Bret Allinott April 2022

# Forecasting Canadian Investment: Replicating and Updating

## 1 Introduction

In a paper from 2007, Islam utilizes the historical flexible accelerator model ('FAM'), an empirical model of aggregate investment and adds time series components.[?] In 1957, Koyck proposed the FAM,[?]

$$I_t = K_t - K_{t-1} = \eta[K_t^* - K_{t-1}]$$

Where $\eta \in (0, 1]$ and $K_t^*$ is the 'desired' level of capital. Essentially, investment each period is some proportion of the gap between the desired level, and last period's level of capital. In 1971, Joregensen showed that you can write the desired level of capital, $K_t^*$, as a fraction of output, such that $K_t^* = \phi Y_t$.[?] Rewriting the FAM gives,

$$I_t = \eta\phi Y_t - \eta K_{t-1}$$

Which I can write in the regression form,

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 K_{t-1} + \varepsilon_t \tag{1}$$

Islam gets their capital and GDP series from Statistics Canada from 1961 to 2000. Both series are in 1992 constant dollars, and $I_t$ is derived from the capital formation, $I_t = K_t - K_{t-1}$.

Islam estimates this model, and then estimates an ARIMA model separately for the error term. In other words, they estimates a MARIMA model, or a regression with ARIMA errors. They specify an ARIMA(8,0,3) for the FAMs error term $\varepsilon_t$, or,

$$(1 - \phi_1 L - ... - \phi_8 L^8)\varepsilon_t = (1 + \theta_1 L + ... + \theta_3 L^3)u_t \tag{2}$$

This paper will replicate their data table to ensure I have similar data, their structural regression (equation 1), and their MARIMA model (equation 2). Unfortunately, Islam's forecasting results are not available directly in the paper, but I will utilize their model to forecast regardless. After replicating their results, I will construct an additional three forecasting models; a FAM with seasonal MARIMA errors ('SMARIMA'), a univariate seasonal ARIMA model ('SARIMA'), and a VAR model. I will evaluate these models by their ability to forecast the 2000 to 2004 period.

## 2 Literature Review

Kopcke, in 1993 used several theoretical, structural models to forecast investment.[?] In particular, they used the accelerator model, with the addition of lagged GDP.

$$I_t = a + \sum_{i=0}^{p} b_i Y_{t-i} + cK_{t-1}$$

They also used another structural alternative, the neoclassical formation of investment,

$$I_t = a + \sum_{i=0}^{p} b_i \frac{Y_{t-i}}{UCC_{t-i-1}} + \sum_{i=0}^{p} c_i \frac{Y_{t-i}}{UCC_{t-i}} + dK_{t-1}$$

Where UCC is the user cost of capital. Finally, moving away from structural dynamics, they considered the simple AR model for investment.

$$I_t = a + \sum_{i=1}^{p} b_i I_{t-i}$$

Their data runs from quarter one 1962 to quarter one 1992. They separately forecast the investment in equipment and non-residential structures. Concerning equipment, the autoregressive model outperforms both structural models, with the neoclassical formation outperforming the modified FAM. Finally, for non-residential structures, the neoclassical model nearly outperforms the AR models in some subsamples but can perform relatively poorly in others. For instance, in a quarter one 1978 to quarter one 1992 sample, the accelerator model outperforms the neoclassical method, while both fail to outperform the AR model.

This paper demonstrates that good structural models can compete with simple time series models. However, the definition of capital, and the sample period used can change how well each structural model performs. In contrast, time series methods do not seem to fluctuate as much concerning capital and sample period definition.

## 3 Data

To properly replicate Islam's paper, I must replicate their capital, GDP, and investment series. As outlined in their data section, they use 'the business and government fixed capital formation,' at '1992 constant prices.' I obtained these series from Statistics Canada and summed them to get my capital series.[?] From their graphs of the series, they are not seasonally adjusted, so I take the unadjusted form as well. The options for prices are 2012 chained dollars and current prices. It is possible that at the time of this paper's release in 2007, there was a 1992 chained dollars (or equivalent) option present. However, since this is not the case now, I will use nominal prices and manually put them in 1992 dollars.

