Return Predictability with Macroeconomic Variables

Asad Dossani
Colorado State University
asad.dossani@colostate.edu

American Economic Association 2026 Annual Meeting

Abstract

Consider a predictive regression with a financial asset, such as a stock market return, as the dependent variable, and a macroeconomic variable, such as consumption growth, as the independent variable or predictor. Such regressions in practice often fail to yield evidence of return predictability. In contrast, if the predictor variable is aggregated at a lower frequency, for example annual consumption growth rather than monthly consumption growth, evidence of return predictability is much stronger, both economically and statistically. The first contribution of the paper is a model to explain this result. I assume that the macroeconomic predictor is measured with error and is subject to adjustment dynamics. For example, aggregate consumption might respond slowly to a shock due to adjustment costs. I further assume that the macroeconomic predictor may contain a long memory or slow moving component. The second contribution of the paper is an instrumental variables estimator to recover the true parameter of the original predictability regression. The instrument is constructed using a lagged low frequency value of the predictor variable. I show that the estimator is robust to the presence of measurement error, adjustment dynamics, and long memory. The consistency of the proposed instrumental estimator relies on the long memory property of the predictor variable. Approximate consistency relies on the sum of autocorrelations being large relative to the covariance between the measurement error and shocks to the time series.

Univariate Model - Data Generating Process

The DGP depends on y_t and p_t, with $t = 0, \dots, T$. x_t^* is the first difference of p_t^*, and is the sum of of two components: μ_t and g_t. μ_t is a covariance stationary long memory process, and g_t is a covariance stationary short memory process. ϵ_t is mean zero, uncorrelated over time, and uncorrelated with x_{t-1}^*.

$$\begin{aligned} x_t^* &\equiv p_t^* - p_{t-1}^* \\ &= \mu_t + g_t \\ y_t &= \alpha + \beta x_{t-1}^* + \epsilon_t \\ &= \alpha + \beta(\mu_{t-1} + g_{t-1}) + \epsilon_t \end{aligned}$$

Suppose p_t^* is measured with error v_t, so that we observe p_t and x_t, where x_t is the first difference of p_t. Let u_t denote shocks to g_t. v_t and u_t are white noise processes that are possibly correlated. Let $\sigma_v^2 \equiv \mathrm{Var}(v_t)$ and $\sigma_{vu} \equiv \mathrm{Cov}(v_t, u_t)$.

$$\begin{aligned} p_t &= p_t^* + v_t \\ x_t &\equiv p_t - p_{t-1} \\ &= \mu_t + g_t + v_t - v_{t-1} \end{aligned}$$

OLS Estimator

The estimable model is a regression of y_t on a constant and x_{t-1}. Let $\rho_j \equiv \mathrm{Cov}(\mu_t, \mu_{t-j})$ and $\gamma_j \equiv \mathrm{Cov}(g_t, g_{t-j})$. The OLS estimate and its probability limit are given by:

$$\hat{\beta}_{OLS} = [\mathrm{Var}(x_{t-1})]^{-1}[\mathrm{Cov}(x_{t-1}, y_t)]$$

$$\text{plim } \hat{\beta}_{OLS} = [\rho_0 + \gamma_0 + 2\sigma_{vu} + 2\sigma_v^2]^{-1}[\rho_0 + \gamma_0 + \sigma_{vu}]\beta$$

In general, the OLS slope coefficient is asymptotically biased towards zero.

IV Estimator

Consider the estimable model summed over t and $t+1$, i.e. a regression of $(y_{t+1} + y_t)$ on a constant and $(x_t + x_{t-1})$. Then, use $\sum_{i=1}^{M} x_{t-i}$ as an instrument for $(x_t + x_{t-1})$. The estimate and plim are given by:

$$\hat{\beta}_{IV} = [\mathrm{Cov}(\sum_{i=1}^{M} x_{t-i}, x_t + x_{t-1})]^{-1}[\mathrm{Cov}(\sum_{i=1}^{M} x_{t-i}, y_{t+1} + y_t)]$$

$$\text{plim } \hat{\beta}_{IV} = [\tilde{\rho}_M + \tilde{\gamma}_M]^{-1}[\tilde{\rho}_M + \tilde{\gamma}_M + \sigma_{vu}]\beta$$

