THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 30_09

Ramazan Gençay
Simon Fraser University, Canada
and
The Rimini Centre for Economic Analysis, Italy

Nikola Gradojevic
Lakehead University, Canada

ERRORS-IN-VARIABLES ESTIMATION WITH NO INSTRUMENTS

The Rimini Centre for Economic Analysis
Legal address: Via Angherà, 22 – Head office: Via Patara, 3 - 47900 Rimini (RN) – Italy
www.rcfea.org - secretary@rcfea.org

Errors-in-Variables Estimation with No Instruments

Ramazan Gençay* Nikola Gradojevic†

June 2009

Abstract

This paper develops a wavelet (spectral) approach to estimate the parameters of a linear regression model where the regressand and the regressors are persistent processes and contain a measurement error. We propose a wavelet filtering approach which does not require instruments and yields unbiased and consistent estimates for the intercept and the slope parameters. Our Monte Carlo results also show that the wavelet approach is particularly effective when measurement errors for the regressand and the regressor are serially correlated. With this paper, we hope to bring a fresh perspective and stimulate further theoretical research in this area.

Keywords: Cointegration, discrete wavelet transformation, maximum overlap wavelet transformation, energy decomposition, errors-in-variables, persistence

JEL No: C1, C2, C12, C22, F31, G0, G1

*Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada. Email: gencay@sfu.ca; Rimini Center for Economic Analysis, Italy and IRMACS, Canada. Ramazan Gençay is grateful to the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada for research support.

†Faculty of Business Administration, Lakehead University, 955 Oliver Road, Thunder Bay, ON P7B 5E1, Canada. Email: ngradoje@lakeheadu.ca; University of Novi Sad, Faculty of Technical Sciences, Serbia. Nikola Gradojevic gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada.

1. Introduction

The problem of measurement error is often ignored, but its implications for standard methods of statistical inference are potentially devastating. In an extreme case, one can argue that in the presence of errors-in-variables the estimates of parameters of interest cannot be given any structural interpretation (Cragg, 1994, 1997). The objective of this paper is to overcome the problem of errors-in-variables by using a novel methodology and provide simulation evidence that demonstrate its effectiveness, when the regressand and the regressor are persistent processes. We propose a wavelet (spectral) approach that does not require instruments and yields unbiased and consistent estimates for the intercept and the slope parameters. Specifically, we filter out the measurement noise and use the filtered regressand and regressors in ordinary least squares (OLS). Furthermore, our approach is particularly suited for highly persistent, i.e., near-integrated regressors[1], and is able to compensate for any biases coming from such processes. We find the instrumental variable (IV) and OLS estimators very sensitive to serially correlated measurement errors in the regressor and the regressand while the wavelet approach presents a more robust framework, especially when the serial correlation increases.

This paper can also be viewed as an extension of the work by Gençay and Fan (2009) that applied wavelets to test the presence of a unit root in a stochastic process. By using Monte Carlo simulations, they demonstrated the comparable power of the wavelet-based tests with reasonable empirical sizes. Similar to Gençay and Fan (2009), we address another important econometric problem with the goal to inspire further theoretical research in this area. We find it worthwhile to emphasize the simplicity of our approach and the fact that our Monte Carlo findings are robust to various specifications of the wavelet filters.

In general, errors of measurement produce biased and inconsistent OLS estimators. More specifically, in the context of a simple bivariate linear model, Cragg (1994) stresses two effects: *attenuation* and *contamination*. The former denotes the slope coefficient being biased towards zero while the latter refers to a bias in the intercept of the opposite sign when the average of the explanatory variable is positive. It may appear that the attenuation effect is not harmful as long as the slope coefficient is significantly different from zero - at worst the estimate will be more conservative than it should be. Moreover, as the magnitude of the attenuation effect is inversely related to the R^2 value, it can be concluded that high R^2 regressions exhibit negligible biases. However, (Cragg, 1994, 1997) shows that for more complex estimations, such as multiple regressions, these conclusions hold only to a limited extent. For instance, when there is more than one independent variable, the direction of bias is difficult to determine (Levi, 1973). In addition, Dagenais and Dagenais (1997) argue that measurement errors adversely affect the size of Type I errors of standard econometric tests. Bound *et al.* (2001) provide a comprehensive survey of measurement errors in survey data. A critical point they emphasize is that traditional methods that are used to alleviate measurement

[1] By "near-integrated" we mean that the process has a root close to but not on the unit root circle.

errors assume that the measurement error is not correlated with the true value of the regressor, which they found often untrue. Violation of this assumption can produce estimates that are even more distorted than OLS that ignore measurement errors. Chen *et al.* (2005) propose a solution to this problem using auxiliary data that contain information about the conditional distribution of the true regressors given the regressors that contain measurement errors.