I use the expenditure-based GDP at final market prices, unadjusted, as my GDP series.[?] Similar to the capital stock series, the options provided are 2012 chained dollars or nominal prices, and again I take the nominal series and convert it manually to 1992 dollars.

Since I derive investment from the capital series, I have the two series I need, and I need to put them in 1992 dollars. Since they say '1992 dollars', I looked for annual price indices. I found one for all items, with 1992 as the base year.[?] This initially seemed the most promising, so I used this CPI to normalize GDP and capital to 1992 prices. However, when I proceeded to replicate summary statistics, these series differed from the series present in Islam's paper. I denote this nominal to real conversion as 'method one.'

There are many subsections of business and government capital formation. Subcategories include construction, machinery and equipment, and intellectual property. No combination of these subsections performed exceptionally well using method one. Therefore, I considered alternative price indices. In particular, Statistics Canada's implicit price index for GDP and government and business fixed capital formation.[?] These series are quarterly and do not have 1992 as the base year. To convert the base year, I computed the average CPI for 1992, divided the series' by this average, and multiplied by 100. I denote using individual, implicit price indices as method two.

## 4 Presentation of Replication Results

To properly replicate Islam's paper, my series must represent theirs. Therefore, I begin by replicating their summary statistics. Table 1 presents the mean and standard deviation for each series for each ten-year subsample. The table includes the values from Islam's original paper and the values from my series, using both method one and method two to convert to real terms.

Regarding proximity to the original paper for GDP, method two seems to outperform method one consistently. Likewise, method two provides a much better series of capital. However, it is still consistently off by approximately 3000 but is relatively close in the 90s, as is method one. Investment is tricky, as method 2 matches the data well in the 70s but performs relatively poor elsewhere, particularly in the 90s (where capital did well).

| Summary Statistic | Time Period | Original Paper | Method 1 | Method 2 |
|---|---|---|---|---|
| GDP (SD GDP) | 1961-1970 | 72495.82 (11300.71) | 74183.45 (13244.06) | 75552.81 (11877.53) |
| Capital (SD Capital) | 1961-1970 | 10592.23 (1962.95) | 16843.85 (3581.05) | 13729.85 (2686.80) |
| Investment (SD Investment) | 1961-1970 | 176.36 (1300.61) | 327.96 (2153.81) | 257.80 (1654.50) |
| GDP (SD GDP) | 1971-1980 | 114707.45 (14390.64) | 125807.37 (20433.50) | 118527.86 (16285.85) |
| Capital (SD Capital) | 1971-1980 | 17536.90 (3244.56) | 28885.74 (5867.75) | 21595.51 (3705.57) |
| Investment (SD Investment) | 1971-1980 | 261.53 (2196.21) | 399.73 (4435.10) | 248.78 (2789.00) |
| GDP (SD GDP) | 1981-1990 | 155196.68 (16757.47) | 165647.12 (15872.83) | 158875.22 (15601.76) |
| Capital (SD Capital) | 1981-1990 | 27640.33 (4978.39) | 36508.51 (4924.83) | 30804.38 (4753.59) |
| Investment (SD Investment) | 1981-1990 | 211.63 (3006.45) | 179.2 (4502.04) | 313.39 (3239.88) |
| GDP (SD GDP) | 1991-2000 | 196561.90 (19704.60) | 197943.23 (17835.49) | 197113.23 (17552.01) |
| Capital (SD Capital) | 1991-2000 | 37171.68 (6917.33) | 38895.34 (4708.39) | 38465.37 (4717.14) |
| Investment (SD Investment) | 1991-2000 | 493.98 (3343.22) | 78.53 (3766.90) | 210.38 (3442.69) |

**Table 1:** Mean and standard deviations of relevant series. Data sources come from original paper, and my own using two different price indices. Capital stock is the sum of government and business fixed capital formation, and GDP is expenditure based GDP, both in 1992 prices. Investment is derived as the change in capital. Method one uses an all item CPI. Method two uses GDP and capital specific CPIs.

