Abstract—In the instrumental variables (IV) regression model, weak instruments can lead to bias in estimators and size distortion in hypothesis tests. This paper examines how weak instruments affect the identification of the elasticity of intertemporal substitution (EIS) through the linearized Euler equation. Conventional IV methods result in an empirical puzzle that the EIS is significantly less than 1 but its reciprocal is not different from 1. This paper shows that weak instruments can explain the puzzle and reports valid confidence intervals for the EIS using pivotal statistics. The EIS is less than 1 and not significantly different from 0 for eleven developed countries.

I. Introduction

The elasticity of intertemporal substitution (EIS) in consumption is a parameter of central importance in macroeconomics and finance. In a basic model of the effects of monetary policy, the EIS is the parameter that relates current and expected future real interest rates to the current level of aggregate demand in the intertemporal IS relation (Woodford, 2003, chapter 4). In the consumption and portfolio choice problem of an infinite-lived investor with Epstein-Zin (1989) preferences, the EIS is the key parameter in the optimal consumption rule (Campbell and Viceira, 1999).

To estimate the EIS, denoted by \( \psi \), one typically uses the regression equation

\[
\Delta c_{t+1} = \tau_i + \psi r_{i,t+1} + \xi_{i,t+1},
\]

where \( \Delta c_{t+1} \) is the consumption growth at time \( t + 1 \), \( r_{i,t+1} \) is the real return on asset \( i \) at \( t + 1 \), and \( \tau_i \) is a constant. The error \( \xi_{i,t+1} \), which is linear in the innovation to consumption growth and asset return, is correlated with the regressor \( r_{i,t+1} \). However, given a vector of instruments \( Z_t \), uncorrelated with the error, \( \psi \) can be identified by the moment restriction

\[
E[Z_t \xi_{i,t+1}] = 0.
\]

Here \( Z_t \) typically consists of economic variables known at time \( t \), such as lagged consumption growth and asset return. Equation (1) can be estimated by two-stage least squares (TSLS) if the error is homoskedastic, or by linear generalized method of moments (GMM) if the error is heteroskedastic.

The regression equation (1) can be written in reversed form as

\[
r_{i,t+1} = \mu_i + \frac{1}{\psi} \Delta c_{t+1} + \eta_{i,t+1},
\]

where \( \mu_i \) is a constant and \( \eta_{i,t+1} \) is the error. The reciprocal of the EIS, which is also the coefficient of relative risk aversion under power utility, is then identified by the moment restriction

\[
E[Z_t \eta_{i,t+1}] = 0.
\]

The moment restrictions (2) and (4) are equivalent up to a linear transformation.

Using equation (1) or (3), numerous papers have estimated the EIS with U.S. data (for example, Hansen and Singleton, 1983; Hall, 1988; and Campbell and Mankiw, 1989) and international data (for example, Campbell, 2003). The general empirical finding is that the EIS estimated by equation (1) is small (Hall, 1988), but its reciprocal estimated by equation (3) is also small (Hansen and Singleton, 1983). For instance, Campbell (2003, table 9) reports a 95% confidence interval of [−0.14, 0.28] for \( \psi \), using quarterly U.S. data (1947–1998) on nondurable consumption and T-bill returns. On the other hand, he reports a 95% confidence interval of [−0.73, 2.14] for \( 1/\psi \).

Therefore, one rejects the null hypothesis \( \psi = 1 \) using equation (1), which instruments for T-bill return, but fails to reject \( \psi = 1 \) using equation (3), which instruments for consumption growth. Whether \( \psi < 1 \) is of economic interest because it has important implications for the relative magnitudes of income and substitution effects in the intertemporal consumption decision of an investor facing time-varying expected returns. Campbell and Viceira (1999) show that when the EIS is less (greater) than 1, the investor’s optimal consumption-wealth ratio is increasing (decreasing) in expected returns.

Although equations (1) and (3) correspond to the same moment restriction up to a linear transformation, GMM is not invariant to such transformations. Therefore, the choice of normalization for the moment restriction can affect point estimates and confidence intervals. According to conventional first-order asymptotic theory, the choice of normalization should be negligible in large samples, leading to (at least approximately) the same inference of the EIS. In practice, however, equations (1) and (3) give very different (even contradictory) confidence intervals for the EIS, as discussed above.

© 2004 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology
The leading explanation for this apparent failure of first-order asymptotics is weak instruments. In order for a vector of instruments $Z_i$ to be valid, it must be not only exogenous but relevant, that is, correlated with the endogenous variable $r_{i,t+1}$ in equation (1) or $\Delta c_{i,t+1}$ in equation (3). As Neely, Roy, and Whiteman (2001) and Campbell (2003) note, weak instruments are a problem in estimating the EIS because both consumption growth and asset returns are notoriously difficult to predict. Weak instruments can cause estimators to be severely biased and the finite-sample distribution of test statistics to depart sharply from the limiting distribution, leading to large size distortions in hypothesis tests (see Nelson and Startz (1990), Staiger and Stock (1997), or Stock, Wright, and Yogo (2002) for a recent survey).

The purpose of this paper is to estimate and make valid inference of the EIS for the eleven developed countries in Campbell’s (2003) data set, taking careful account of problems caused by weak instruments. The idea that weak instruments are a problem in estimating the EIS is not new; the paper that is closest to this one is Neely et al. (2001). Showing that weak instruments may account for the discrepancy between small values of $\frac{1}{\theta}$ estimated by equation (1) and small values of $1/\psi$ estimated by equation (3), Neely et al. (2001, p. 403) conclude that “prior beliefs grounded in economic theory seem to be necessary to settle the consumption CAPM debate over small versus large risk aversion” because of identification failure.\(^1\)

Compared to Neely et al. (2001), this paper goes a step further, to estimate the EIS despite near identification failure. I am able to make progress on the EIS debate due to recent methods that have been developed to handle weak instruments. Stock and Yogo (2003) have developed a pretest, based on the first-stage F-statistic, to test formally whether given instruments are weak. Some instrumental variables (IV) estimators, such as the limited information maximum likelihood (LIML) estimator, provide more reliable point estimates and inferences with weak instruments than does TSLS (Hausman, Hahn, & Kuersteiner, 2001; Stock & Yogo, 2003). Kleibergen (2002) and Moreira (2001, 2003) have developed pivotal statistics to test coefficients in the structural equation, which result in tests with correct size regardless of the strength of identification. Using these methods, I conclude that the EIS is small across the eleven developed countries, which agrees with Hall’s (1988) finding for the United States.

The rest of the paper is organized as follows. Section II reviews the assumptions necessary to derive the regression equations (1) and (3) from the Euler equation for Epstein-Zin (1989, 1991) preferences. Section III outlines the relevant econometric methods when instruments are weak. Section IV applies these econometric methods to data from eleven developed countries and discusses the empirical findings. Section V concludes.

