MIDAS Instruments

Eric Ghysels* **Jonathan H. Wright†**

Abstract

Estimation of time series models by instrumental variables has been considered by many authors, including Hayashi and Sims (1983) and Hansen and Singleton (1996). Identification is achieved from the error term being orthogonal to the entire history of one or more time series. Adding more lags as instruments can only improve efficiency, asymptotically. However, in small samples this is clearly not true. If we wish to keep the information in a large number of lags, but avoid overparameterization of the model, we need some device to reduce dimensionality. The idea of using functions of instruments to reduce the dimension of the instrument set is well known; for example Donald, Imbens, and Newey (2003) proposed using splines as instruments. In a time series context, this paper proposes using MIDAS polynomials of candidate basic instrumental variables to reduce dimensionality and get better small-sample performance. The approach may be especially useful where the sampling frequency is mixed.

*Department of Finance, Kenan-Flagler School of Business and Department of Economics University of North Carolina, Gardner Hall CB 3305,Gardner Hall CB 3305, Chapel Hill, NC 27599-3305, phone: (919) 966-5325, e-mail: eghysels@unc.edu.

†Department of Economics, Johns Hopkins University, Baltimore MD 21202, phone: (410) 516-5728, e-mail: wrightj@jhu.edu.

1 Introduction

In time series applications of instrumental variable (IV) or Generalized Method of Moments (GMM) estimation, the selection of the strongest instruments is essentially a forecasting problem. It entails finding the best forecast of the endogenous variable, or more generally of the score of the underlying moment condition. However, it is common practice to simply use a few lags as instruments, even when this is nowhere near best practice in forecasting. For example, in estimating a forward-looking New Keynesian Phillips curve or a forward-looking Taylor rule, it is standard to instrument future inflation by a few lags of inflation and some measure of economic activity. But, this is far from best practice in predicting inflation. In this and many other contexts, insights from the forecasting literature may in turn give us stronger instruments.

Mixed data sampling (MIDAS) regressions have been found to be quite useful in many forecasting problems, as argued in Ghysels, Santa-Clara, and Valkanov (2006), Ghysels and Wright (2009), among others.[1] As in a very old distributed lags literature, they have the benefit of parsimony. They also bring forecasting improvements when the time series are observed at different frequencies. But if MIDAS regressions give good forecasts, then this paper argues that they should also provide useful instruments.

Estimation of time series models by instrumental variables has been considered by many authors, including Hayashi and Sims (1983), Hansen (1985), Hansen, Heaton and Ogaki (1988), Hansen and Singleton (1996), West (2001), Hansen and West (2002) and Anatolyev (2007). Identification is achieved from the error term being orthogonal to the entire history of one or more time series. Increasing the number of lags can only

[1] Recent papers documenting forecasting gains with MIDAS include Forsberg and Ghysels (2006), Galvão (2006), Clements, Galvão, and Kim (2007), Ghysels, Plazzi, and Valkanov (2007), Armesto, Hernandez-Murillo, Owyang, and Piger (2008), Chen and Ghysels (2008), Clements and Galvão (2008, 2009), Schumacher and Breitung (2008), Andreou, Ghysels, and Kourtellos (2009) and Ghysels and Valkanov (2009).

improve efficiency asymptotically, but in small samples this is clearly not necessarily true. If we wish to keep the information from using a large number of lags, but avoid overparameterization of the model, we may need some device to reduce dimensionality. The idea of using functions of instruments to reduce the dimension of the instrument set is well known; for example Donald, Imbens, and Newey (2003) proposed using splines as instruments. This paper proposes using MIDAS polynomials of candidate instrumental variables to reduce dimensionality and get better small-sample performance. The approach is especially useful in contexts where the sampling frequency is mixed.

The plan for the remainder of this paper is as follows. The proposed methodology for using MIDAS polynomials as instruments is laid out in section 2. Section 3 contains Monte-Carlo simulations. Some applications are presented in section 4. Section 5 concludes.

2 Methodology

Consider the moment condition $E_{t-1}(h(Y_t,\theta_0))=0$ where θ is a scalar parameter, θ_0 is its true value, $\{Y_t\}_{t=1}^T$ is a time series and $h(.,.)$ is a scalar function. Let $f(Y_t,\theta,\alpha) = h(Y_t,\theta)b(\alpha,L)Z_t$ where $\{Z_{t-1},Z_{t-2},...\}$ is a set of possible instruments, $b(\alpha,L)=\Sigma_{j=1}^{\bar{k}}b_jL^j$ and $b_j=(\exp(\alpha_1 j+\alpha_2 j^2))/(\Sigma_{i=1}^{\bar{k}}\exp(\alpha_1 i+\alpha_2 i^2))$. Each of the basic instruments (the Z_ts) is assumed to be a qx1 vector, $q\geq 1$.

Consider the moment condition $E(f(Y_t,\theta,\alpha))=0$, which holds for any value of α. The GMM estimator corresponding to this condition is

$$\hat{\theta}(\alpha)=\arg\min_{\theta}\Sigma_{t=1}^T f(Y_t,\theta,\alpha)'W\Sigma_{t=1}^T f(Y_t,\theta,\alpha) \quad (2.1)$$

where W is a weight matrix, which is the identity matrix for the one-step estimator. For the continuous-updating estimator, it is $\hat{\Omega}(\theta,\alpha)^{-1}$, where $\hat{\Omega}(\theta,\alpha)$ is a consistent estimate of the zero-frequency spectral density of $f(Y_t,\theta,\alpha)$. And for the two-step estimator, it is $\hat{\Omega}(\hat{\theta}_{OS}(\alpha),\alpha)^{-1}$where $\hat{\theta}_{OS}(\alpha)$ denotes the one-step estimator. We assume that either the continuous-updating or two-step estimator is being employed. Under the usual conditions, for any fixed q, the asymptotic distribution of the GMM estimator is then

$$\sqrt{T}(\hat{\theta}(\alpha)-\theta_0)\longrightarrow_d N(0,(D(\alpha,\theta_0)'\Omega(\alpha,\theta_0)^{-1}D(\alpha,\theta_0))^{-1}) \tag{2.2}$$

where $\Omega(\theta,\alpha)$is the zero-frequency spectral density of $f(Y_t,\theta,\alpha)$and $D(\alpha,\theta)=E((\partial f(Y_t,\theta,\alpha))/(\partial\theta))$. The optimal choice of α is

$$\alpha_0(\theta)=\arg\max_{\alpha} D(\alpha,\theta)'\Omega(\alpha,\theta)^{-1}D(\alpha,\theta) \tag{2.3}$$

This maximizes the concentration parameter, that measures the strength of identification. So the strategy can be to iterate between (2.1) and

$$\hat{\alpha}(\theta)=\arg\max_{\alpha}\hat{D}(\alpha,\theta)'\hat{\Omega}(\alpha,\theta)^{-1}\hat{D}(\alpha,\theta) \tag{2.4}$$

where $\hat{D}(\alpha,\theta)=T^{-1}\Sigma_{t=1}^{T}(\partial f(Y_t,\theta,\alpha))/(\partial\theta)$ until convergence of both optimization problems. This allows us to compute $\hat{\theta}^*$ an optimal estimator of θ within the class of moment conditions considered here. Call this the MIDAS-IV estimator. It's asymptotic distribution is

$$\sqrt{T}(\hat{\theta}^*-\theta_0)\longrightarrow_d N(0,(D(\alpha_0,\theta_0)'\Omega(\alpha_0,\theta_0)^{-1}D(\alpha_0,\theta_0))^{-1}) \tag{2.5}$$

Asymptotically, using all the instruments instead of a particular function of the instruments has to be optimal, but we suspect that in small samples, using this lower-dimensional representation is likely to work better. Still it is important to note that the MIDAS-IV estimator will not generally be asymptotically efficient.