Overall, method two tends to perform better. While both series get close for capital and GDP in the 90s, method two is more consistently close in the other subperiods and in investment.

Since method one performs poorly, I think it is unlikely that Islam obtained these series in nominal terms and put them in real terms himself. Instead, I believe when Islam wrote the paper in 2007, these series were available in 1992 dollars of some form. These 1992 dollars, similar to method two,

| | Investment | | |
|---|---|---|---|
| | (Original Paper) | (Method One) | (Method Two) |
| Intercept | $-4649.21^{***}$ | $2621.900^{***}$ | $-2717.976^{***}$ |
| | (661.8) | (739.426) | (465.319) |
| $Y_t$ | $0.148^{***}$ | $0.144^{***}$ | $0.211^{***}$ |
| | (0.015) | (0.013) | (0.013) |
| $K_{t-1}$ | $-0.655^{***}$ | $-0.756^{***}$ | $-1.004^{***}$ |
| | (0.068) | (0.063) | (0.062) |
| $R^2$ | 0.378 | 0.486 | 0.636 |
| DW | 1.58 | 1.39 | 1.22 |
| F Statistic | $47.08^{***}$ | $71.928^{***}$ | $132.763^{***}$ |
| *Note:* | | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

**Table 2:** OLS results of equation 1. $Y_t$ is GDP in 1992 dollars, $K_{t-1}$ is lagged fixed capital formation in 1992 dollars. DW is the Durbin-Watson statistic. Method one uses an all-item CPI, and method two uses a item-specific CPI.

would be uniquely weighted and not based on an overall, all-item type index as used in method one.

Next, I replicate their OLS model, and table 2 presents the results. Here method one performs quite well to match coefficients on the series but has a large, positive intercept. Method two does not match the coefficients as well but does have the large negative intercept from Islam's model. However, the coefficient on $K_{t-1}$ in method two suggests that $\eta = 1.004$, which exceeds its theoretical bounds. In methods one and two, both capital and GDP are the same nominal series and differ only in their conversion to real terms. This seemingly small difference has led to a significant difference in their estimation regarding the intercept. Overall, it is unclear which data performs better. Method two provides a better intercept but estimates parameters exceeding their theoretical bounds, and method 2 provides parameters in their theoretical bounds, but with a positive intercept.

| | Investment | | |
| --- | --- | --- | --- |
| | (Original Paper) | (Method 1) | (Method 2) |
| Intercept | $-5631.08^{***}$ | 1906.707 | $-1650.8129^{***}$ |
| | (677.42) | (2498.111) | (862.570) |
| $K_{t-1}$ | $-0.7167^{***}$ | $-0.8844^{***}$ | $-0.5998^{***}$ |
| | (0.070) | (0.1144) | (0.2400) |
| $Y_t$ | $0.16587^{***}$ | $0.1755^{***}$ | $0.1270^{***}$ |
| | (0.0159) | (0.0234) | (0.0495) |
| $\phi_1$ | 0.0805 | −0.0139 | −0.0473 |
| $\phi_2$ | 0.4572 | 0.3387 | 0.2624 |
| $\phi_3$ | −0.1332 | 0.0113 | −0.1806 |
| $\phi_4$ | 0.6664 | 0.8161 | 0.6439 |
| $\phi_5$ | −0.0889 | −0.0460 | −0.0630 |
| $\phi_6$ | −0.4401 | −0.3825 | −0.3627 |
| $\phi_7$ | 0.1155 | −0.0531 | 0.0658 |
| $\phi_8$ | 0.2836 | 0.1131 | 0.2263 |
| $\theta_1$ | 0.5285 | 0.8147 | 0.3924 |
| $\theta_2$ | 0.1225 | 0.5477 | 0.1661 |
| $\theta_3$ | 0.5685 | 0.4276 | 0.4474 |
| $R^2$ | 0.935 | 0.978 | 0.973 |
| *Note:* | | | $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$ |

**Table 3:** MARIMA(8,0,3) results for original author (Islam), and two replicating data sets. Method 1 uses an all-item CPI to normalize prices to 1992 dollars. Method 2 utilizes an item-specific CPI. No standard errors are given for the ARMA parameters, as they are not provided in the original paper.