II. Linearized Euler Equation

Let $\delta$ be the subjective discount factor and $\gamma$ be the coefficient of relative risk aversion, and define $\theta = (1 - \gamma)/(1 - 1/\psi)$. The Epstein-Zin (1989, 1991) objective function is defined recursively by

$$U_t = [(1 - \delta)C_t^{1-\gamma/\theta} + \delta(\mathbb{E}U_{t+1}^{1-\gamma})^{\theta/(1-\gamma)}],$$

(5)

where $C_t$ is consumption at time $t$. In the special case $\gamma = 1/\psi$, (5) reduces to the familiar time-separable power utility model with period utility $U(C_t) = C_t^{1-\gamma}/(1 - \gamma)$. The representative household maximizes the objective function (5) subject to the intertemporal budget constraint

$$W_{t+1} = (1 + R_{w,t+1})(W_t - C_t),$$

(6)

where $W_{t+1}$ is the household’s wealth and $1 + R_{w,t+1}$ is the gross real return on the portfolio of all invested wealth at $t + 1$. Epstein and Zin (1991) show that equations (5) and (6) together imply an Euler equation of the form

$$\mathbb{E}_t\left[\frac{(C_{t+1}/C_t)^{1-\psi}}{1 + R_{w,t+1}} \left(1 + R_{t+1}ight)\right] = 1,$$

(7)

where $1 + R_{t+1}$ is the gross real return on asset $i$.

A. Conditional Homoskedasticity

Let lowercase letters denote the logarithms of the corresponding uppercase variables [for example, $r_{i,t+1} = \log (1 + R_{i,t+1})$]. Assuming that asset returns and consumption are homoskedastic and jointly log normal conditional on information at time $t$, the Euler equation (7) can be linearized as

$$\mathbb{E}_t r_{i,t+1} = \mu_i + \frac{1}{\psi} \mathbb{E}_t \Delta c_{i,t+1},$$

(8)

$$\mu_i = -\log \delta + \frac{\theta - 1}{2} \text{Var} (r_{w,t+1} - \mathbb{E}_t r_{w,t+1})$$

$$- \frac{\theta}{2\psi^2} \text{Var} (\Delta c_{i,t+1} - \mathbb{E}_t \Delta c_{i,t+1}),$$

$$\mu_i = \mu_i - \frac{1}{2} \text{Var} (r_{i,t+1} - \mathbb{E}_t r_{i,t+1})$$

$$+ \frac{\theta}{\psi} \text{Cov} (r_{i,t+1} - \mathbb{E}_t r_{i,t+1}, \Delta c_{i+1} - \mathbb{E}_t \Delta c_{i+1})$$

$$+ (1 - \theta) \text{Cov} (r_{i,t+1} - \mathbb{E}_t r_{i,t+1}, r_{w,t+1} - \mathbb{E}_t r_{w,t+1}).$$

[See Campbell (2003) or Campbell and Viceira (2002, chapter 2) for a textbook treatment.] Without the assumption

\(^1\) Neely et al. assume power utility, so the risk aversion is the reciprocal of the EIS.
of log-normality, equation (8) holds as a second-order log linear approximation of equation (7). For a conditionally risk-free asset, equation (8) reduces to

$$r_{t+1} = \mu_f + \frac{1}{\psi} E_r \Delta c_{t+1}. \quad (11)$$

The regression equation (3) is obtained from equation (8) by setting

$$\eta_{t+1} = r_{t+1} - \frac{1}{\psi} (\Delta c_{t+1} - E_r \Delta c_{t+1}).$$

The error $\eta_{t+1}$ is conditionally homoskedastic by the same assumption used to linearize the Euler equation. It is straightforward to show that $\eta_{t+1}$ is serially uncorrelated and satisfies the moment restriction (4). The efficient two-step GMM estimator is TSLS in this case. In order for the instruments to be relevant (that is, not weak), they must be correlated with consumption growth $\Delta c_{t+1}$.

The regression equation (1) is obtained by rearranging equation (3), which implies that

$$\xi_{t+1} = \Delta c_{t+1} - E_r \Delta c_{t+1} - \psi (r_{t+1} - E_r r_{t+1}).$$

Because the moment restriction (2) is satisfied, $\psi$ can be estimated by TSLS. In this normalization, the instruments are weak if they are weakly correlated with asset return $r_{t+1}$.

B. Conditional Heteroskedasticity

If asset returns and consumption are conditionally heteroskedastic, the Euler equation (7) can still be linearized as equation (8). The only difference is that the variance and covariance terms that appear in the intercept $\mu_c$ must be replaced by conditional variances and covariances. In this section, I show that the EIS can still be identified by the same moment restrictions.

To simplify the notation, consider the linearized Euler equation for the risk-free asset (11) in the special case of power utility (that is, $\gamma = 1/\psi$ and $\theta = 1$),

$$r_{t+1} = \mu_f + \gamma E_r \Delta c_{t+1}, \quad (12)$$

$$\mu_f = -\log \delta - \frac{\gamma^2}{2} \operatorname{Var}_r (\Delta c_{t+1} - E_r \Delta c_{t+1}). \quad (13)$$

The intercept $\mu_f$ is now subscripted by $t$ to allow for conditional heteroskedasticity in consumption, which represents precautionary savings. As long as the vector of instruments $Z_t$ is uncorrelated with the innovation to the conditional variance of consumption, that is, $E[Z_t (\mu_f - \mu)]=0$, the reciprocal of the EIS (the coefficient of relative risk aversion in this case) is identified by the moment restriction (4). In this case,

TSLS is consistent but is no longer the efficient two-step GMM estimator.

This suggests that even if instruments $Z_t$ are correlated with the conditional variance and covariance terms that appear in $\mu_{t+1}$, a vector of twice lagged instruments $Z_{t-1}$ satisfies the moment restriction $E[Z_{t-1} \eta_{t+1}] = 0$, where

$$\eta_{t+1} = \mu_c - \mu_f + \gamma r_{t+1} - \frac{1}{\psi} (\Delta c_{t+1} - E_r \Delta c_{t+1}).$$

In other words, the reciprocal of the EIS can still be identified by the regression equation (3) although inference must now account for conditional heteroskedasticity in the error $\eta_{t+1}$. A similar point has been made by Attanasio and Low (2000) in the context of estimating the linearized Euler equation on household data. They argue that the coefficient of relative risk aversion can be identified with sufficiently long time series data in response to Carroll’s (2001) criticism that it cannot be estimated consistently on a cross section of households.

C. Estimation of the Nonlinear Euler Equation

This paper focuses on estimation of the EIS based on the linearized Euler equation. In this section, I briefly compare this approach with estimation based on the nonlinear Euler equation.

Given a vector of instruments, the preference parameters $\delta$, $\gamma$, and $\psi$ can be estimated by GMM through the nonlinear Euler equation (7). This is the approach taken by Hansen and Singleton (1982) for the power utility case, and by Epstein and Zin (1991) for Epstein-Zin preferences. As noted by Epstein and Zin (1991), the difficulty with this approach is that it requires knowledge of returns on the wealth portfolio, which includes returns on human capital. Hence, Roll’s (1977) critique of the testability of CAPM applies. In contrast, the EIS can be estimated from the linearized Euler equation (8) without knowledge of returns on the wealth portfolio.

Aside from this practical advantage, the reason for focusing on the linearized Euler equation is that much more is known about weak instruments in the linear IV regression model. Many of the recent econometric methods that handle weak instruments (for example, Stock & Yogo, 2003; Kleibergen, 2002; and Moreira, 2003) apply to the linear IV model with conditional homoskedasticity. I therefore impose the assumption of conditional homoskedasticity for most of the empirical work in section IV, although I also check that the results are robust to heteroskedasticity.

The main disadvantage of the linearized Euler equation is that the discount factor $\delta$ cannot be identified, because it enters additively in the intercept (10), along with unknown second moments of innovations to consumption and asset returns (Attanasio & Low, 2000). Nevertheless, the study of weak instruments in the linear model is interesting because
of the large existing literature that uses this methodology, starting with Hansen and Singleton (1983) and Hall (1988). For those interested in the nonlinear model, I refer to a related study by Stock and Wright (2000). They develop GMM asymptotic theory under weak identification and apply it to estimation of preference parameters through the nonlinear Euler equation.