$$\tilde{\rho}_M \equiv \rho_0 + 2\sum_{i=1}^{M-1} \rho_i + \rho_M$$

$$\tilde{\gamma}_M \equiv \gamma_0 + 2\sum_{i=1}^{M-1} \gamma_i + \gamma_M$$

We now consider the limit of plim $\hat{\beta}_{IV}$ as M increases. $\tilde{\gamma}_M$ converges to a constant. $\tilde{\rho}_M$ does not converge to a constant due to the long memory property of μ_t. As M increases, $[\tilde{\rho}_M + \tilde{\gamma}_M]^{-1}[\tilde{\rho}_M + \tilde{\gamma}_M + \sigma_{vu}]$ converges to unity. Note we assume that M does not grow as a fast as T, so that standard large sample asymptotics hold as M increases. Thus $\lim_{M\to\infty} \text{plim } \hat{\beta}_{IV} = \beta$

Multivariate Model

We generalize the model to allow for multiple predictors. The probability limit of the OLS and IV estimators are given by:

$$\text{plim } \hat{\boldsymbol{\beta}}_{OLS} = [\boldsymbol{\rho}_0 + \boldsymbol{\gamma}_0 + \boldsymbol{\Sigma}_{\boldsymbol{vu}} + \boldsymbol{\Sigma}_{\boldsymbol{vu}}' + 2\boldsymbol{\Sigma}_{\boldsymbol{v}}]^{-1}[\boldsymbol{\rho}_0 + \boldsymbol{\gamma}_0 + \boldsymbol{\Sigma}_{\boldsymbol{vu}}]\boldsymbol{\beta}$$

$$\text{plim } \hat{\boldsymbol{\beta}}_{IV} = [\tilde{\boldsymbol{\rho}}_M + \tilde{\boldsymbol{\gamma}}_M + \boldsymbol{\Sigma}_{\boldsymbol{vu}} - \boldsymbol{\Sigma}_{\boldsymbol{vu}}']^{-1}[\tilde{\boldsymbol{\rho}}_M + \tilde{\boldsymbol{\gamma}}_M + \boldsymbol{\Sigma}_{\boldsymbol{vu}}]\boldsymbol{\beta}$$

$$\tilde{\boldsymbol{\rho}}_M \equiv \boldsymbol{\rho}_0 + 2\sum_{i=1}^{M-1} \boldsymbol{\rho}_i + \boldsymbol{\rho}_M$$

$$\tilde{\boldsymbol{\gamma}}_M \equiv \boldsymbol{\gamma}_0 + 2\sum_{i=1}^{M-1} \boldsymbol{\gamma}_i + \boldsymbol{\gamma}_M$$

As with the univariate case, plim $\hat{\boldsymbol{\beta}}_{IV}$ converges to $\boldsymbol{\beta}$ as M increases.

Inference

Due to the aggregation of the estimable model over t and $(t+1)$, the residuals are autocorrelated at one lag. Let e_t denote the residual from the IV regression, $\boldsymbol{X}_t$ denote the regressors, $\boldsymbol{Z}_t$ denote the instruments, and $\boldsymbol{B}$ denote the parameters. The distribution of the IV estimator is given by:

$$\sqrt{T - (M+1)}\,(\hat{\boldsymbol{B}} - \boldsymbol{B}) \overset{a}{\sim} \mathrm{N}(\boldsymbol{0}, \boldsymbol{V})$$