Islam's coefficients for their MARIMA(8,0,3) model are presented in table 3, along with my coefficients for the same model, using method one and method two. Method two performs better here, as the introduction of the ARIMA errors brings coefficients closer to the original paper, except for the intercept, which moves in the opposite direction as their.[1] For the most part, both methods get reasonably close to the original paper concerning the ARIMA coefficients. Overall, method two once again outperforms method one.

# 5 Presentation of New Results

For this section, I will be strictly using the method 2 dataset, as it has outperformed method one for the most part, and is therefore more representative of the data Islam had. I present 3 new forecasting models. First, I estimate a MARIMA model with seasonal components in the error term ('SMARIMA'), then a univariate seasonal model ('SARIMA'), and finally a VAR model. Sections 5.1 through 5.2 discuss the model estimation procedure, and section 5.3 presents forecasting results.

## 5.1 Seasonal Models

Since Islam utilizes seasonally unadjusted data, the most natural extension is to add seasonal terms to their MARIMA model ('SMARIMA'). Given the quarterly frequency of my data, the SMARIMA(p,d,q)(P,D,Q)$_s$ model for my regressions error term, $\varepsilon_t$, can be written as,

$$\underbrace{\phi_p(B)}_{AR}\,\underbrace{\Phi_P(B^4)}_{SAR}\,\underbrace{(1-B)^d}_{\text{Differencing}}\,\underbrace{(1-B^4)^D}_{\text{Seasonal Diff.}}\,\varepsilon_t=\underbrace{\theta_q(B)}_{MA}\,\underbrace{\Theta_Q(B^4)}_{SMA}\,u_t \tag{3}$$

I used a cross-validated exhaustive search to select the appropriate SMARIMA model orders. I set the max order for $p$ and $q$ to five, $P$ and $Q$ to two, and $d$ and $D$ to one.[2] Models were trained on the 1961 to 1999 dataset and forecasted five years, from 2000 to 2004. I obtained the squared

[1] IE, my intercept gets closer to zero and their gets further from zero.

[2] Therefore, I trained and tested $6^2 * 3^2 * 2^2 = 1296$ models.

residuals between these forecasts and the actual value for each model and computed their mean squared error. The model with the lowest mean squared error, a SMARIMA(0,0,0)(1,1,1)$_4$ model, was selected as the optimal model. Interestingly, cross-validation selected no non-seasonal ARIMA parameters. Instead, the model seems focused on removing seasonal effects from the error term.

Utilizing the same procedure as above, I performed exhaustive search cross-validation to determine the optimal univariate seasonal ARIMA ('SARIMA') investment model. This formulation is functionally identical to equation 3 but uses SARIMA parameters on the investment itself, not on a linear regressions error term,

$$\phi_p(B)\Phi_P(B^4)(1-B)^d(1-B^4)^D I_t = \theta_q(B)\Theta_Q(B^4)\varepsilon_t$$

Cross-validation chose the SARIMA(0,1,1)(0,1,1)$_4$ model, with no AR parameters selected. However, this model selects non-seasonal differencing and seasonal differencing. This is particularly interesting since $I_t$ is already a differenced series, $I_t = K_t - K_{t-1}$.

### 5.2 VAR Model

Each VAR is estimated equation by equation via OLS. There is perfect multicollinearity in the investment equation with both capital and investment present.

$$I_t = \Gamma_{11} \underbrace{I_{t-1}}_{=K_{t-1}-K_{t-2}} + ... + \Gamma_{21}K_{t-1} + \Gamma_{22}K_{t-2} + ...$$

There is no multicollinearity under equation (1), but adding lagged values of investment and capital introduces this issue. Therefore, I remove capital and instead only use investment and GDP to construct the VAR. The AIC and BIC select 8 and 5 lags, respectively, using in-sample selection criteria to determine the optimal lag length.