III. Econometric Methods for Weak Instruments

Following the notation in Staiger and Stock (1997), the linear IV regression model is

\[ y = Y\beta + X\gamma + u, \]

\[ Y = Z\Pi + X\Phi + V, \]  

(14)

(15)

where equation (14) is the structural equation of interest and equation (15) is the reduced form for the \( PX \) regressors, equation (15) is the structural equation of interest and is used for projection onto matrices other than \( \Pi \). The three special cases of interest in this paper are

1. TSLS with \( k = 1 \);
2. LIML with \( k = LIML \), where \( LIML \) is the smallest root of the determinantal equation \( [Y'V_1Y - kY_2V_2Y] = 0 \);
3. Fuller-\( k \) (Fuller, 1977) with \( k = LIML - 1/(T - K_1 - K_2) \).

The Wald statistic for testing the null hypothesis \( \beta = \beta_0 \) is

\[ W(k) = \frac{\hat{\beta}(k) - \beta_0}{\hat{\sigma}_{\beta_0}} \left[ \frac{1}{T} \sum (Y - \hat{Y})' \left( I - kM_k \right) (Y - \hat{Y}) \right] \frac{1}{n\hat{\sigma}_{\beta_0}}, \]

(20)

where \( \hat{\sigma}_{\beta_0} = \hat{\sigma}_{\beta_0} = \hat{\sigma}_{\beta_0} (k) \) and \( \hat{\sigma}_{\beta_0} = \hat{\sigma}_{\beta_0} (k) (T - K_1 - n) \) and \( \hat{\sigma}_{\beta_0} = y - \hat{Y} \hat{\beta} \).

Under conventional first-order asymptotics, the three \( k \)-class estimators and the corresponding Wald statistics have the same asymptotic distribution (see Amemiya, 1985, pp. 236–238). However, first-order asymptotics is a poor approximation in finite samples when instruments are weak (Nelson & Startz, 1990). Staiger and Stock (1997) develop an alternative asymptotic framework, weak-instrument asymptotics, which accurately approximates the sampling distribution of estimators and test statistics even when instruments are weak.

Under weak-instrument asymptotics, the three estimators and the corresponding Wald statistics have nonstandard limiting distributions that differ from one another. Both TSLS and Fuller-\( k \) are biased, but the bias of Fuller-\( k \) is less severe for given population parameters. Similarly, the size distortion of the LIML Wald test is less severe than that of the TSLS Wald test (Stock and Yogo, 2003). Hence, Fuller-\( k \) and LIML can be thought of as estimators that are more robust to weak instruments than TSLS (see Stock et al., 2002, section 6).

B. Test for Weak Instruments

Suppose there is only one endogenous regressor in the structural equation (that is, \( n = 1 \)). Then the key population parameter that measures the relevance of the instruments is the concentration parameter,

\[ \mu^2 = \frac{\Pi'Z'Z\Pi}{\Sigma_{VV}}. \]

(21)

Following the discussion in Rothenberg (1984, section 6), \( \mu^2 \) can be thought of as the sample size in simultaneous-equation models. When \( \mu^2 \) is large, the TSLS estimator is approximately unbiased, and the distribution of the its \( t \)-statistic is approximately standard normal. When \( \mu^2 \) is small, the TSLS estimator can be badly biased, and the distribution of the its \( t \)-statistic can be highly skewed (see Stock et al., 2002, figure 1).

This suggests that one can test whether instruments are weak by testing whether \( \mu^2 \) is sufficiently small to cause bias or size distortion. To test the null hypothesis that
instruments are weak, Stock and Yogo (2003) propose using the first-stage $F$-statistic,

$$F = \frac{\hat{\Omega} Z' Z \hat{\Omega}}{K_2 \hat{\Sigma}_{VV}},$$  \hspace{1cm} (22)

where $\hat{\Omega} = (Z' Z)^{-\frac{1}{2}} Z' Y$ and $\hat{\Sigma}_{VV} = Y' M_Z Y/(T - K_1 - K_2)$. Note that the $F$-statistic is the sample analog of the concentration parameter (21), scaled by $K_2$. The null hypotheses that I consider in this paper are:

1. The bias of TSLS as a fraction of OLS bias is greater than 10\% (10.27).
2. The actual size of the TSLS $t$-test at 5\% significance can be greater than 10\% (24.58).
3. The bias of Fuller-$k$ as a fraction of OLS bias is greater than 10\% (6.37).
4. The actual size of the LIML $t$-test at 5\% significance can be greater than 10\% (5.44).

The numbers in parentheses are the critical values of the test at 5\% significance when $K_2 = 4$, taken from Stock and Yogo (2003, tables 1–4). For instance, to assure that TSLS relative bias is no greater than 10\%, the $F$-statistic must be greater than 10.27. That the critical value for TSLS is greater than the critical value for Fuller-$k$ is a reflection of the fact that the latter is more robust to weak instruments. Likewise, LIML is less prone to size distortion than TSLS for the same level of instrument relevance, which results in a lower critical value.

C. Similar Tests

The pretest described in the last section can detect weak instruments, protecting the researcher from biased estimates and misleadings inferences. However, a researcher may be interested in making valid inference of the structural parameter $\beta$ despite having weak instruments. In this section, I outline methods fully robust to weak instruments that accomplish this task.

Moreira (2001, 2003) has characterized the family of similar tests in the IV regression model when instruments are fixed (that is, $Z$ is nonrandom), the reduced-form errors $\tilde{V}_i$ are independently and identically distributed normal, and the reduced-form covariance matrix $\Omega$ is known. Under these assumptions, he showed that there is a pair of independent sufficient statistics, $\mathcal{F}$ and $\mathcal{T}$, for the unknown parameters $\beta$ and $\Pi$. Under the null hypothesis $\beta = \beta_0$, $\mathcal{F}$ is pivotal (that is, its distribution does not depend on $\Pi$), and $\mathcal{T}$ is sufficient for nuisance parameter $\Pi$. Hence, any non-pivotal statistic $\phi(\mathcal{F}, \mathcal{T}, \beta_0)$ becomes a pivotal statistic conditional on $\mathcal{T} = \tau$. Let $c(\tau, \beta_0, \alpha)$ be the upper $\alpha$-quantile of the null distribution of $\phi(\mathcal{F}, \tau, \beta_0)$. The test that rejects the null if $\phi(\mathcal{F}, \tau, \beta_0) > c(\tau, \beta_0, \alpha)$ is similar at the level $\alpha$ (see Moreira, 2003, theorem 1).