The parameter θ is a scalar here. If it were instead a vector, then the criterion function for the optimal choice of α would require some weighting of the different parameters..

2.1 The Linear IV model

Consider the linear IV model, in which $y_t = \beta x_t + u_t$ is a single equation where both variables are endogenous. The structural parameter is $\theta = \beta$ and we have $h(Y_t \ , \theta)$ $= y_t - \beta x_t$ and $\hat{D}(\alpha, \theta) = T^{-1} \Sigma_{t=1}^{T} x_t \tilde{Z}_t$ where $\tilde{Z}_t = b(\alpha, L) Z_t$. Suppose moreover that the instruments are strictly exogenous and $(y_t, x_t)'$ is i.i.d. so that $\Omega(\alpha, \theta)$ is the variance of $h(Y_t, \theta) = y_t - \beta x_t$ times the variance of $\tilde{Z}_t$. In this special case, $\Omega(\alpha, \theta)$ is a product of a function of θ alone and a function of α alone and so $\hat{\alpha}$ is simply the estimate of α in a nonlinear least squares regression of x_t onto $\tilde{Z}_t = b(\alpha, L) Z_t$. In other words, in this special case, the optimal moment condition uses an instrument that is simply the fitted value in a MIDAS regression of x_t on $Z_{t-1}, Z_{t-2}, \ldots, Z_{t-\bar{k}}$. But in time series contexts, we can seldom, if ever, think of the endogenous variables as being i.i.d. or of the instruments as being strictly exogenous.

The models considered in the Monte-Carlo simulations and in the empirical work in this paper are all linear models. However, in all cases we estimate the zero-frequency spectral density of $f(Y_t, \theta, \alpha)$ by the Newey-West estimator, and do not make any assumption of strict exogeneity of the instruments.

2.2 Weak Identification

All this reasoning works as long as the identification is strong, i.e. $D(\alpha, \theta)$ is of full column rank. Otherwise, however, the parameter θ is not identified. Moreover, if $D(\alpha, \theta)$ is nearly rank deficient, then selecting the moment conditions to maximize the correlation between endogenous variables and instruments could well make matters worse in small samples (Hall, Rudebusch, and Wilcox (1996)). If $D(\alpha, \theta)$ is rank deficient, or nearly so, then a better approach is to give up on point estimation and treat θ as instead being set-identified. Inference then entails constructing a confidence set for this parameter vector (see, for example, Stock, Wright and Yogo (2002)). But we would like this confidence set to be as tight as possible. In this subsection, we offer some thoughts on how this might be implemented.

Suppose first that $q = 1$. A confidence set for θ is formed by inverting the acceptance region of a test of the null hypothesis $H_0 : \theta = \theta_0$. Let J be the zero-frequency spectral density matrix of $(f(Y_t, \theta, \alpha)\ \partial f(Y_t, \theta, \alpha)/\partial\theta)'$. Let

$$\hat{\alpha} = \arg\max_{\alpha} \frac{D^*(\alpha, \theta_0)^2}{\hat{\Omega}(\alpha, \theta_0)} \tag{2.6}$$

where $D^*(\alpha, \theta) = \hat{D}(\alpha, \theta) - J_{21}(\alpha, \theta) J_{11}^{-1}(\alpha, \theta) T^{-1} \Sigma_{t=1}^{T} f(Y_t, \theta, \alpha)$. We assume that

$$\left(\ \sqrt{T}(T^{-1}\Sigma_{t=1}^{T} f(Y_t, \theta, \alpha) - E(f(Y_t, \theta, \alpha)\ \ \hat{D}(\alpha, \theta) - D(\alpha, \theta))'\ \right) \longrightarrow_d N(0, J) \tag{2.7}$$

which implies that $D^*(\alpha, \theta)$ and hence $\hat{\alpha}$ are asymptotically independent of $\Sigma_{t=1}^{T} f(Y_t, \theta_0, \alpha)$. No assumption about identification is being made: $E(f(Y_t, \theta_0, \alpha))$ could be uniformly equal to zero, in which case the model is completely unidentified. This device of orthogonalizing the score with respect to the moment function was proposed by Kleibergen (2002) and Kleibergen (2005).

Under the null, $H_0 : \theta = \theta_0$, conditional on $\hat{\alpha}$.

$$K = \frac{\{T^{-1/2}\Sigma_{t=1}^{T} f(Y_t, \theta_0, \hat{\alpha})\}^2}{\hat{\Omega}(\alpha, \theta_0)} \longrightarrow_d \chi^2(1) \tag{2.8}$$

But $\hat{\alpha}$ is asymptotically independent of $\Sigma_{t=1}^{T} f(Y_t, \theta_0, \alpha)$ and so this limiting result holds unconditionally as well. The test statistic can then be compared to the critical values of a $\chi^2(1)$ at each of a grid of values of θ, and the confidence set for θ is simply the inverse of the acceptance region of this test.

If instead there are $q > 1$ basic instruments, then J will be a $2q$x$2q$ matrix. We can define:

$$\hat{\alpha} = \arg\max_{\alpha} D^*(\alpha, \theta_0)' \hat{\Omega}(\alpha, \theta_0)^{-1} D^*(\alpha, \theta_0) \tag{2.9}$$

Under the null, we can test the hypothesis $\theta = \theta_0$ using the statistic

$$K = T^{-1}\Sigma_{t=1}^{T} f(Y_t, \theta_0, \hat{\alpha})' \hat{\Omega}(\alpha, \theta_0)^{-1/2} P_{\hat{\Omega}(\alpha, \theta_0)^{-1/2} D^*(\alpha, \theta_0)} \hat{\Omega}(\alpha, \theta_0)^{-1/2} \Sigma_{t=1}^{T} f(Y_t, \theta_0, \hat{\alpha}) \tag{2.10}$$

where $P_A = A(A'A)^{-1}A'$ which is again asymptotically $\chi^2(1)$ distributed under the null.

Nevertheless, in the remainder of this paper, we do not consider the weak-identification version of MIDAS instruments, and instead assume that $D(\alpha, \theta)$ is of full rank, as in conventional asymptotic theory. The focus is therefore on reduction in the dimensionality of the instruments, rather than in dealing with a failure or near-failure of identification.