However, as an alternative and given how I determined the optimal models in section 5.1, I also used cross-validation as a criterion to select the optimal lag length. Since the data is not seasonally adjusted, I added seasonal dummies. However, the best model with seasonal dummies had a mean squared error approximately 100,000 higher than the best model without seasonal dummies. I seasonally differenced the data and reran this model, both with and without seasonal dummies. These two models did not provide any increases in mean squared error and were therefore not utilized.

I trained a VAR(1) through VAR(10) on the 1961 to 1999 dataset, forecasted five years, and calculated mean squared errors. Figure 1 presents the mean squared error for each lag order.[3] Cross-validation selects the VAR(6) model. The VAR(6) has a noticeably smaller mean squared error than the VAR(5), 739321.8 versus 817940.0. Therefore, I select the VAR(6) model as my optimal model since I care primarily about forecasting ability. The VAR(5) model is more parsimonious than the VAR(6), but the VAR(6) has more predictive power.

Finally, to make gains over the univariate case, GDP should be useful in helping predict investment. I can test the usefulness of GDP in predicting investment by a Granger-causality test. Theoretically, GDP should help predict investment, as investment is a portion of output due to the national accounting identity.

[3]Note this was not done for models in section 5.1, as they were varying across multiple dimensions, and the optimal VAR only varies along one.

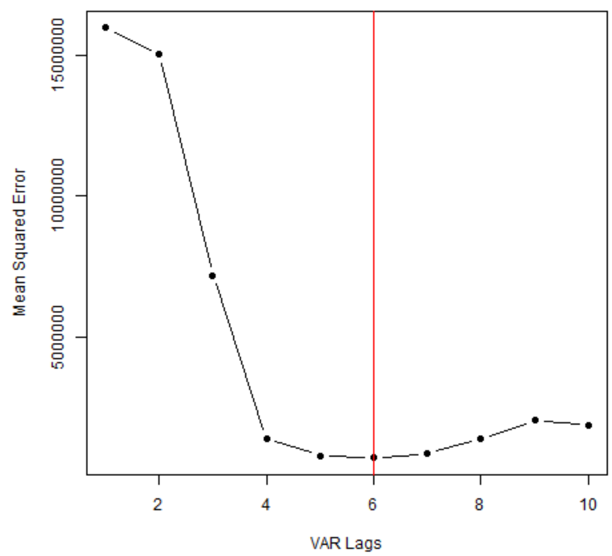


**Figure 1:** Mean squared error for the VAR(1) through VAR(10) model. Model is trained on the 1961-1999 data and mean squared error is calculated on its forecasts of 2000-2004. Lowest mean squared error is given by the VAR(6) model.

| | Granger Causality Test | |
|---|---|---|
| | F-Statistic | p-value |
| $H_0$: GDP does not Granger-cause investment | 3.8443 | 0.002199 |

**Table 4:** Granger causality test for the null hypothesis that GDP does not granger cause investment. Results come from the estimated VAR(5) model with GDP and investment as the only two variables.

Table 4 presents the results that GDP granger-causes investment. I reject the null hypothesis that GDP does not Granger-cause investment at the 99% significance level. Therefore, GDP should add predictive power to my VAR investment forecast and should ideally provide gains over the univariate case.