In the general IV regression model (stochastic regressors, non-Gaussian errors, and unknown $\Omega$), Moreira’s exact finite-sample results hold asymptotically under weak-instrument asymptotics (Moreira, 2003, theorem 2). This result is not surprising, for weak-instrument asymptotics corresponds to the finite-sample distribution theory for the simultaneous-equation model with fixed regressors, Gaussian errors, and known reduced-form covariance matrix. Consequently, the family of similar tests forms a basis for fully robust inference in the presence of weak instruments.2

To characterize these tests, define the vectors $a_0 = (\beta_0, 1)'$ and $b_0 = (1, -\beta_0)'$ and the statistics

$$\mathcal{F} = \frac{(Z' Z)^{-\frac{1}{2}} Z' \tilde{Y} b_0}{(b_0 \hat{\Omega} b_0)^{1/2}},$$  \hspace{1cm} (23)

$$\mathcal{T} = \frac{(Z' Z)^{-\frac{1}{2}} Z' \hat{\tilde{Y}} \hat{\Omega}^{-1} a_0}{(a_0 \hat{\Omega}^{-1} a_0)^{1/2}},$$  \hspace{1cm} (24)

where $\hat{\Omega} = \hat{\tilde{Y}} M_Z \hat{\tilde{Y}}/(T - K_1 - K_2)$ is a consistent estimator of $\Omega$. In this paper, I consider three Gaussian similar tests:

1. The Anderson-Rubin (AR) test (Anderson and Rubin, 1949) based on the statistic

$$AR(\beta_0) = \frac{\mathcal{F}' \mathcal{F}}{K_2},$$  \hspace{1cm} (25)

which is asymptotically distributed $\chi^2_{K_2}/K_2$ under the null (Staiger and Stock, 1997, theorem 5).
2. The Lagrange multiplier (LM) test (Kleibergen, 2002) based on the statistic

$$LM(\beta_0) = \frac{(\mathcal{F}' \mathcal{F})^2}{\mathcal{T}' \mathcal{T}},$$  \hspace{1cm} (26)

which is asymptotically distributed $\chi^2_1$ under the null (Kleibergen, 2002, theorem 1).
3. The conditional likelihood ratio (LR) test (Moreira, 2003) based on the statistic

$$LR(\beta_0) = \frac{1}{2} \left((\mathcal{F}' \mathcal{F} - \mathcal{T}' \mathcal{T}) + \sqrt{(\mathcal{F}' \mathcal{F} + \mathcal{T}' \mathcal{T})^2 - 4(\mathcal{F}' \mathcal{F})(\mathcal{T}' \mathcal{T}) - (\mathcal{F}' \mathcal{F})^2} \right),$$  \hspace{1cm} (27)

whose critical values can be computed as a function of $K_2$ and $\mathcal{T}' \mathcal{T}$ by Monte Carlo simulation.

It is well known that the AR test is invariant to linear transformations of the GMM moment restriction. In the

\footnote{For simplicity, I use the terminology “similar test” instead of “asymptotically similar test,” hopefully without confusion.}
context of the moment restrictions (2) and (4), the AR test rejects the null hypothesis \( \psi = \psi_0 \) based on equation (2) if and only if it rejects \( 1/\psi = 1/\psi_0 \) based on equation (4). It can be readily verified that the LM and the conditional LR tests share this invariance property. In contrast, two-step GMM is not invariant to linear transformations of the moment restriction, which results in contradictory inference about the EIS depending on whether one uses the moment restriction (2) or (4).

These similar tests can be inverted to construct confidence regions for \( \beta \). For instance, one can construct a \((1 - \alpha)100\%\) confidence region based on the AR test as

\[
\{ \beta_0 \in \mathcal{B} | AR(\beta_0) < \chi^2_{k_2,\alpha}/K_2 \},
\]

where \( \mathcal{B} \) is the parameter space for \( \beta \) and \( \chi^2_{k_2,\alpha} \) is the upper \( \alpha \)-quantile of the \( \chi^2_{k_2} \) distribution. If the parameter space is unrestricted, \( \mathcal{B} \) is the set of reals; if \( \beta \) is restricted to be positive, which may be the case for the EIS, \( \mathcal{B} \) is the set of positive reals. In general, these confidence regions can consist of disjoint intervals. By taking the minimum and maximum values of \( \beta \) in the confidence region, one obtains a confidence interval that has coverage of at least \((1 - \alpha)100\%\).

D. Power of Similar Tests

Because more powerful tests lead to tighter confidence intervals, one would like to use the most powerful similar test to construct confidence intervals that are robust to weak instruments. Unfortunately, there is no uniformly most powerful test (Moreira, 2001), so the three similar tests have power advantages in different regions of the parameter space. In this section, I discuss their power properties.

A natural way to evaluate the power of similar tests is to consider their asymptotic power under weak-instrument asymptotics. Asymptotically, the power functions depend on the scaled concentration parameter \( \mu^2/K_2 \) and the degree of endogeneity \( \rho = \Sigma_{v_v}/(\Sigma_{v_v} \sigma_{uu})^{1/2} \). In an earlier version of this paper, I reported the power functions, which have since been published by Stock et al. (2002, figure 2). To avoid redundancy, I refer to their figures in the following discussion. The appendix details the asymptotic results used to plot the power functions, which were omitted from Stock et al. (2002) to save space.

Stock et al. plot the power functions for two levels of instrument relevance \((\mu^2/K_2 = 1, 5)\) and two levels of endogeneity \((\rho = 0.5, 0.99)\). When instruments are very weak \((\mu^2/K_2 = 1)\), all three similar tests have poor power in the sense that the power is far less than 1 even at distant alternatives. As a consequence, confidence intervals based on similar tests can be unbounded when instruments are very weak. When \( \rho = 0.5 \), the AR and conditional LR tests have better power than the LM test. When \( \rho = 0.99 \), the LM and conditional LR tests have better power than the AR test. When instruments are moderately weak \((\mu^2/K_2 = 5)\), the similar tests have much better power; their power approaches 1 for distant alternatives. Among the three similar tests, the conditional LR test has the best power properties; its power function comes close to the power envelope for similar tests at both values of \( \rho \). This suggests that the conditional LR test should usually result in confidence intervals that are tightest among the three similar tests.

E. Heteroskedasticity and Simultaneous Estimation

In this section, I discuss econometric methods for the more general GMM setting. This generalization allows for conditional heteroskedasticity and simultaneous estimation. For instance, if the linearized Euler equation (8) holds for both the interest rate and stock return, the GMM allows for simultaneous estimation of the EIS using both moment restrictions. The cost of going to the more general setup is that methods designed to handle weak instruments are much less developed. Generalizations of the test to weak instruments and similar tests for the GMM are topics of ongoing research.

To be concrete with notation, suppose there are two \( K_2 \times 1 \) vectors of linear moment restrictions, which is the relevant case for this paper. Let \( y_1, y_2, Y_1 \), and \( Y_2 \) be \( T \times 1 \) vectors of \( T \) observations on jointly endogenous variables. As before, \( Z \) is a \( T \times K_2 \) matrix of instruments, and superscript \( \perp \) denotes the residual from projection onto the included exogenous regressors \( X \). To simplify notation, let \( X \) include a column of 1’s so that the residuals from the projection have mean zero. Define the \( 2K_2 \times 1 \) vector

\[
\phi(\beta) = \left[ \begin{array}{c} Z_1^\top(y_1 - Y_1\beta) \\ Z_2^\top(y_2 - Y_2\beta) \end{array} \right].
\]

The moment restriction is \( E[\phi(\beta_0)] = 0 \). Define the heteroskedasticity robust weighting matrix

\[
V(\beta) = T^{-1} \sum_{t=1}^T \phi(\beta)\phi(\beta)'
\]

and the objective function

\[
S(\beta, \tilde{\beta}) = \left[ T^{-1/2} \sum_{t=1}^T \phi(\beta) \right] V(\beta)^{-1/2} \left[ T^{-1/2} \sum_{t=1}^T \phi(\beta) \right].
\]