3 Monte-Carlo Design

We investigate the performance of the proposed MIDAS-IV estimator in three simple time series models.

3.1 Instrumental Variables Estimation of an ARMA process

Consider the ARMA(1,1) time series model

$$y_t = ay_{t-1} + u_t + bu_{t-1} \tag{3.11}$$

where u_t is i.i.d. $N(0, \sigma^2)$ and the sample size is T. It is well known that this model may be estimated by instrumental variables, using any subset of $\{y_{t-2}, y_{t-3}, ..\}$ as instruments. In this procedure, the moving average component of the model is not estimated. Although increasing the number of lags used as instruments can only help asymptotic efficiency, using a lower-dimensional set of instruments is likely to work better in small samples.

Hansen, Heaton and Ogaki (1988) provide asymptotic efficiency bounds for IV estimators in this model. They show that for any choice of instruments, $T^{1/2}(\hat{a} - a) \rightarrow_d N(0, V_a)$ where $V_a \geq ((1 + ab)^2(1 - a^2))/(a + b)^2$. Figures 1 and 2 show the simulated mean square error for the IV estimator using as instruments, $\{y_{t-2}, y_{t-3}, ..y_{t-k+1}\}$, with sample sizes of $T = 500$ and $T = 1000$, respectively. Results are shown for different values of a and b, and the simulated mean-square error is plotted against the number of instruments, k. The figures also shows the mean-square error of the proposed MIDAS-IV estimator, with the lag truncation parameter $\bar{k}$ set at 15, and the efficiency bound. In all these simulations, $\sigma^2 = 1$.

In both sample sizes, adding instruments reduces mean-square error up to a point, but beyond that further increases in the number of lags pushes mean-square error up. Using the MIDAS-IV estimator usually gives a lower mean square error than using any number of lags as instruments. It actually gets quite close to the efficiency bound. This indicates that in this application the MIDAS polynomial is doing an effective job of combining the information in the lags, without introducing an excessive number of instruments.

3.2 Instrumental Variables Estimation of a Forecasting Equation

Next we consider a stylized forecasting model. The model specifies that

$$y_t = \beta x_{t-1} + \varepsilon_t \quad (3.12)$$

$$\varepsilon_t = v_t + b v_{t-1} \quad (3.13)$$

$$x_t = a x_{t-1} + u_t \quad (3.14)$$

where u_t is i.i.d $N(0, \sigma_u^2)$, v_t is i.i.d. $N(0, \sigma_v^2)$ and these errors are mutually uncorrelated. The objective is to estimate β and hence to find the forecast of y_t given lags of x_t. Clearly, β can be estimated by an OLS regression of y_t on x_t which amounts to the GMM estimator using the moment condition $E(x_{t-1}\varepsilon_t) = 0$. However, the inclusion of other moment conditions of the form $E(x_{t-s}\varepsilon_t)$, $s = 1, 2, ..$ can also improve efficiency, at least asymptotically. Indeed, Hansen and West (2002) show that the estimator using the single moment condition $E(\tilde{x}_{t-1}\varepsilon_t) = 0$ where $\tilde{x}_{t-1} = \Sigma_{j=0}^{\infty}(-b)^j x_{t-1-j}$ is optimal in the sense of achieving the efficiency bound $V_{\hat{\beta}} = \frac{\sigma_v^2}{\sigma_u^2}(1 - a^2)(1 + ab)^2$, though it is not feasible since b is unknown.[2] Meanwhile, the standard OLS estimator has asymptotic variance $\frac{\sigma_v^2}{\sigma_u^2}(1-a^2)(1+a^2+2ab)$.Many forecasting models with multi-period horizons have forecast errors with a moving average structure along with a highly persistent regressor, as in this model. Nonetheless, it is standard practice to estimate these forecasting equations by OLS, apparently because of the concern that including more moment conditions will just induce bias or impair efficiency in small samples. As an alternative, we might use $b(\alpha, L)x_{t-1}$ as an instrument where $b(\alpha, L) = \Sigma_{j=1}^{\bar{k}} b_j L^j$ and $b_j = (\alpha_1 j + \alpha_2 j^2)/(\Sigma_{i=1}^{\bar{k}} |\alpha_1 i + \alpha_2 i^2|)$. This is a MIDAS-IV estimator, though using a slightly different lag polynomial than before, to capture the fact that the weights on an optimal instrument oscillate in sign in this kind of application.

Figures 3 and 4 show the simulated mean square error for the IV estimator using k lags as instruments, $\{x_{t-1}, x_{t-2}, ..x_{t-k}\}$ with sample sizes of $T = 500$ and $T = 1000$, respectively. Of course, the case $k = 1$ corresponds to OLS. The figures also shows the mean-square error of the proposed MIDAS-IV estimator, with the lag truncation parameter $\bar{k}$ set at 15, and the efficiency bound. In these simulations, $\sigma_u^2 = \sigma_v^2 = 1$ and the true value of $\beta = 0$.

In both sample sizes, the choice of k that gives the smallest mean-square error is either 1 (OLS) or a small number. But using the MIDAS-IV estimator usually gives a lower mean square error than using any number of lags as instruments. Indeed, it gets very close to the efficiency bound, especially if a is not too large.

3.3 Instrumental Variables Estimation of a Rational Expectations Model

Finally we consider a stylized forward-looking rational expectations model that can be estimated by instrumental variables. The model specifies that

[2] The general principle is that the optimal instrument is the expectation of the forward-filtered regressor conditional on the set of instruments (Hayashi and Sims (1983)).

$$
\begin{aligned}
y_t &= \beta E_{t-1} x_t + u_t & (3.15)\\
x_t &= a x_{t-1} + v_t + b v_{t-1} & (3.16)
\end{aligned}
$$

where u_t is i.i.d $N(0, \sigma_u^2)$, v_t is i.i.d. $N(0, \sigma_v^2)$ and these errors have contemporaneous correlation δ. The econometrician does not know the data generating process for x_t. Nonetheless, the agents in this model use it to form the conditional expectation of x_t, and so $E_{t-1}(x_t) = (a+b)\Sigma_{j=0}^{\infty}(-b)^j x_{t-j-1}$. The objective is to estimate β by a regression of y_t onto realized values of x_t

$$
y_t = \beta x_t + \xi_t \tag{3.17}
$$

where $\xi_t = u_t - \beta(x_t - E_{t-1}(x_t))$ by instrumental variables, using some subset of $\{x_{t-1}, x_{t-2}, ...\}$ as instruments. The idea of replacing conditional expectations by their realized values and then estimating the equation by instrumental variables is widely used in the estimation of forward-looking macroeconomic models. Instrumental variables have to be used because the composite error, ξ_t, is correlated with x_t.