### 5.3 Forecasting Results

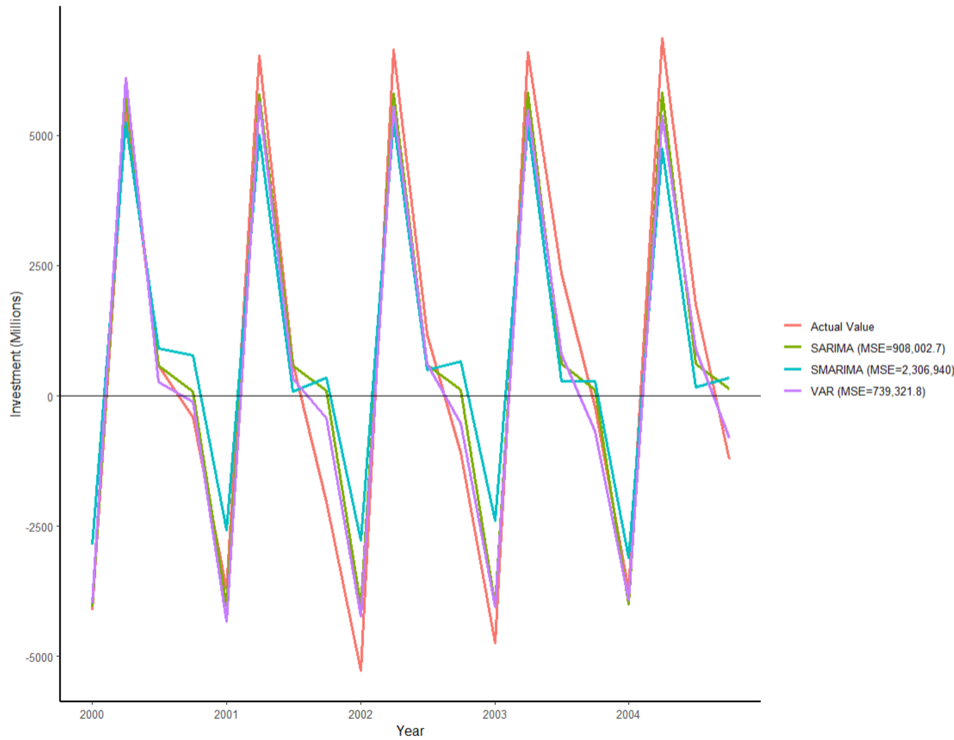


**Figure 2:** Forecasted values of investment for 2000 to 2004. SARIMA is a univariate seasonal ARIMA model. SMARIMA is a regression of GDP and lagged capital on investment, with seasonal ARIMA errors. VAR is a bivariate VAR model with investment and GDP.

Figure 2 presents the test sample forecast results for the SARIMA, SMARIMA, and VAR models.[4] All models tend to match the overall cycle quite well, with the VAR having the lowest mean squared error. The most significant difference between the models and the observed data is their forecasts for quarters three and four. The observed investment value follows a steep cycle, whereas the models have 'hitches' in the third and fourth quarter, most noticeably in the SMARIMA model.

Since the models seem to predict certain quarters accurately, table 5 presents quarterly specific mean squared errors. The SARIMA model outperforms the VAR in quarter two, but the VAR

[4] Islam's MARIMA model is omitted as it performs poorly and makes the figure hard to read.

outperforms the SARIMA model in all other quarters. There is no quarter where the SMARIMA model outperforms the VAR or SARIMA models. Islam's model gives more accurate forecasts in quarter four than any other model besides the VAR but performs significantly worse than all other models in other quarters.

| Forecast Type | Q1 MSE | Q2 MSE | Q3 MSE | Q4 MSE | MSE Total |
|---|---|---|---|---|---|
| MARIMA | 7,048,325.7 | 25,964,924.7 | 10,024,485.0 | 1,221,837.6 | 11,064,893.2 |
| SMARIMA | 2,990,832.0 | 2,133,007.3 | 1,522,193.9 | 2,581,725 | 2,306,939.8 |
| SARIMA | 456,815.1 | 599,838.9 | 927,183.6 | 1,648,173 | 908,002.7 |
| VAR | 401,249.4 | 1,150,672.7 | 730,928.7 | 674,436.2 | 739,321.8 |

**Table 5:** Mean squared errors for several forecast models, for a given quarter of the testing period. Testing period runs from 2000 to 2004. MARIMA is the original authors model, and the other three are alternative methods formulated in this paper.