The efficient two-step GMM estimator \( \tilde{\beta}_2 \) minimizes \( S(\beta, \tilde{\beta}) \) for some consistent first-step estimator \( \tilde{\beta}_1 \). For example, the first-step estimator can be obtained by minimizing equation (30) using the weighting matrix \( V(\beta) = I_2 \otimes 3 \) Figure 2 in Stock et al. (2002) is for \( K_2 = 5 \) instruments, whereas \( K_2 = 4 \) in the empirical application of this paper. However, this is not a substantive difference, because the power functions for \( K_2 = 4 \) and \( K_2 = 5 \) are essentially the same.
(Z^{1}_tZ^{2}_t). Because the objective function is quadratic in this case, the two-step estimator has the closed form

\[ \hat{\beta}_2 = \left( \begin{bmatrix} Z^{1}_t & Y^{1}_t \\ Z^{2}_t & Y^{2}_t \end{bmatrix} \right)^{-1} \begin{bmatrix} Z^{1}_t \hat{y}^{1}_t \\ Z^{2}_t \hat{y}^{2}_t \end{bmatrix} \times \left( \begin{bmatrix} Z^{1}_t & Y^{1}_t \\ Z^{2}_t & Y^{2}_t \end{bmatrix} \right)^{-1} \begin{bmatrix} Z^{1}_t \hat{y}^{1}_t \\ Z^{2}_t \hat{y}^{2}_t \end{bmatrix}. \]  

(31)

An alternative to the two-step GMM is the continuous updating estimator (CUE), which minimizes the objective function \( S(\beta, \hat{\beta}) \). In a Monte Carlo experiment designed to simulate estimation of the linearized Euler equation, Hansen, Heaton, and Yaron (1996) find that the CUE is less biased and its confidence intervals have better coverage rates than the two-step GMM. The intuition for this result is that the CUE is a generalization of LIML to GMM, just as the two-step GMM is a generalization of TSLS.

In the GMM setting, the analog of weak instruments is

\[ \text{weak identification (Stock and Wright, 2000),} \]

loosely speaking, occurs if \[ \text{E}[\phi_{t}(B)] \approx 0 \text{ even when } \beta \neq \beta_0. \]

Confidence intervals for \( \beta \) with the correct coverage can be constructed by inverting the objective function (30) evaluated at \( \beta_0 \). By Stock and Wright (2000, theorem 2), \( S(\beta_0, \hat{\beta}_0) \) is asymptotically distributed \( \chi_{2}^2 \). This test, which I refer to as the \( S \)-test, is a generalization of the AR test to the GMM.

IV. Empirical Results

A. Data

The data set that I use is from Campbell (2003). It consists of quarterly data on equity markets at an aggregate level and macroeconomic variables for eleven developed countries: Australia (AUL), Canada (CAN), France (FR), Germany (GER), Italy (ITA), Japan (JAP), Netherlands (NTH), Sweden (SWD), Switzerland (SWT), the United Kingdom (UK), and the United States (USA). In addition, a longer time series is available at annual frequency for Sweden, the United Kingdom, and the United States. The primary sources of international data are Morgan Stanley Capital International and the International Financial Statistics of the International Monetary Fund. The sample periods vary by country and frequency, as reported in table 1. With the exception of the United States, quarterly data are only available starting in 1970. For the quarterly U.S. series, I report the results for both the full sample, which starts in 1947, and a truncated sample that starts in 1970. For a full description of the data set, see Campbell (2003) and the accompanying data appendix (Campbell, 1998).

For each country, I estimate the EIS using two asset returns: the real interest rate, denoted by \( r_t \), and the real aggregate stock return, denoted by \( r_s \). The real stock return is constructed as log of the gross stock return deflated by the consumer price index. The real interest rate is constructed in the same way, using an available proxy for the short-term interest rate. Real consumption growth is the first difference in log real consumption per capita. For all quarterly series except for the U.S., the consumption measure is total consumption rather than nondurables and services, in view of data availability. The timing convention used for consumption is beginning of the period, following Campbell (2003). In other words, I assume that the consumption for a given time period is the flow measured at the beginning of the period rather than at the end.

B. Test for Weak Instruments

The coefficients of interest are the EIS \( \psi \), estimated by equation (1), and its reciprocal \( 1/\psi \), estimated by equation (3). The instruments that I use for the endogenous regressor \( \Delta c_{t+1} \) in equation (3) and \( r_{t+1} \) in equation (1) are the nominal interest rate, inflation, consumption growth, and log dividend-price ratio. All instruments are lagged twice to avoid problems with time aggregation in consumption data (Hall, 1988). As discussed in section II, this also assures that instruments are exogenous even if consumption or asset returns are conditionally heteroskedastic.

Assuming that the error is conditionally homoskedastic, equations (1) and (3) can be estimated by TSLS. In table 1, I report the first-stage F-statistic for each of the possible endogenous regressors (consumption growth, interest rate, and stock return), which is the relevant statistic to test for weak instruments. Next to the F-statistic, I report the p-values of the test. A p-value less than 0.05 means that the test would reject the null hypothesis of weak instruments at the 5% significance level. As explained in section III, the p-value depends on the type of estimator (TSLS, Fuller-k, or LIML) used for estimation or inference.

At the quarterly frequency, consumption growth and stock return both have low predictability as evidenced by low F-statistics, so the test fails to reject the null of weak instruments. Hence, a researcher should suspect that the TSLS estimator is biased and the TSLS t-test is size-distorted. In fact, instruments are so weak in this case that estimation and inference based on Fuller-k or LIML are also suspect. On the other hand, the interest rate appears to be more predictable for all countries. The F-statistic is large enough that the Fuller-k estimator is approximately unbiased. In addition, the LIML t-test leads to approximately correct inference, although the TSLS t-test may be size-distorted.

For the annual series, none of the regressors appear to be sufficiently predictable to avoid problems with weak instruments. The only possible exceptions are the United Kingdom and the United States, where the instruments are

4 Campbell (2003) uses the real interest rate, instead of the nominal interest rate and inflation, for a total of three instruments. For many countries, the nominal interest rate appears to contain important information about future real asset returns.
somewhat relevant in predicting the interest rate. The LIML $t$-test should lead to approximately correct inference, because the test for weak instruments rejects at the 10% level.

### C. Estimates of the EIS Using the Interest Rate

In the first three columns of Table 2, I report the point estimate and standard error of $1/\lambda$ with the interest rate as the dependent variable in equation (3). I report results using TSLS, Fuller-$k$, and LIML. The first fact to note is that the three estimators give very different results. Under conventional first-order asymptotics, the three estimators have the same asymptotic distribution. Therefore, the fact that the three estimators give very different results is indirect evidence for weak instruments. In general, the magnitude of both the coefficient and the standard error increases from TSLS to Fuller-$k$ and from Fuller-$k$ to LIML. The 95% confidence intervals for $1/\lambda$ based on these estimators in-
include rather large values of the EIS. In particular, one cannot reject the null hypothesis \( \psi = 1 \), except for Canada and Switzerland.

In the last three columns of table 2, I report estimates of the EIS using equation (1) with the interest rate as the endogenous regressor. In contrast to inference based on equation (3), which requires that the instruments predict consumption growth, weak instruments are not a problem, because the interest rate is sufficiently predictable, as documented in table 1. Consequently, the three estimators give very similar coefficients and standard errors. The point estimates of \( \psi \) are small, although sometimes negative. The 95% confidence intervals based on these estimators reject large values of the EIS, in particular 1.