Figure 5 plots the mean-square error of the IV estimator where the basic instruments are $\{x_{t-1}, x_{t-2}, ...x_{t-k}\}$ against the number of instruments k. Results are shown for the sample size of $T = 500$, with parameters $\beta = 0$, $\delta = 0.9$ and different values of a and b. As in our earlier simulations, increasing the number of instruments initially reduces the mean-square error, but then increasing it further causes it to rise. Figure 5 also shows the mean-square error of the MIDAS-IV estimator with the lag truncation parameter $\bar{k}$ set at 15. This estimator has a low mean-square error–in most cases lower than can be obtained with any set of the basic instruments.

Precision of point estimates is not the only criterion that one might wish to consider. Figures 6 and 7 show the simulated effective coverage and average width of 95 percent confidence intervals for β using k basic instruments (for different values of k) and also using the MIDAS-IV estimator. In all cases, conventional asymptotic Newey-West standard errors are used. Using a small number of basic instruments gives an effective coverage rate that is close to the nominal level, but the confidence intervals are quite wide. Increasing the number of instruments, k, reduces the width of the confidence intervals, but the coverage quickly falls very far below the nominal level. The MIDAS-IV estimator has an effective coverage rate that is close to the nominal level, while giving an average width of the confidence intervals that is smaller than can be obtained with any other IV estimator that has reasonable coverage.

4 Empirical Applications

We next demonstrate the proposed approach to IV estimation in three important applications.

4.1 Euler equation for output

The first is a standard forward-looking macroeconomic model estimated by instrumental variables methods; the forward-looking Euler equation for output (Fuhrer and Rudebusch (2004)). This model specifies that:

$$y_t = \rho_0 + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \beta E_t(y_{t+1}) + \sigma \frac{1}{k} \Sigma_{i=0}^{k-1} E_t(r_{t+i} - \pi_{t+1+i}) + \varepsilon_t \tag{4.18}$$

where y_t is the output gap, r_t denotes the short-term interest rate and π_t is the inflation rate. The model relates the output gap to lagged output, expectations of future output, and the real interest rate. The simplest form of the model would set ρ_1 and ρ_2 to zero and $k = 1$, but models of the form of equation 4.18 appear to fir the data better. If we let $u_{1,t} = y_{t+1} - E_t(y_{t+1})$ and $u_{2,t} = \frac{1}{k} \Sigma_{i=0}^{k-1} (r_{t+i} - \pi_{t+1+i} - E_t(r_{t+i} - \pi_{t+1+i})$, we can substitute out the unobserved expectations, writing the model as

$$y_t = \rho_0 + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \beta y_{t+1} + \sigma \frac{1}{k} \Sigma_{i=0}^{k-1} (r_{t+i} - \pi_{t+1+i}) + w_t \tag{4.19}$$

where $w_t = \varepsilon_t - \beta u_{1t} - \sigma u_{2,t}$. The equation cannot be estimated by OLS, because future inflation and the error term will be correlated, but it can be estimated by instrumental variables, as w_t is orthogonal to anything in the information set at time t.

We estimate this equation using monthly U.S. data from January 1960-February 2009. The short-term interest rate is measured by the federal funds rate, the output gap is proxied by quadratically-detrended log industrial production, and inflation is measured using the month-over-month change in the total CPI (seasonally adjusted). The parameter k is set to 6, meaning that spending decision are affected by expectations of real-short term interest rates over the next 6 months.

We use four different sets of instruments:

I1: A constant, two lags of the output gap, one lag of inflation and one lag of the short-term interest rate,

I2: A constant and four lags of the output gap, inflation and the short-term interest rate, and

I3: A constant and twelve lags of the output gap, inflation and the short-term interest rate, and

I4: A constant, four lags of the output gap and the same MIDAS function of inflation and the short-term interest rate over the previous 48 months.

The results are shown in Table 1. The first instrument set produces the largest standard errors, indicating that adding more lags may sharpen estimation precision. The conventional standard errors are smallest using instruments I3, but this instrument set has 37 instruments, and so these standard errors surely overstate the precision of the estimates. Instrument set I2 has 13 elements and the MIDAS-IV estimator (with instrument set I4) has just 7 instruments. Neverthless, the MIDAS-IV estimator gives standard errors for the parameters β and σ that are smaller than can be obtained with instrument set I2.

The upper panel of Figure 8 plots the estimated MIDAS function of lagged inflation and short-term interest rates that are used by the MIDAS-IV estimator. As can be seen, the weights die off fairly slowly, with data from about two years earlier still contributing to the identification.

4.2 Estimating the C-CAPM Euler equation

The second application is to the problem of estimating a log-linearized Euler equation in the consumption CAPM. With constant relative risk aversion, this log-linearized Euler equation is:

$$E_t(\log(\delta) + r_{t+1} - \gamma c_{t+1}) = 0 \tag{4.20}$$

where r_{t+1} is the log stock return, δ is the discount factor, γ is the coefficient of risk aversion and c_{t+1} is the growth rate of real per-capita consumption. Following Campbell and Mankiw (1989), the equation can be rearranged as follows:

$$c_{t+1} = \alpha + \sigma r_{t+1} + u_{t+1} \tag{4.21}$$

where $\sigma = 1/\gamma$ is the coefficient of intertemporal substitution and the error term is orthogonal to anything in the information set at time t. The equation can then be estimated by instrumental variables–the motivation for rearranging the equation so that returns are on the right-hand side is that returns are more predictable than consumption growth.

We estimate this equation using annual data from Campbell and Shiller (1987), update to cover the years 1889-2004, with three different sets of instruments:
I1: An intercept and two lags of annual consumption growth and returns,
I2: As in I1 plus the realized volatility constructed from daily stock returns over the previous 220 days (i.e. the sum of the daily squared returns over these days), and
I3: As in I1 plus the MIDAS function of the squared daily stock returns over the previous 220 days.

The instruments in I2 and I3 require daily stock returns, which were obtained from Schwert (1990) up to 1962 spliced onto CRSP value-weighted data since then. Note that this is an application in which the basic instruments are observed at a higher frequency than the endogenous variables. Instruments I1 are the familiar instruments for estimating Euler equations. The motivation for including the extra instruments in I2 and I3 is that one might expect a positive correlation between stock returns and lagged volatility. As further motivation for the use of instruments I3, note that Ghysels, Santa-Clara and Valkanov (2005) find a positive relationship between returns and volatility forecasts from MIDAS models.

The results are shown in Table 2. The point estimates for σ are all around 0.15. However, the standard errors using instruments I3 (the MIDAS instruments) are smaller than using I2, which are in turn smaller than using I1. This indicates that lagged volatility is helping to sharpen identification, and that the MIDAS function incorporates lagged volatility in an efficient way. The bottom panel of Figure 8 plots the estimated MIDAS function of squared daily stock returns used with instruments I3. The estimated weights decay fairly rapidly–most of the weight is on squared returns within the previous two months.

5 Conclusion

This paper has proposed a particular device for reducing the dimensionality of instruments in a time series context. It is most likely to be helpful if a researcher believes that many lags of one or time series may be useful as instruments, but that the asymptotic efficiency gains from including all of them as instruments are likely to be offset by the finite-sample problems associated with increasing the number of instruments. The idea is to use a tightly parameterized lag polynomial function of the basic instruments. The approach may be especially useful where the sampling frequency is mixed.