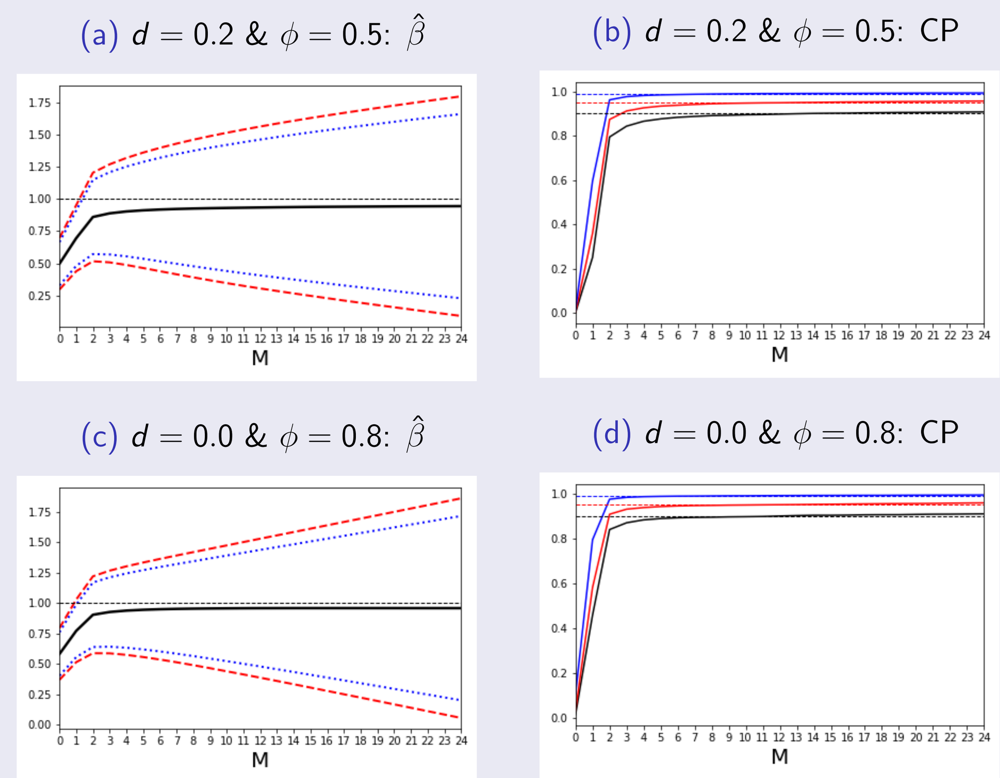
$$\begin{aligned} \boldsymbol{V} &= \left(\mathrm{E}[\boldsymbol{Z}_t\boldsymbol{X}_t']\right)^{-1} \boldsymbol{S} \left(\mathrm{E}[\boldsymbol{X}_t\boldsymbol{Z}_t']\right)^{-1} \\ \boldsymbol{S} &\equiv \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_1' \\ \boldsymbol{\Gamma}_0 &\equiv \mathrm{E}[\boldsymbol{Z}_t\boldsymbol{Z}_t' e_t^2] \\ \boldsymbol{\Gamma}_1 &\equiv \mathrm{E}[\boldsymbol{Z}_t e_t e_{t-1} \boldsymbol{Z}_{t-1}'] \\ &= \frac{1}{2}\mathrm{E}(e_t^2)\mathrm{E}[\boldsymbol{Z}_t\boldsymbol{Z}_{t-1}'] \end{aligned}$$

The covariance matrix can be consistently estimated using the sample analog of the above expressions. $\boldsymbol{\Gamma}_1$ is efficiently estimated as the sample analog of $\frac{1}{2}\mathrm{E}(e_t^2)\mathrm{E}[\boldsymbol{Z}_t\boldsymbol{Z}_{t-1}']$, as this takes into account autocorrelation structure of the residual.

Simulation

Using simulated data, I examine the performance of the proposed estimator. x_t^* follows a mean zero fractionally integrated autoregressive model of order one. I consider two simulations: In the first, $d = 0.2$ and $\phi = 0.5$. This corresponds to a long memory process with a moderately persistent short memory component. In the second, $d = 0.0$ and $\phi = 0.8$. This corresponds to a standard short memory process, but with greater persistence. I set $\sigma_u^2 = \sigma_v^2 = 1$, and $\rho_{vu} \equiv \frac{\sigma_{vu}}{\sigma_v \sigma_u} = -0.5$. A negative correlation in shocks to g_t and p_t is consistent with slow adjustment of macroeconomic variables to their equilibrium values. In the return predictability regression, I set $\alpha = 0$ and $\beta = 1$, $R^2 = 0.05$, the sample size is $T = 1,000$, and the total number of simulations is $100,000$.

Figure: Simulation Results



(a) $d = 0.2$ & $\phi = 0.5$: $\hat{\beta}$ (b) $d = 0.2$ & $\phi = 0.5$: CP (c) $d = 0.0$ & $\phi = 0.8$: $\hat{\beta}$ (d) $d = 0.0$ & $\phi = 0.8$: CP

Left - estimated slope coefficient for each M, with the 90% and 95% confidence intervals. Right - the 90% (black), 95% (red), and 99% (blue) coverage probabilities. $M = 0$ corresponds to the OLS regression, while $M = 1, \dots, 24$ are the IV regressions.