To summarize the results in table 2, one would conclude that the EIS is small and significantly less than 1 whereas its inverse is not significantly different from 1. The hypothesis \( \psi = 1 \) is of economic interest because with Epstein-Zin preferences, an investor’s optimal consumption choice is a constant fraction of wealth when the EIS is equal to 1. Moreover, in the special case of power utility where the EIS is equal to the reciprocal of the risk aversion, \( \gamma = 1/\psi = 1 \) leads to myopic portfolio choice (see Campbell and Viceira, 2002, chapter 2). This apparent empirical puzzle, emphasized by Neely et al. (2001), can be accounted for by weak instruments. The regression equation (3) leads to biased estimates and confidence intervals with poor coverage because the instruments cannot predict consumption growth adequately to identify \( 1/\psi \). On the other hand, estimation by equation (1) leads to valid inference, because the instruments are not weak for the interest rate.

The sensitivity of inference to the particular normalization of the moment restriction is an unattractive property of \( k \)-class estimators.\(^5\) In contrast, confidence intervals based on the similar tests (AR, LM, and conditional LR) are invariant to this normalization. Moreover, because these methods are fully robust to weak instruments, there is no

\(^5\) The point estimate of LIML is invariant to normalization, but its confidence interval is not.
need for a pretest to make sure that the instruments are relevant.

In table 3, I report the 95% confidence intervals for the EIS constructed from the similar tests. Of the three similar tests, the conditional LR test tends to give the tightest confidence intervals, consistent with the fact that it has the best power properties. Focusing on the quarterly series and the conditional LR confidence interval, the EIS is less than 0.5 across all eleven countries. For the annual series, the EIS is similarly small for Sweden and the United Kingdom. There appears to be identification failure for the annual U.S. series, as evidenced by the uninformative confidence intervals $[-\infty, \infty]$ for both the LM and the conditional LR tests, although the AR test gives small estimates of the EIS. In summary, the weak-instrument-robust confidence intervals indicate that the EIS is small and not significantly different from 0 for the eleven developed countries.

### D. Estimates of the EIS Using the Stock Return

In tables 4 and 5, I report the same set of results as tables 2 and 3, using the stock return instead of the interest rate. As demonstrated in table 1, both consumption growth and stock return are difficult to predict, so either normalization of the moment restriction, equation (2) or (4), runs into problems with weak instruments. This is evidenced by the fact that most of the weak-instrument-robust confidence intervals in table 5 are uninformative. For a few of the countries (Canada, France, and Japan at quarterly frequency, and the United Kingdom at annual frequency), the confidence intervals are informative. Note that these are precisely the series for which the first-stage $F$-statistics for stock return were relatively large in table 1, ranging from 2.51 to 4.18. The confidence intervals indicate that the EIS is small, agreeing with the results for the interest rate in table 3.

In tables 3 and 5, I have reported the unconstrained confidence intervals. Constrained confidence intervals that restrict the EIS to be nonnegative can usually be obtained by truncating the unconstrained confidence intervals at zero. Although the truncated confidence interval has the correct coverage rate provided that $\psi \geq 0$, it may be conservative when identification is sufficiently weak so that the confidence region is disjoint. In other words, the truncated confidence interval may not coincide with the actual constrained confidence interval. In table 5, this occurs only for the quarterly (1970.3–1998.4) U.S. series, where the actual constrained confidence intervals are $[0.05, \infty]$ and $[0.02, \infty]$ for the AR and conditional LR tests, respectively, because these tests reject $\psi = 0$.

### E. Heteroskedasticity and Simultaneous Estimation

In the first two columns of table 6, I report GMM and CUE estimates of the EIS using the moment restriction (2) for the interest rate. These estimates are heteroskedasticity-robust versions of the estimates by TSLS and LIML, reported in table 2. Comparing GMM and TSLS, the point estimates and standard errors are quite similar, so heteroskedasticity does not appear to be a dominant feature of the data. Likewise, the LIML and CUE estimates are quite similar. In the third column, I report weak-instrument-robust confidence intervals computed by inverting the S-test. These are heteroskedasticity-robust versions of the AR confidence intervals reported in table 3. In general, these confidence intervals are comparable to those reported in table 3, indicating that the EIS is small and less than 1. The only exceptions are the United States, which contains only negative values in the confidence interval, and Germany, whose confidence interval cannot exclude 1.

The last three columns of table 6 report the same set of results using the moment restriction (2) for both the interest rate and the stock return. With four instruments and two assets, the moment restriction has dimension eight. Note that GMM gives point estimates and standard errors that are very small. This appears to be a consequence of weak identification from the low correlation between the instruments and the stock return. The CUE, which is more robust to weak instruments than the GMM, gives estimates that are similar to those obtained using just the interest rate. Likewise, the weak-instrument-robust confidence intervals based on the S-test are similar to those obtained using just the interest rate. The confidence intervals are actually slightly wider for many of the series. This is not surprising in light of table 5, which showed that the moment restriction implied by the stock return contains little information that is useful for identifying the EIS.

For the United States, the model is entirely rejected, as indicated by empty confidence intervals. One possible explanation of this result is that the Euler equation for stock return does not hold for the representative consumer because of limited participation in asset markets. Vissing-Jørgensen (2002) has found some evidence for this theory using the Consumer Expenditure Survey.

### Table 3: Weak-Instrument-Robust Confidence Intervals for the EIS Using the Interest Rate

<table>
<thead>
<tr>
<th>Country</th>
<th>Sample Period</th>
<th>AR LM Cond. LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>1947.3–1998.4</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>AUL</td>
<td>1970.3–1998.4</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>CAN</td>
<td>1970.3–1999.1</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>FR</td>
<td>1970.3–1998.3</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>GER</td>
<td>1970.1–1998.3</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>ITA</td>
<td>1971.4–1998.1</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>JAP</td>
<td>1970.3–1998.4</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>NTH</td>
<td>1977.3–1998.4</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>SWE</td>
<td>1970.3–1999.2</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>UK</td>
<td>1970.3–1999.1</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>USA</td>
<td>1970.3–1998.4</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>SWD</td>
<td>1921–1994</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>UK</td>
<td>1921–1994</td>
<td>(0.01, 0.70)</td>
</tr>
<tr>
<td>USA</td>
<td>1891–1995</td>
<td>(0.01, 0.70)</td>
</tr>
</tbody>
</table>

The table reports 95% confidence intervals for the EIS, constructed from AR, LM, and conditional LR tests. $\Box$ indicates an empty confidence interval. The instruments are twice lagged nominal interest rate, inflation, consumption growth, and log dividend-price ratio.
F. Implications for the Equity Premium Puzzle

In the power utility model, the coefficient of relative risk aversion is equal to the reciprocal of the EIS. In that case,

Table 5.—Weak-Instrument-Robust Confidence Intervals for the EIS Using the Stock Return