We have showed evidence in both Monte-Carlo simulations and empirical applications that this approach allows the information in a large number of instruments to be incorporated efficiently. In simulations, we find that the MIDAS-IV estimator proposed in this paper is often more efficient in finite samples than using any set (small or large) of the basic instruments directly. Finally, we find that it tightens inference in two important time-series applications of instrumental variables.

References

Anatolyev, S. (2007), "Optimal instruments in time series: a survey," *Journal of Economic Surveys*, 21, 143–173.

Andreou, E., Ghysels, E., and Kourtellos, A. (2009), "Why should macroeconomic forecasters look at daily financial data," Discussion Paper UNC and University of Cyprus.

Armesto, M., Hernandez-Murillo, R., Owyang, M., and Piger, J. (2009), "Measuring the Information Content of the Beige Book: A Mixed Data Sampling Approach," *Journal of Money, Credit and Banking*, 41, 35-55.

Campbell, J.Y. and R.J. Shiller (1987): "Cointegration and Tests of Present Value Models", *Journal of Political Economy*, 95, 1062-1088.

Campbell, J.Y. and N.G. Mankiw (1989), "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence", *NBER Macroeconomics Annual*, 185-216.

Chen, X. and Ghysels, E. (2008), News-Good or Bad-and its Impact Over Multiple Horizons," Discussion Paper UNC.

Clements, M.P. and Galvão, A.B. (2009), "Forecasting US output growth using Leading Indicators: An appraisal using MIDAS models," *Journal of Applied Econometrics,* (forthcoming).

——— (2008), "Macroeconomic Forecasting with Mixed Frequency Data: Forecasting US output growth," *Journal of Business and Economic Statistics*, 26, 546-554.

Clements, M.P., Galvão, A.B., and Kim, J. (2007), "Quantile forecasts of daily exchange rate returns from forecasts of realized volatility," *Journal of Empirical Finance*, 15, 729-750.

Donald, S., Imbens, G., and Newey, W. (2003), "Empirical likelihood estimation and consistent tests with conditional moment restrictions," *Journal of Econometrics*, 117, 55-93.

Forsberg, L. and Ghysels, E. (2006), "Why do absolute returns predict volatility so well?"*Journal of Financial Econometrics*, 6, 31-67.

Fuhrer, J.C. and G.D. Rudebusch (2004), "Estimating the Euler Equation for Output,"*Journal of Monetary Economics*, 51, 1133-1153.

Galvão, A.B. (2006), "Changes in Predictive Ability with Mixed Frequency Data," Discussion Paper Queen Mary.

Ghysels, E., Plazzi, A., and Valkanov, R. (2007), "Valuation in US Commercial Real Estate," *European Financial Management*, 13, 472-497.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2006), "Predicting volatility: getting the most out of return data sampled at different frequencies," *Journal of Econometrics*, 131, 59-95.

Ghysels, E. and Valkanov, R. (2009), "Granger Causality Tests with Mixed Data Frequencies," Discussion Paper UNC and University of California, San Diego.

Ghysels, E. and Wright, J.H. (2009), "Forecasting Professional Forecasters," *Journal of Business and Economic Statistics* (forthcoming).

Hall, A., Rudebusch, G., and Wilcox, D. (1996), "Judging Instrument Relevance in Instrumental Variables Estimation," *International Economic Review*, 37, 283-298.

Hansen, B.E. and K.D. West (2002), "Generalized Method of Moments and Macroeconomics," *Journal of Business and Economic Statistics*, 20, 460-469.

Hansen, L.P. (1985), "A Method for Calculating Bounds on the Asymptotic Covariance Matrices of Generalized Method of Moments Estimators," *Journal of Econometrics*, 30, 203-238.

Hansen, L.P., Heaton, J., and Ogaki, M. (1988), "Efficiency Bounds Implied by Multiperiod Conditional Moment Restrictions," *Journal of the American Statistical Association*, 83, 863-871.

Hansen, L. P. and Singleton, K. J. (1996), "Efficient Estimation of Linear Asset-Pricing Models with Moving Average Errors," *Journal of Business and Economic Statistics*, 14, 53-68.

Hayashi, F. and Sims, C. (1983), "Nearly Efficient Estimation of Time Series Models with Predetermined, but not Exogenous, Instruments," *Econometrica*, 51, 783-798.

Kleibergen, F. (2002), "Pivotal Statistics for Testing Structural Parameters in Instrumental

Variables Regression," *Econometrica*, 70, 1781-1803.

——— (2005), "Testing Parameters in GMM Without Assuming that They Are Identified,"*Econometrica*, 73, 1103-1123.

Kleibergen, F. and Mavroeidis, S. (2008), "Weak instrument robust tests in GMM and the new Keynesian Phillips curve," Working paper, Brown University.

Schumacher, C. and Breitung, J. (2008), "Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data," *International Journal of Forecasting*, 24, 386-398.

Schwert, G.W. (1989), "Why Does Stock Volatility Change Over Time?" *Journal of Finance*, 44, 1115-1153.

Stock, J., Wright, J.H., and Yogo, M. (2002), "A Survey of Weak Instrument and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 20, 518-529.

West, K.D. (2001), "On Optimal Instrumental Variables Estimation of Stationary Time Series Models," *International Economic Review*, 42, 1043-1050.

Table 1: Alternative estimates of the forward-looking Taylor rule

	Instrument Set			
	I1	I2	I3	I4
ρ_0	0.0629	0.0175	0.0021	0.0146
	(0.0388)	(0.0196)	(0.0096)	(0.0163)
ρ_1	0.8075	0.58	0.5569	0.5644
	(0.1249)	(0.0579)	(0.0399)	(0.0626)
ρ_2	-0.1316	-0.0383	-0.0383	-0.0388
	(0.0598)	(0.0339)	(0.0246)	(0.0395)
β	0.3179	0.4581	0.4828	0.4744
	(0.0832)	(0.0313)	(0.0191)	(0.0299)
σ	-0.027	-0.0079	0.0004	-0.0054
	(0.0151)	(0.0076)	(0.0036)	(0.0059)

