Forecasting Sales: A Model and Some Evidence from the Retail Industry

Russell Lundholm
Sarah McVay
Taylor Randall
Why forecast financial statements?

- Seems obvious, but two common criticisms:
  - “Who cares, can’t we can look at the analyst forecasts?”
  - “Surely in the ‘real world’ analysts use much more sophisticated models than we could ever imagine.”
Why forecast sales?

- The sales forecast is the most important forecast in most financial forecasting applications.
  - Combine with margin to get income forecast.
  - Combine with turnover to get asset forecast.

- A 1% change in the sales growth estimate can change the equity value estimate by about 15%.

- Our goal is to:
  - Develop a reasonably general sales-forecasting model.
  - Show how to estimate it on a retail firm.
  - Test its predictive ability in-sample and out-of-sample.
  - Compare it to analyst forecasts.
What needs to be modeled?

- how do we mix together beliefs about
  - the number of new sales-generating units, and
  - the growth in sales from existing units?

- Half of the puzzle - # of stores - is commonly estimated/disclosed by the firm.
  - what does this imply about the sales forecast?

- Comp growth rates are more complicated than you might have thought
The Model – sales rates / store

\[ R_t^N = \text{the average sales/store rate for the new stores in year } t, \text{ (annualized in results)} \]

\[ R_t^M = \text{the average sales/store rate for the mid stores in year } t, \text{ and} \]

\[ R_t^O = \text{the average sales/store rate for old stores in year } t. \]

- \[ Sales_t = O_t R_t^O + M_t R_t^M + N_t R_t^N \]
Comparable store sales growth is the growth in sales from stores that were open at the beginning of the prior year and are currently still open.
The algebra of comps

1 + C_t = \frac{O_t R_t^O}{(O_{t-1} - D_t) R_{t-1}^O + M_{t-1} R_{t-1}^M}

- Recall that \( O_t = O_{t-1} + M_{t-1} - D_t \) so same stores are in numerator and denominator.

- Note that comps are sensitive to the movement of stores from mid to old (if the rates are different).

- If \( R_t^M = R_t^O \) each period then \( 1 + C_t = \frac{R_t^O}{R_{t-1}^O} \).
Using the comp growth to “clean up” the estimation

\[ Sales_t = O_t R_t^O + M_t R_t^M + N_t R_t^N \]

we know the # of stores of each type

but the rates are unknown. However, we know something about the evolution of the rates over time.

The trick in the estimation is to account for the changes in past rates as part of the independent variable so that we can estimate the most recent rates as parameters.
The Model – something we can estimate?

\[ \Delta \text{Sales}_t = (O_t R^O_t - O_{t-1} R^O_{t-1}) + (M_t R^M_t - M_{t-1} R^M_{t-1}) + (N_t R^N_t - N_{t-1} R^N_{t-1}) \]

- observe all the # of stores and control for the changing rates over time using the comps (plus some assumptions).

- Estimate \( R^N_T R^M_T R^O_T \) where T is most recent year in data.

- Estimate by firm over time. Estimate in changes to control for non-stationarity.
Model 1

- after the end of first year, store is fully mature
- all rates evolve at the rate \((1+C_t)\)

\[
\Delta Sales_T = \left( O_T + M_T \right) R_T^O - \left( O_{T-1} + M_{T-1} \right) R_{T-1}^O + \left[ N_T R_T^N - N_{T-1} R_{T-1}^N \right]
\]

\[\Delta Sales_T = \left[ \left( O_T + M_T \right) - \frac{\left( O_{T-1} + M_{T-1} \right)}{1 + C_T} \right] R_T^O + \left[ N_T - \frac{N_{T-1}}{1 + C_T} \right] R_T^N\]

know that \(R_t^O = R_{t-1}^O (1 + C_t)\)

assume that \(R_t^M = R_t^O\)

and \(R_t^N = R_{t-1}^N (1 + C_t)\)
Model 1 – its nicer than it looks!

\[ \Delta Sales_{T-\tau} = \left[ \frac{(1 + C_{T-\tau})(O_{T-\tau} + M_{T-\tau}) - (O_{T-\tau-1} + M_{T-\tau-1})}{\prod_{i=T-\tau}^{T} (1 + C_i)} \right] R_T^O + \left[ \frac{(1 + C_{T-\tau})N_{T-\tau} - N_{T-\tau-1}}{\prod_{i=T-\tau}^{T} (1 + C_i)} \right] R_T^N \]

- change in # of stores plus comp effect in current year
- bring prior year data to year T basis by deflating with historical comps
- everything in brackets is known so can estimate the rates.
Model 2

\[ R_t^M = kR_t^O \]

\[ R_t^N = R_{t-1}^N (1 + C_t) Q_t, \text{ where} \]

\[ Q_t = \frac{O_{t-1} + kM_{t-1} - D_t}{O_t} \]

- mid rates and new rates differ by proportion \( k \)
- all rates evolve at the rate \( (1 + C_t)Q_t = \frac{R_t^O}{R_{t-1}^O} \)
  - if \( k > 1 \) (so \( Q > 1 \)) then prolonged new store “honeymoon period”.
  - if \( k < 1 \) (so \( Q < 1 \)) then prolonged time to maturity.
models 3-6