<table>
<thead>
<tr>
<th>Country</th>
<th>Sample Period</th>
<th>AR</th>
<th>LM</th>
<th>Cond. LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>1947.3–1998.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUL</td>
<td>1970.3–1998.4</td>
<td>−1.75</td>
<td>−12.94</td>
<td>−0.03</td>
</tr>
<tr>
<td>CAN</td>
<td>1970.3–1999.1</td>
<td>5.28</td>
<td>29.64</td>
<td>0.04</td>
</tr>
<tr>
<td>FR</td>
<td>1970.3–1999.3</td>
<td>−0.40</td>
<td>−0.29</td>
<td>0.05</td>
</tr>
<tr>
<td>ITA</td>
<td>1971.4–1998.1</td>
<td>1.92</td>
<td>8.05</td>
<td>0.03</td>
</tr>
<tr>
<td>JAP</td>
<td>1970.3–1998.4</td>
<td>10.16</td>
<td>17.20</td>
<td>0.05</td>
</tr>
<tr>
<td>NTH</td>
<td>1977.3–1998.4</td>
<td>1.29</td>
<td>4.20</td>
<td>0.03</td>
</tr>
<tr>
<td>SWD</td>
<td>1973.3–1999.2</td>
<td>−5.87</td>
<td>−64.89</td>
<td>−0.01</td>
</tr>
<tr>
<td>SWT</td>
<td>1976.2–1998.4</td>
<td>−0.35</td>
<td>−0.29</td>
<td>0.05</td>
</tr>
<tr>
<td>UK</td>
<td>1970.3–1999.1</td>
<td>−0.68</td>
<td>−9.24</td>
<td>0.01</td>
</tr>
<tr>
<td>USA</td>
<td>1970.3–1998.4</td>
<td>6.92</td>
<td>8.05</td>
<td>0.03</td>
</tr>
</tbody>
</table>

The reciprocal of the EIS is estimated from \( \frac{1}{1 + \eta} = \kappa + \frac{1}{\psi} \Delta_1 + \psi \Delta_2 + \xi \), and the EIS is estimated from \( \Delta_1 = \kappa + \psi \Delta_2 + \xi \). The instruments are the Tweedie lognormal interval rate, inflation, consumption growth, and log dividend-price ratio. Standard errors in parentheses.

the small estimates of the EIS reported in this paper are evidence for large values of risk aversion. For instance, the confidence intervals in table 3 indicate that the EIS is in the range \([0, 0.5]\) across the eleven countries, which implies that risk aversion is in the range \([2, \infty]\). Though this is consistent with evidence from the large literature on the equity premium puzzle (Mehra and Prescott, 1985), I hesitate to draw conclusions about risk aversion based on the estimates of the EIS.

V. Conclusion

The econometric lesson to take away from this paper is that weak instruments are relevant in practice and that conventional t-tests can lead to misleading inference. Various methods are now available for handling weak instruments, from the simple pretest based on the first-stage F-statistic to fully robust confidence intervals based on similar tests. These methods are not necessarily a cure for weak instruments, for the resulting confidence intervals are often uninformative when identification is poor; but they prevent the researcher from making erroneous inferences.
The economic lesson to take away from this paper is that the EIS is small and not significantly different from 0. In particular, the EIS appears to be less than 1, which implies that an investor’s optimal consumption-wealth ratio is increasing in expected returns. In my preferred estimates, reported in table 3, the upper end of the 95% confidence interval for the EIS is never greater than 0.5 across eleven developed countries. For the United States, the value is approximately 0.2, which remarkably agrees with Hall (1988, p. 350): “My overall conclusion . . . is that the evidence points in the direction of a low value for the intertemporal elasticity. The value may even be zero and is probably not above .2.”

The references cited in the document include:


REFERENCES

Table 6.—Heteroskedasticity-Robust Estimates of the EIS

<table>
<thead>
<tr>
<th>Country</th>
<th>Sample Period</th>
<th>Interest Rate (GMM)</th>
<th>95% CI</th>
<th>Interest Rate (CUE)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>1947.3–1998.4</td>
<td>0.05 (0.09)</td>
<td>−0.11 (0.11)</td>
<td>−0.14, −0.08 (0.00)</td>
<td>−0.36 (0.09)</td>
</tr>
<tr>
<td>AUL</td>
<td>1970.3–1998.4</td>
<td>0.09 (0.12)</td>
<td>0.08 (0.12)</td>
<td>−0.17, 0.30 (0.01)</td>
<td>0.12 (0.04)</td>
</tr>
<tr>
<td>CAN</td>
<td>1973.3–1999.1</td>
<td>−0.34 (0.17)</td>
<td>−0.33 (0.17)</td>
<td>−0.77, 0.11 (0.02)</td>
<td>−0.19 (0.06)</td>
</tr>
<tr>
<td>FR</td>
<td>1973.3–1998.3</td>
<td>−0.12 (0.14)</td>
<td>−0.12 (0.14)</td>
<td>−0.57, 0.36 (0.00)</td>
<td>−0.16 (0.05)</td>
</tr>
<tr>
<td>GER</td>
<td>1979.1–1998.3</td>
<td>−0.44 (0.43)</td>
<td>−0.48 (0.43)</td>
<td>−1.95, 1.63 (0.00)</td>
<td>−0.55 (0.20)</td>
</tr>
<tr>
<td>ITA</td>
<td>1971.4–1998.1</td>
<td>−0.08 (0.08)</td>
<td>−0.07 (0.08)</td>
<td>−0.34, 0.20 (0.00)</td>
<td>−0.09 (0.05)</td>
</tr>
<tr>
<td>JAP</td>
<td>1973.3–1998.4</td>
<td>−0.18 (0.21)</td>
<td>−0.21 (0.21)</td>
<td>−0.93, 0.39 (0.00)</td>
<td>−0.25 (0.08)</td>
</tr>
<tr>
<td>NTH</td>
<td>1977.3–1998.4</td>
<td>−0.25 (0.20)</td>
<td>−0.28 (0.20)</td>
<td>−0.57, 0.09 (0.00)</td>
<td>−0.25 (0.09)</td>
</tr>
<tr>
<td>SWD</td>
<td>1973.3–1999.2</td>
<td>0.01 (0.09)</td>
<td>0.00 (0.09)</td>
<td>−0.28, 0.28 (0.00)</td>
<td>−0.02 (0.01)</td>
</tr>
<tr>
<td>SWT</td>
<td>1976.2–1998.4</td>
<td>−0.39 (0.25)</td>
<td>−0.41 (0.25)</td>
<td>−1.42, 0.50 (0.22)</td>
<td>−0.44 (0.24)</td>
</tr>
<tr>
<td>UK</td>
<td>1973.3–1999.1</td>
<td>0.22 (0.12)</td>
<td>0.28 (0.12)</td>
<td>−0.45, 0.51 (0.00)</td>
<td>0.17 (0.07)</td>
</tr>
<tr>
<td>USA</td>
<td>1973.3–1998.4</td>
<td>0.02 (0.08)</td>
<td>−0.09 (0.09)</td>
<td>−0.14, −0.02 (0.01)</td>
<td>−0.05 (0.02)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Country</th>
<th>Sample Period</th>
<th>Interest Rate (GMM)</th>
<th>95% CI</th>
<th>Interest Rate &amp; Stock Return (GMM)</th>
<th>95% CI</th>
<th>Interest Rate &amp; Stock Return (CUE)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWD</td>
<td>1921–1994</td>
<td>0.00 (0.10)</td>
<td>−0.09 (0.10)</td>
<td>−0.51, 0.51 (0.00)</td>
<td>−0.07 (0.04)</td>
<td>−0.61, 0.74</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>1921–1994</td>
<td>0.25 (0.09)</td>
<td>0.27 (0.09)</td>
<td>[0.01, 0.64] (0.03)</td>
<td>0.39 (0.08)</td>
<td>−0.05, 0.82</td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>1891–1995</td>
<td>−0.02 (0.06)</td>
<td>−0.01 (0.06)</td>
<td>−0.21, 0.15 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

The table reports the EIS estimated by two-step GMM and CUE with standard errors in parentheses. The 95% confidence interval is constructed from the S-test. Ø indicates an empty confidence interval. The instruments are the twice lagged nominal interest rate, inflation, consumption growth, and log dividend-price ratio.
**ESTIMATING THE ELASTICITY OF INTERTEMPORAL SUBSTITUTION**


**APPENDIX**

This appendix derives the asymptotic distributions of similar tests under weak-instrument asymptotics (Staiger and Stock, 1997). The asymptotic representations can then be used to plot the power functions for the similar tests that appear in Stock et al. (2002, figures 2, 3). Let denote convergence in probability, and denote convergence in distribution. Following Staiger and Stock, I make the following assumptions.