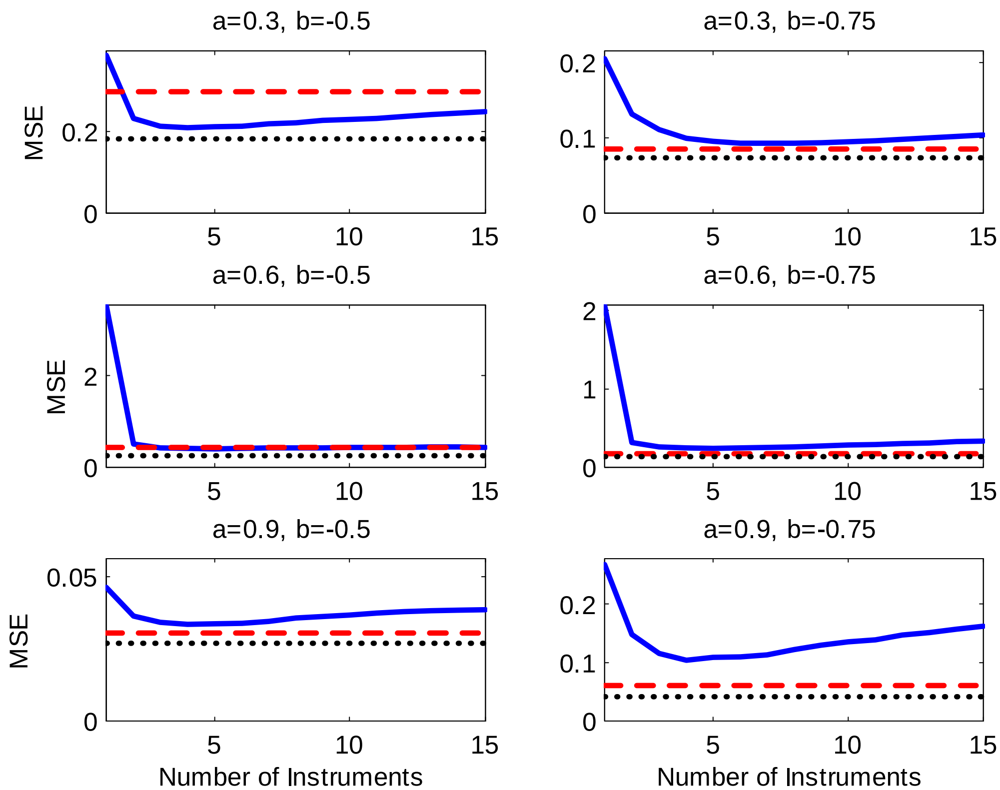
Notes: This table shows the estimates of the forward-looking Taylor rule (equation 4.19 in the text), along with Newey-West standard errors, using the instrument sets I1 (a constant, two lags of the output gap and one lags of inflation and the short-term interest rate), I2 (a constant and four lags of the output gap, inflation and the short-term interest rate), I3 (a constant and twelve lags of the output gap inflation and the short-term interest rate), and I4 (a constant, four lags of the output gap and a MIDAS function of 36 lags of inflation and the short-term interest rate). The equation is estimated at the monthly frequency, using data from January 1960 to February 2009. One, two and three asterisks denote significance at the 10, 5 and 1 percent significance levels, respectively. k denotes the number of moment conditions.

Table 2: Alternative estimates of the log-linearized Euler equation

	Instrument Set		
	I1	I2	I3
α	0.012	0.013	0.012
	(0.005)	(0.004)	(0.004)
σ	0.153^{**}	0.121^{**}	0.151^{***}
	(0.074)	(0.057)	(0.052)

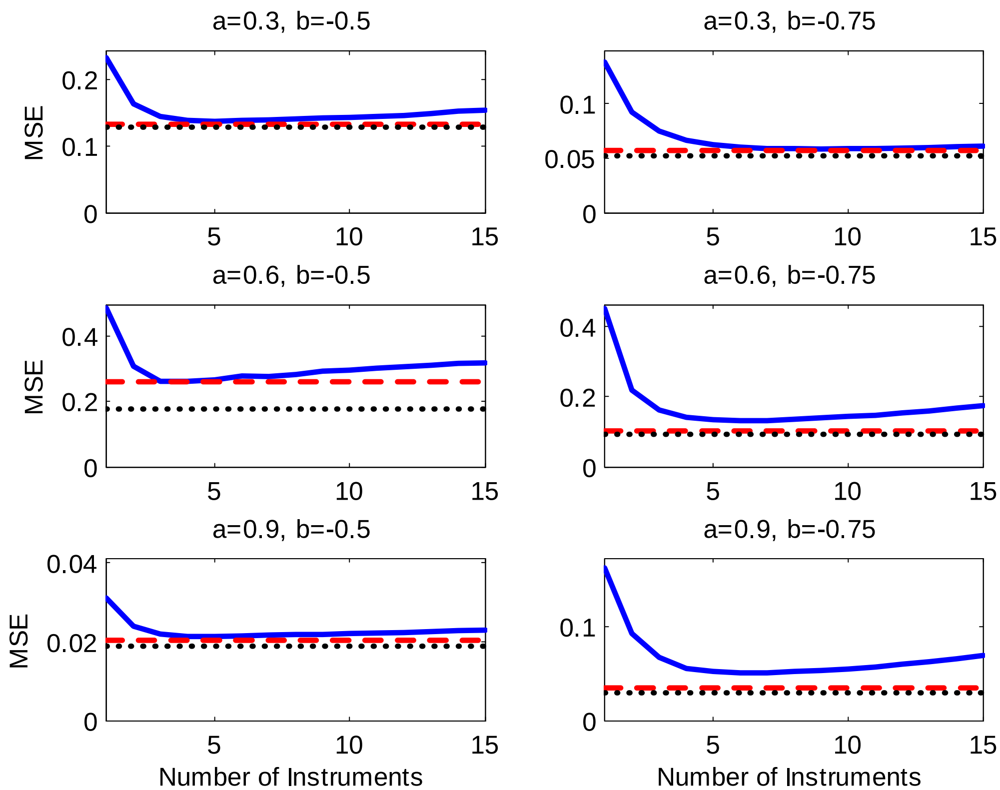
Notes: This table shows the estimates of the log-linearized CAPM Euler equation (equation 4.21 in the text), along with Newey-West standard errors, using the instrument sets I1 (two lags of returns and consumption growth), I2 (I1 plus lagged realized volatility) and I3 (I1 plus a MIDAS function of daily squared returns). The equation is estimated at the annual frequency, using data from 1889 to 2004. One, two and three asterisks denote significance at the 10, 5 and 1 percent significance levels, respectively.

Figure 1: MSE of IV Estimators of ARMA models using k lags as instruments: T=500.