Table: Simulation Results

	$d = 0.2$ & $\phi = 0.5$		$d = 0.0$ & $\phi = 0.8$	
M	$\hat{\beta}$ ($\hat{se}$) ($\hat{mse}$)	CP-90 CP-95 CP-99	$\hat{\beta}$ ($\hat{se}$) ($\hat{mse}$)	CP-90 CP-95 CP-99
0	0.496	0.001	0.581	0.016
	(0.102)	0.002	(0.108)	0.033
	(0.264)	0.013	(0.187)	0.108
3	0.889	0.843	0.926	0.870
	(0.194)	0.912	(0.173)	0.931
	(0.050)	0.977	(0.035)	0.983
6	0.918	0.883	0.949	0.892
	(0.244)	0.938	(0.210)	0.945
	(0.066)	0.986	(0.047)	0.989

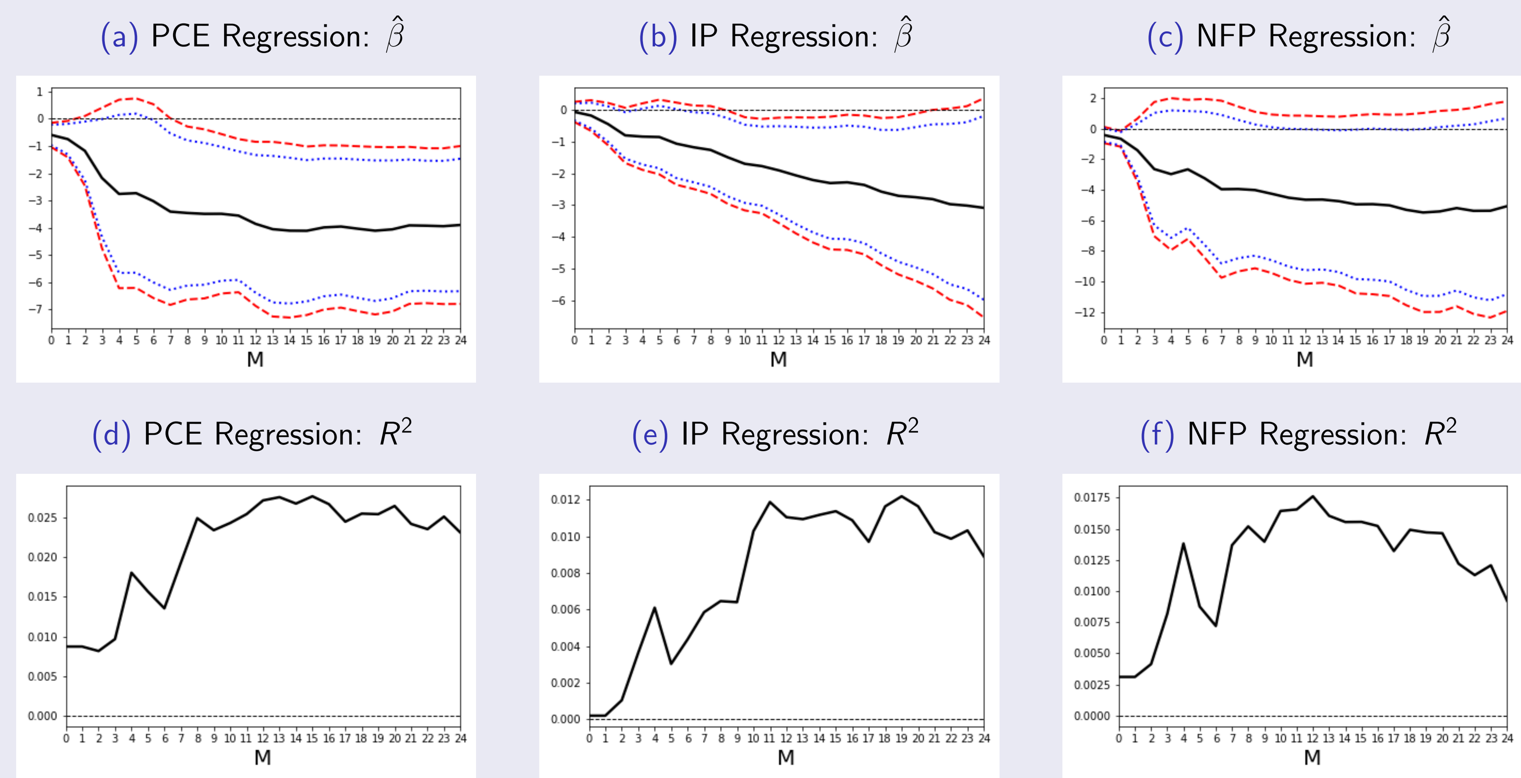
	$d = 0.2$ & $\phi = 0.5$		$d = 0.0$ & $\phi = 0.8$	
M	$\hat{\beta}$ ($\hat{se}$) ($\hat{mse}$)	CP-90 CP-95 CP-99	$\hat{\beta}$ ($\hat{se}$) ($\hat{mse}$)	CP-90 CP-95 CP-99
12	0.934	0.898	0.957	0.900
	(0.321)	0.949	(0.291)	0.950
	(0.107)	0.991	(0.086)	0.991
18	0.941	0.904	0.958	0.906
	(0.381)	0.954	(0.374)	0.955
	(0.149)	0.992	(0.142)	0.992
24	0.945	0.908	0.958	0.910
	(0.435)	0.957	(0.459)	0.959
	(0.193)	0.993	(0.213)	0.994