**ASSUMPTION 1 (LOCAL-TO-ZERO):** $\Pi = C/\sqrt{T}$, where $C$ is a fixed $K_x \times 1$ vector.

**ASSUMPTION 2 (MOMENT CONDITIONS):** The following limits hold jointly:

1. $(u' T V u' T V V') \to (\sigma_{uu} \Sigma_{VV} \Sigma_{VV})$;

2. $Z'/T \to Q = \begin{bmatrix} Q_x & Q_z \end{bmatrix}$.

3. $(X'/U/V/ \sqrt{T}, X'/V/ \sqrt{T}, Z'/V/ \sqrt{T}) \to d (\Psi_{XX}, \Psi_{ZZ}, \Psi_{XZ}, \psi_{ZV})$, where $\Psi = (\Psi_{XX}, \Psi_{ZZ}, \Psi_{XZ}, \psi_{ZV})' - N(0, \Sigma \otimes \Omega)$.

As noted by Staiger and Stock, assumption 2 can be derived from weak instrumental assumptions that are reasonable in the present context of estimating the linearized Euler equation. Let $Y = Q_{ZZ} - Q_{ZQ}(Q_{XX}Q_{ZZ}$ and $\lambda = Y^{-1/2} \Sigma_{VV}$. Define the vector

$$z_v = \begin{bmatrix} Y^{-1/2} \Sigma_{ZZ} - Q_{ZQ}Q_{XX} \Sigma_{VV} \sigma_{u \Sigma_{VV}} & Y^{-1/2} \Sigma_{ZZ} - Q_{ZQ}Q_{XX} \Sigma_{VV} \sigma_{u \Sigma_{VV}} \\
\end{bmatrix} 
- \begin{bmatrix} 0, 1 \rho 
\end{bmatrix} \otimes I_k).$$

Under assumptions 1 and 2, and Staiger and Stock (1997, theorem 1(e)) show that $d \to (\lambda + \bar{z}_v)'(\lambda + \bar{z}_v)/K_x$. In other words, the first-stage $F$-statistic is asymptotically $Q_{XX}(1)$ under weak-instrument asymptotics. In contrast, the $F$-statistic becomes arbitrarily large under the conventional first-order asymptotics with fixed $\Pi$.

The following lemma derives the weak instrument asymptotic distributions of the statistics $\mathcal{J}$ and $\mathcal{J}*$ (see equations (23) and (24)).

**Lemma 1.** Suppose that Assumptions 1 and 2 hold. Let $\bar{\lambda} = \sigma_{u \Sigma_{VV}} \Sigma_{VV} (b_0 - \beta)$ and $S_1(s) = 1 - 2\rho s + s^2$ for any scalar $s$. Then

$$d \to \mathcal{J} = \frac{\lambda + \bar{z}_v - \bar{\lambda}}{S_1(\bar{\lambda})^{1/2}} - N(-\bar{\lambda}/S_1(\bar{\lambda})^{1/2}, I_k).$$

$$d \to \mathcal{J}^* = \frac{(\lambda + \bar{z}_v) + \rho \bar{\lambda} - \rho \bar{\lambda} - (\lambda + z_v)(\lambda + \bar{\lambda})}{(1 - \rho^2)S_1(\bar{\lambda})^{1/2}} - N\left(\frac{\lambda \rho \bar{\lambda}}{(1 - \rho^2)S_1(\bar{\lambda})^{1/2}}, I_k\right).$$

where $\mathcal{J}$ and $\mathcal{J}^*$ are independent.

**Proof:** Applying Staiger and Stock (1997, lemma A1), we have $\mathcal{J} \to \Omega$ and $(Z'/T Z')^{-1/2} Z'/T Y' \to w'/T$, where

$$w = \begin{bmatrix} z_v + \sigma_{u \Sigma_{VV}} \Sigma_{VV} / \rho (\lambda + z_v) \\
\sigma_{u \Sigma_{VV}} \Sigma_{VV} / \rho (\lambda + z_v) \end{bmatrix}.$$ 

This then implies that

$$d \to \mathcal{J} = \frac{w' b_0}{(b_0') b_0)^{1/2}}.$$ (A-4)
\[ \mathcal{F} \xrightarrow{d} \mathcal{F}^* = \frac{w' \Omega^{-1} a_0}{(a_0' \Omega^{-1} a_0)^{1/2}}. \] (A-5)

Establishing the equivalence of these expressions to those that appear in the statement of the lemma requires the intermediate steps \( b_0' \Omega b_0 = \sigma w S_1(\Delta), \ a_0' \Omega^{-1} a_0 = [(1 - \rho^2) \Sigma_{VV}]^{-1} S_1(\Delta), \) and

\[ \Omega^{-1} a_0 = \frac{1}{(1 - \rho^2) \Sigma_{VV}} \begin{bmatrix} \sigma \nu \Sigma_{VV} \nu (\Delta - \bar{\Delta}) \\ 1 - \rho \Delta - \sigma \nu \Sigma_{VV} (\bar{\Delta} - \rho \bar{\Delta}) \end{bmatrix}. \]

Note that \( \mathcal{F} \) is asymptotically pivotal and independent of \( \mathcal{F}^* \) under the null hypothesis (\( \Delta = 0 \)). A straightforward application of Lemma 1 to the AR statistic (25) results in

\[ AR(b_0) \xrightarrow{d} \mathcal{F}^* \sim \frac{\chi^2(\Delta' \Lambda \Delta S_1(\Delta))}{\Delta' \Lambda \Delta S_1(\Delta)} \]. (A-6)

which was shown by Staiger and Stock (1997, theorem 5). The asymptotic distributions of the LM and LR statistics can similarly be obtained by application of lemma 1 to equations (26) and (27), respectively.

Note that the asymptotic distributions of AR, LM, and LR statistics are completely determined by the matrix \( \left[ \Delta' \Lambda \Delta S_1(\Delta) \right] \), which has a noncentral Wishart distribution \( \mathcal{W}_2(\Delta^2, 2, I_2) \) (see Phillips, 1983) with noncentrality matrix

\[ \Lambda = \Lambda' \Lambda = \begin{bmatrix} \Delta^2 & -\Delta \Delta^2 \\ \Delta \Delta^2 & (1 - \rho^2) \Delta^2 \end{bmatrix}. \] (A-7)

Hence, the asymptotic distributions only depend on the number of instruments \( K_2 \), the concentration parameter \( \lambda' \lambda \), the degree of simultaneity \( \rho \), and \( \Delta \). The parameter \( \Delta \) has a natural interpretation as the distance between the null hypothesis \( \Delta = 0 \) and the true value \( \Delta \) when the IV regression model, (14) and (15), is normalized to have unit variance.