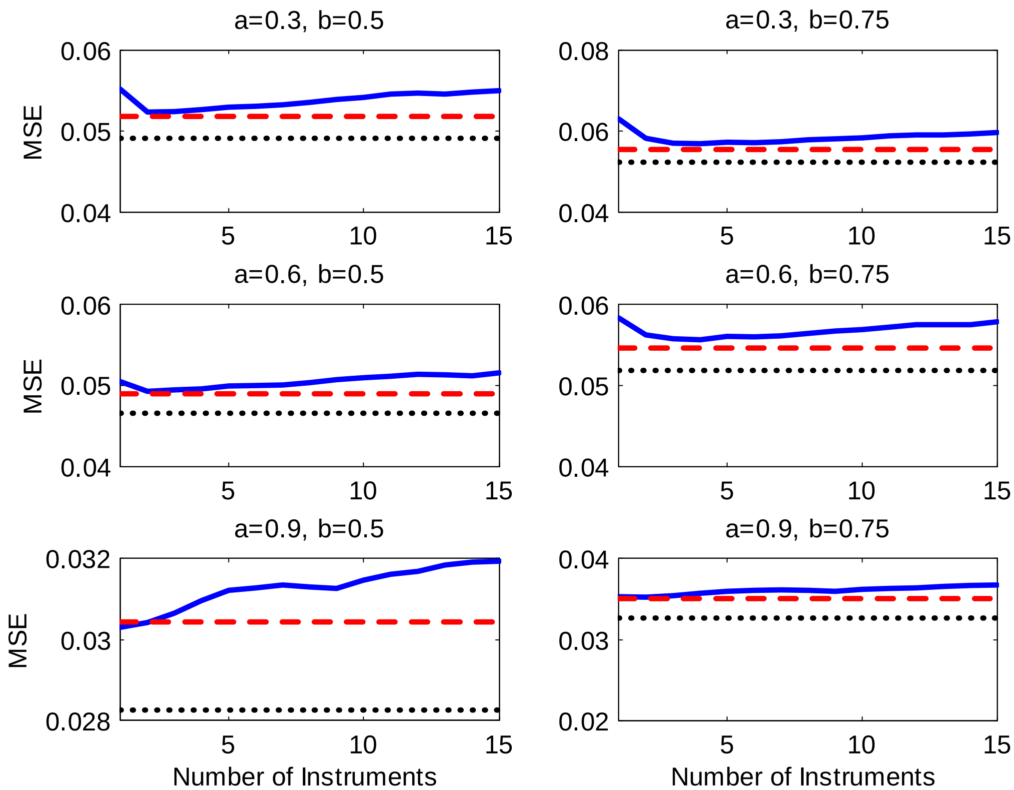
The solid blue line shows the MSE using k lags as instruments. The red dashed line shows the MSE of the MIDAS-IV estimator. The black dotted line is the efficiency bound.

Figure 2: MSE of IV Estimators of ARMA models using k lags as instruments: T=1000.



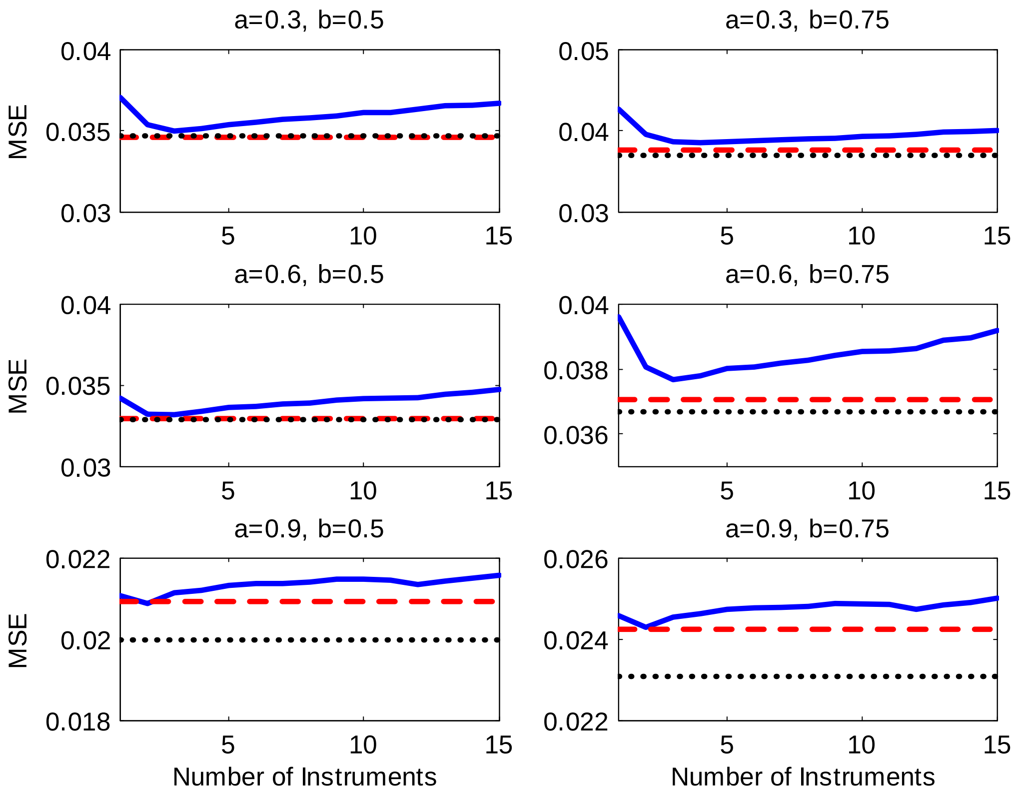
The solid blue line shows the MSE using k lags as instruments. The red dashed line shows the MSE of the MIDAS-IV estimator. The black dotted line is the efficiency bound.

Figure 3: MSE of IV Estimators of Forecasting models using k lags as instruments: T=500.



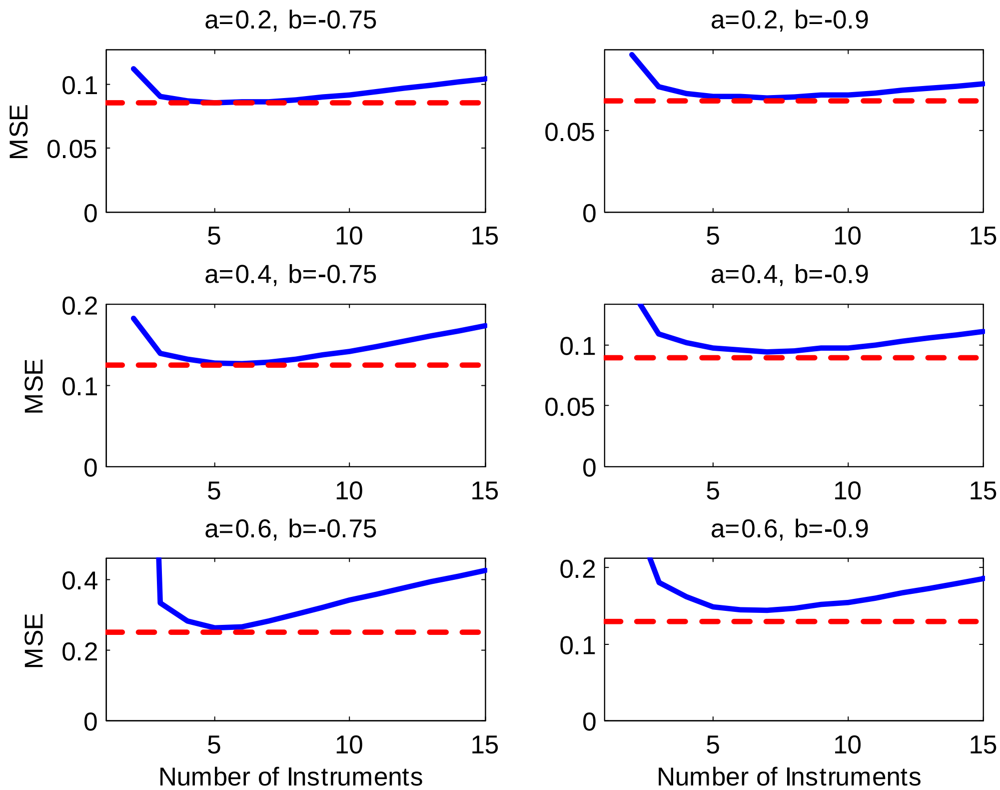
The solid blue line shows the MSE using k lags as instruments. The red dashed line shows the MSE of the MIDAS-IV estimator. The black dotted line is the efficiency bound.