In each table, the left column reports the estimated slope coefficient $\hat{\beta}$, with the standard error ($\hat{se}$) and mean squared error ($\hat{mse}$) in the rows below. The right column reports the 90%, 95%, and 99% coverage probabilities.

Empirical Example

The dependent variable is the real excess return on the value weighted market index. The first predictor is the log change in consumption growth. The second predictor is the log change in industrial production. The third predictor is the log change in non-farm payrolls. The data start in January 1959 and end in December 2024. The data are sampled monthly, and $T = 792$. I estimate the predictive regression using OLS, and using the proposed IV method, for $M = 1, \dots, 24$. I collect the estimated slope coefficient, associated standard error, and the r-squared. The results show that evidence of return predictability is stronger when using the proposed instrumental variables estimator, relative to OLS. In general, evidence of predictability is greater as M increases, with diminishing effects at larger values of M.

Figure: Return Predictability Regression Results



(a) PCE Regression: $\hat{\beta}$ (b) IP Regression: $\hat{\beta}$ (c) NFP Regression: $\hat{\beta}$ (d) PCE Regression: R^2 (e) IP Regression: R^2 (f) NFP Regression: R^2

I plot the estimated slope coefficients and r-squared for each M. The dotted blue and dashed red lines are the 90% and 95% confidence intervals, respectively. $M = 0$ corresponds to the OLS regression, while $M = 1, \dots, 24$ are the IV regressions.

Table: Return Predictability Regression Results

	PCE Regression		IP Regression		NFP Regression	
M	$\hat{\beta}$ ($\hat{se}$)	R^2 (p)	$\hat{\beta}$ ($\hat{se}$)	R^2 (p)	$\hat{\beta}$ ($\hat{se}$)	R^2 (p)
0	−0.587	0.009	−0.062	0.000	−0.424	0.003
	(0.223)	(0.009)	(0.159)	(0.696)	(0.270)	(0.117)
2	−1.176	0.008	−0.449	0.001	−1.421	0.004
	(0.653)	(0.072)	(0.339)	(0.185)	(1.053)	(0.177)
4	−2.759	0.018	−0.843	0.006	−2.978	0.014
	(1.764)	(0.118)	(0.532)	(0.114)	(2.527)	(0.239)
6	−3.020	0.014	−1.066	0.004	−3.252	0.007
	(1.813)	(0.096)	(0.659)	(0.106)	(2.639)	(0.218)

	PCE Regression		IP Regression		NFP Regression	
M	$\hat{\beta}$ ($\hat{se}$)	R^2 (p)	$\hat{\beta}$ ($\hat{se}$)	R^2 (p)	$\hat{\beta}$ ($\hat{se}$)	R^2 (p)
9	−3.491	0.023	−1.494	0.006	−4.019	0.014
	(1.582)	(0.028)	(0.746)	(0.046)	(2.602)	(0.123)
12	−3.858	0.027	−1.907	0.011	−4.650	0.018
	(1.536)	(0.012)	(0.845)	(0.024)	(2.797)	(0.097)
18	−4.037	0.026	−2.577	0.012	−5.311	0.015
	(1.544)	(0.009)	(1.181)	(0.029)	(3.173)	(0.095)
24	−3.899	0.023	−3.086	0.009	−5.070	0.009
	(1.478)	(0.008)	(1.759)	(0.080)	(3.486)	(0.146)

For each regression, the left column reports the estimated slop coefficient $\hat{\beta}$, with the standard error ($\hat{se}$) in parentheses below. The right column reports the R^2, with the p-value (p) testing the null hypothesis that $\beta = 0$ in parentheses below.