| | Diebold-Mariano Tests | |
|---|---|---|
| | DM Statistic | p-value |
| $H_0$: VAR is as good as SARIMA | -1.0145 | 0.1616 |
| $H_0$: VAR is as good as SMARIMA | -4.5368 | 0.0001128 |
| $H_0$: VAR is as good as MARIMA | -3.7348 | 0.0007019 |
| $H_0$: SMARIMA is as good as MARIMA | -3.1938 | 0.00239 |
| $H_0$: SARIMA is as good as MARIMA | -3.5501 | 0.001069 |

**Table 6:** Results from the Diebold-Mariano test comparing forecasting accuracy from several models. Squared errors are used in the loss functions. One sided hypothesis are conducted. This tests if method 2 is less accurate than method 1. Method 1 is the first model presented in each null hypothesis.

To test each model's predictive power, I obtain a vector of residuals from each model to perform the Diebold-Mariano (DM) test.[5] The DM test compares forecast accuracy between two models to determine if one model has more predictive power than the other. Table 6 presents the results of several DM tests. From these tests, I conclude that the bivariate VAR model is not better than the SARIMA model. Even though the VAR model has a mean squared error approximately 160,000 lower, it does not significantly outperform the univariate method.

Each alternative method formulated in this paper outperforms Islam's MARIMA model in terms of predictive power. In particular, just adding seasonal components to get the SMARIMA model leads to a significantly better model. However, it is possible that due to data differences, their MARIMA model might have fit their dataset better than it fits mine. Therefore, some of these gains in predictive power might be due to different datasets.

## 6 Conclusion

I was able to replicate Islam's results with moderate success. My data series differed, particularly in earlier periods, which led to differences in my OLS coefficients, but relatively small differences in my MARIMA coefficients. I hypothesize that the differences between my series are primarily due to differences in converting from nominal to real terms.

Both models that utilized the structural relationship of the FAM perform worse than pure time series methods in forecasting ability. This result suggests that the structural relationship is ill-fitting, or that the time series methods are simply more powerful. Alternative specifications, such as the FAM

[5] The Diebold-Mariano test provides a more robust result than just observing differences in the mean squared errors.

model utilized by Kopcke, might be more appropriate since it added several lags of GDP, which might add predictive power. Finally, it may be worth abandoning the FAM entirely and using an alternative model such as the neoclassical, also suggested by Kopcke.

The VAR model could also improve by adding other informative series that explain investment. One could estimate a form of neoclassical investment in VAR form by adding a series on the user cost of capital. This series would likely be a powerful predictor, as GDP can be considered income to spend on investment but the current specification has no information on the costs of investment, which the user cost of capital would provide.

## References


[1] Islam, A (2007). *Explaining and Forecasting Investment Expenditure in Canada: Combined Structural and Time Series Approaches, 1961-2000.* Applied Econometrics and International Development. Vol. 7, No. 1.

[2] Koyck, LM (1954). *Distributed Lags and Investment Analysis.* North Holland, Amsterdam.

[3] Joregensen, DW (1971). *Econometric Studies of Investment Behaviour: A Survey.* Journal of Economic Literature, Vol. 9, No. 4, pp. 1111-1147.

[4] Kopckhe, RW (1993). *Forecasting Investment with Models and Surveys of Capital Spending.* Federal Reserve Bank of Boston.

[5] Statistics Canada (2022). *Gross fixed capital formation, quarterly, Canada (x 1,000,000).* Table 36-10-0108-01.

[6] Statistics Canada (2022). *Gross domestic product, expenditure-based, Canada, quarterly (x 1,000,000)*. Table 36-10-0104-01.

[7] Statistics Canada (2022). *Consumer Price Index, annual average, not seasonally adjusted*. Table 18-10-0005-01.

[8] Statistics Canada (2022). *Gross Domestic Product Price Indexes, quarterly*. Table 36-10-0106-01.