$. We estimate values for the smoothing parameters $\theta = (\alpha, \gamma)$ and initial states $x_0 = (l_0, s_0, s_{-1}, \dots, s_{-m+1})$ using observations in the pre-policy period using Maximum Likelihood estimation:

$$\mathcal{L}(\theta, x_0) = n \log(\textstyle\sum_{t=1}^{n} \epsilon_t^2) + 2 \sum_{t=1}^{n} \log \ |r(x_{t-1}). \quad (A7)$$

We use our estimates of l_t and s_t, as well as α, and γ, to generate a h-step-ahead point forecast for the post-policy period using the following set of equations:

$$\hat{y}_{t+h} = l_t + s_{t-m+h_m^+} \qquad (A8)$$

$$l_t = \alpha(Y_t - s_{t-m}) + (1-\alpha)l_{t-1} \qquad (A9)$$

$$s_t = \gamma(Y_t - l_{t-1}) + (1-\gamma)s_{t-m} \qquad (A10)$$

where $h_m^+ = [(h-1) \mathrm{mod}\ m] + 1$.

Following Ord et al. (1997) and Hyndman et al. (2002), we calculate a prediction interval for the point estimates by simulating multiple ($M = 5{,}000$) forecasting paths conditional on the final (pre-treatment) state x_n and random draws of the disturbance, and identify the 0.025 and 0.975 quantiles of the simulated values.

ITSA results visualization

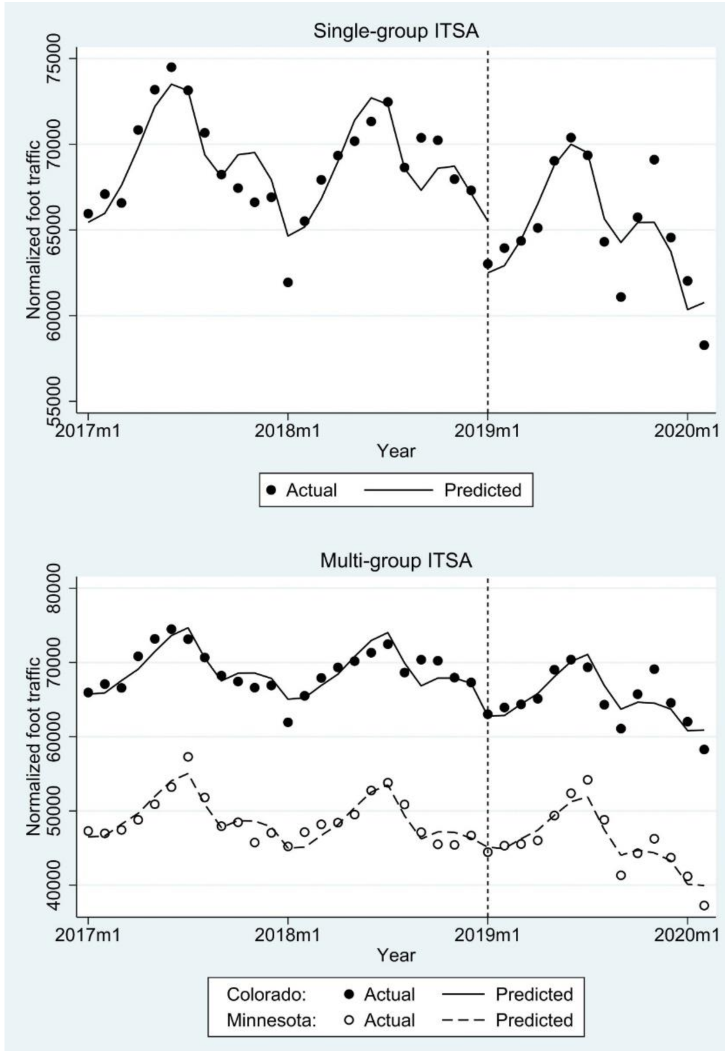

Figure A.2. Estimated and actual visits to liquor stores in single- and multi-group ITSA. The vertical dashed line indicates that the policy change took effect January 1st, 2019. Observed values appear as points along with a line that represents the fitted values from our regression.

Figure A.2 provides a visual representation of the ITSA models estimated according to equation (1) and (2). Liquor store foot traffic shows seasonal variation with two peaks coinciding with the summer and the holiday season. The level of foot traffic appears to decrease slightly in Colorado immediately following the policy change and remains consistent in Minnesota.

References

Hyndman, Rob J, Anne B Koehler, Ralph D Snyder, and Simone Grose. 2002. "A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods." *International Journal of Forecasting* 18 (3): 439–54. https://doi.org/10.1016/S0169-2070(01)00110-8.

Ord, J. K., A. B. Koehler, and R. D. Snyder. 1997. "Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models." *Journal of the American Statistical Association* 92 (440): 1621–29. https://doi.org/10.1080/01621459.1997.10473684.

SafeGraph. 2020. "Core Patterns." https://docs.safegraph.com/docs/monthly-patterns.