Figure 4: MSE of IV Estimators of Forecasting models using k lags as instruments: T=1000.



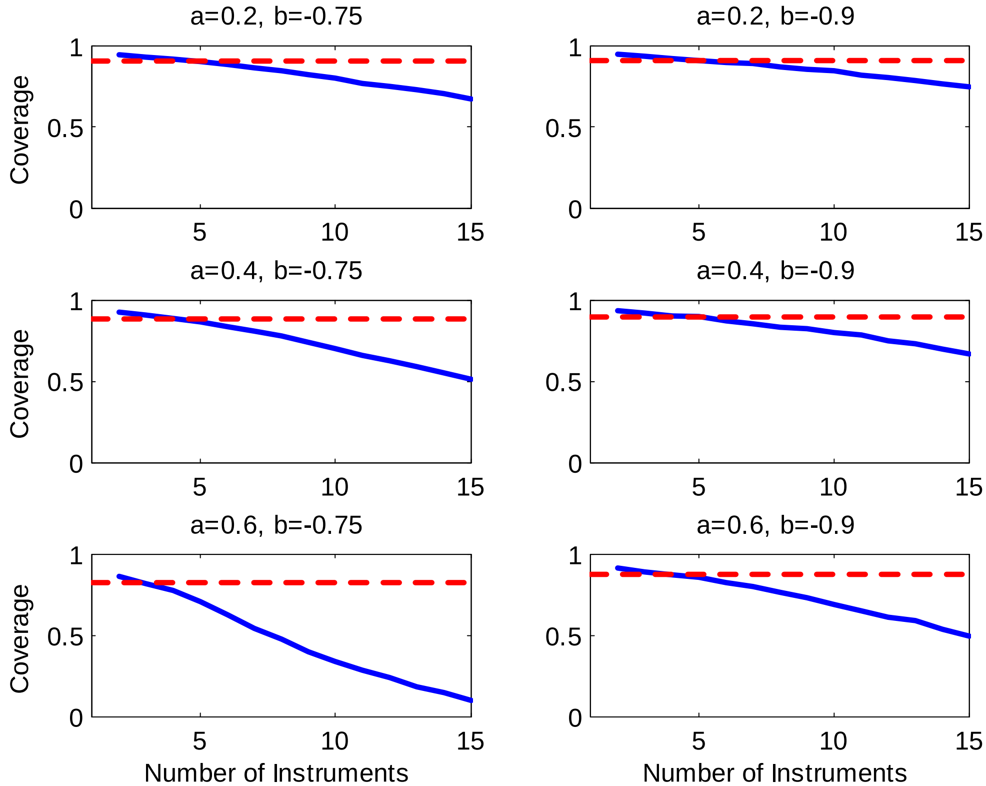
The solid blue line shows the MSE using k lags as instruments. The red dashed line shows the MSE of the MIDAS-IV estimator. The black dotted line is the efficiency bound.

Figure 5: MSE of IV Estimators of the Rational Expectations Model using k lags as instruments: T=500.



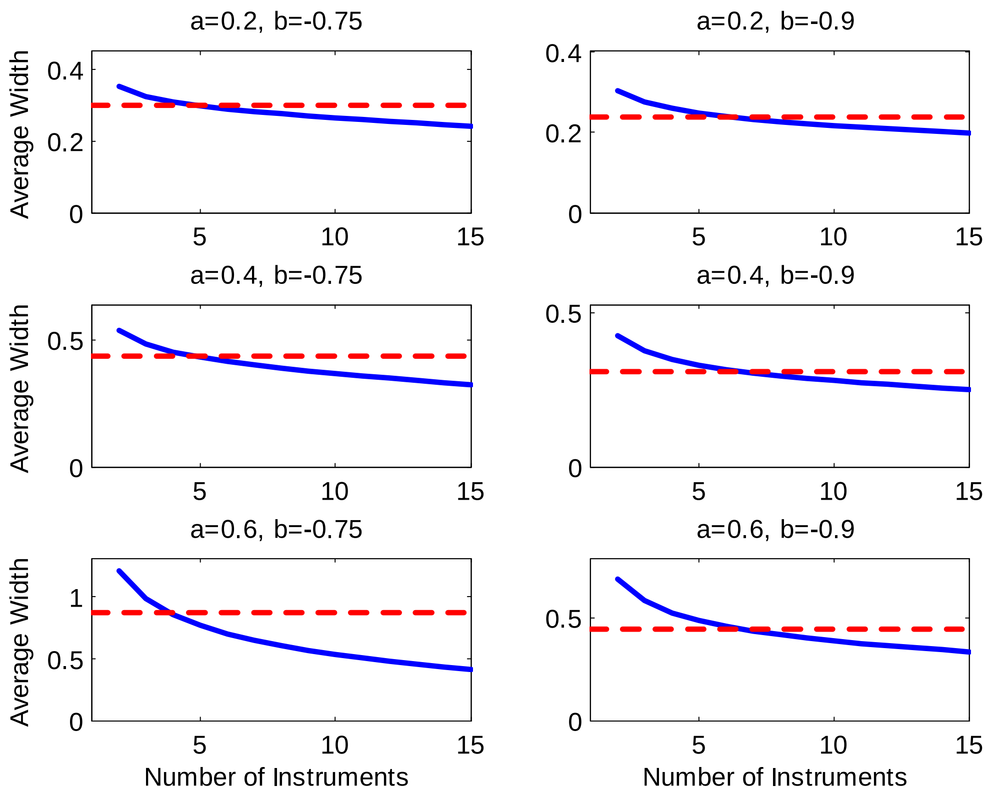
The solid blue line shows the MSE using k lags as instruments. The red dashed line shows the MSE of the MIDAS-IV estimator.

Figure 6: Coverage of IV-Based Confidence Intervals in the Rational Expectations Model using k lags as instruments: T=500.



The solid blue line shows the MSE using k lags as instruments. The red dashed line shows the MSE of the MIDAS-IV estimator.

Figure 7: Average Width of IV-Based Confidence Intervals in the Rational Expectations Model using k lags as instruments: T=500.



The upper panel shows the weights on lagged monthly inflation in forming the MIDAS instrument for the forward-looking Taylor rule. The lower panel shows the weights on lagged daily squared returns in forming the MIDAS instrument for the Euler equation.

Figure 8: Estimated MIDAS Polynomials in the Empirical Applications