- **model 3**: \( \%\Delta \text{Sales}_{T-\tau} = (1+C_{T-\tau})(1+G_{T-\tau})-1 \)
  - assumes \( R_t^M = R_t^O = R_t^N \)

- **model 4**: \( \Delta \text{Sales}_{T-\tau} = \text{Sales}_{T-\tau-1}SG_j \) where \( SG_j \) is historical sales growth for that decile (Nissim/Penman 2001)

- **model 5**: \( \Delta \text{Sales}_{T-\tau} = \beta(\Delta \text{ total \# stores}) + \varepsilon \)

- **model 6**: \( \Delta \text{Sales}_{T-\tau} = \gamma_1\Delta \text{ existing stores} + \gamma_2\Delta \text{ new stores} + \varepsilon \)
The Sample

- 87 retail firms (1036 firm-years) with at least 6 years of data on
  - sales
  - number of stores at year end
  - stores opened during the year
  - stores closed during the year
  - comparable store growth rate for the year
  - expected number of store openings/closings for the following year
## Summary Regression Results (estimated firm-by-firm)

<table>
<thead>
<tr>
<th>Models</th>
<th>Median p values for Estimated Sales Rates</th>
<th>Percent Positive Sales Rates</th>
<th>Mean(Median) $k \in (0.8, 1.2)$</th>
<th>Median Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: New/Old, $k = 1$ (Full Sample)</td>
<td>(0.001) (0.014)</td>
<td>98.9% 92.0%</td>
<td>na</td>
<td>91.7%</td>
</tr>
<tr>
<td>Model 2: New/Old, $k \in (0.8, 1.2)$</td>
<td>(0.001) (0.021)</td>
<td>98.9% 92.0%</td>
<td>1.0(1.0)</td>
<td>92.8%</td>
</tr>
</tbody>
</table>
## In-Sample comparison of models

<table>
<thead>
<tr>
<th></th>
<th>Model 1 (k=1)</th>
<th>Model 2 (k \in (.8, 1.2))</th>
<th>Model 3 ((1+G)(1+C)-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median of Median</td>
<td>2.49%</td>
<td>1.99%</td>
<td>3.24%</td>
</tr>
<tr>
<td>[Residual]/Sales</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Model 4 Mean Reversion</th>
<th>Model 5 ChgStores</th>
<th>Model 6 ChgNew/Old Stores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median of Median</td>
<td>4.50%</td>
<td>4.61%</td>
<td>4.04%</td>
</tr>
<tr>
<td>[Residual]/Sales</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
examples

<table>
<thead>
<tr>
<th></th>
<th>old rate</th>
<th>new rate</th>
<th>k</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starbucks</td>
<td>1.18</td>
<td>0.66</td>
<td>k=1</td>
<td>98.9</td>
</tr>
<tr>
<td>Whole Foods</td>
<td>29.85</td>
<td>43.33</td>
<td>k=1</td>
<td>97.5</td>
</tr>
<tr>
<td>Walmart</td>
<td>111.65</td>
<td>322.74</td>
<td>k=1.2</td>
<td>94.0</td>
</tr>
</tbody>
</table>
building an out-of-sample forecast

- need the company forecast of next year’s # of new stores (already know old and mid)

- need an estimate of the old rate and new rate based on past data (use past 5 years)

- need an estimate of next year’s comp
  - after much searching, best model is simply
  - \( C_t - .02 = b(C_{t-1} - .02) \).
out-of-sample results

<table>
<thead>
<tr>
<th>Foresight Assumption</th>
<th>Model 1: k = 1</th>
<th>Model 2: k ∈ (.8, 1.2)</th>
<th>Model 3: (1+g)(1+c) – 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect Foresight for both Comp Growth and Change in Stores</td>
<td>2.39%</td>
<td>2.80%</td>
<td>2.95%</td>
</tr>
<tr>
<td>Company Forecasts of Change in Stores and Comp Growth = b*(C_{t-1} - .02)</td>
<td>3.79%</td>
<td>3.88%</td>
<td>3.81%</td>
</tr>
</tbody>
</table>
model 1 versus analyst forecasts

Median IBES analyst absolute forecast error is 3.20% of sales.
Median model 1 absolute forecast error is 3.79% of sales.
incremental value relative to analyst forecasts

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient (t-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>68.02 (1.50)</td>
</tr>
<tr>
<td>Model 1 Change in Sales Estimate</td>
<td>0.069 (6.09)</td>
</tr>
<tr>
<td>I/B/E/S Change in Sales Estimate</td>
<td>0.743 (45.34)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>94.34%</td>
</tr>
<tr>
<td>Number of firm-year observations</td>
<td>439</td>
</tr>
</tbody>
</table>
Financial Statement Analysis and Valuation: Forecasting Firm and Industry Fundamentals

October 22-23, 2010

Mendoza College of Business  The University of Notre Dame

Submission Deadline: May 5, 2010