In linear models, measurement errors are typically handled with the IV approach, but one needs to find instruments correlated with the true value of the regressor and at the same time not correlated with the measurement error. It is worth noting that a good candidate for an instrument thus could be the lagged value of the regressor because it is usually correlated with the original regressor, but not contemporaneously correlated with the measurement error (assuming the measurement error is not autocorrelated). Unfortunately, as noted by Amemiya (1985), Cragg (1994) and Schennach (2004), standard IV approach breaks down when the specification is non-linear. In this setting the measurement error cannot be considered as an additively separable disturbance and finding an adequate instrument becomes extremely difficult.

To cope with measurement errors in non-linear models various approaches have been devised. Hausman *et al.* (1991) generalized the IV method to polynomial functions establishing identification and providing a consistent estimator. Wang and Hsiao (2003) and Newey (2001) used distributional assumptions on the measurement error to obtain a general framework for non-linear models not limited to polynomials, when no auxiliary data are present. Nevertheless, Schennach (2004) showed that although the approaches by Hausman *et al.* (1991), Wang and Hsiao (2003) and Newey (2001) do provide the identification of the non-linear errors-in-variables model using instruments, this holds only in a limited number of specific cases. In contrast to this strand of research, Horowitz and Manski (1995) relaxed the assumption of classical errors and conceptualized a model of measurement error in which they assume that the observed sample is contaminated or corrupted.[2] Other papers allowing for non-classical measurement errors involve the existence of true validation data, i.e., subsamples of the primary data (e.g., Sepanski and Carroll, 1993 and Lee and Sepanski, 1985). The methods that use validation data are essentially able to correct biases and obtain consistent estimates from primary data without any distributional assumptions. Phillips (1987) presented asymptotic results for unit root and near-unit root processes using a unified framework to explain the properties of regressions involving borderline-stationary variables. His results suggest very similar finite-sample behavior in unit root and near-integrated processes. However, failure to properly identify variables that have a unit root can result in a "near-unit root bias" - standard estimators such as OLS are significantly downwardly biased in finite samples (Cashin *et al.*, 2004). In addition, Elliott and Stock (1994) provide Monte Carlo evidence which indicates substantial size distortions of the one-sided t-test on the slope coefficient when the regressor is highly persistent.

This paper introduces a novel approach to tackling the problem of errors-in-variables in a

[2] For more information on this approach see Molinari (2008).

regression with highly persistent regressors.[3] By employing extensive Monte Carlo simulations, we demonstrate that the IV approach dominates the OLS approach which consistently produces biased estimates. However, the IV approach becomes unreliable when the persistence of regressors decreases and the serial correlation in the measurement error increases. The Monte Carlo results further show that the wavelet approach dominates both the OLS and IV estimation methods and is particularly effective for high levels of serial correlation in the measurement error.

The reminder of this paper is organized as follows. Section 2 briefly explains the problem of measurement errors. Section 3 presents the wavelet methodology and its applications. Section 4 reports the results of our extensive Monte Carlo experiments and compares the OLS, IV and wavelet approaches. Section 5 concludes.

2. The Problem

Consider a linear regression model

$$y_t^* = \alpha + \beta x_t^* + \epsilon_t^*, \quad \epsilon_t^* \sim iid(0, \sigma^2_{\epsilon^*}) \tag{1}$$

where y_t^* and x_t^* are unobserved persistent processes, ϵ_t is identically and independently distributed (iid) with variance $\sigma^2_{\epsilon^*}$. The observables are

$$x_t = x_t^* + v_{1t} \qquad \text{and} \qquad y_t = y_t^* + v_{2t} \tag{2}$$

where $v_{1t} \sim iid(0, \sigma^2_{v1})$ and $v_{2t} \sim iid(0, \sigma^2_{v2})$ are measurement errors. Substituting $y_t^* = y_t - v_{1t}$ and $x_t^* = x_t - v_{2t}$ from Equation (2) into Equation (1) yields

$$\begin{aligned} y_t &= \alpha + \beta(x_t - v_{1t}) + \epsilon_t^* + v_{2t} \\ &= \alpha + \beta x_t + \epsilon_t^* + v_{2t} - \beta v_{1t} \\ &= \alpha + \beta x_t + \epsilon_t \end{aligned} \tag{3}$$

where $\epsilon_t \equiv \epsilon_t^* + v_{2t} - \beta v_{1t}$ where $\epsilon_t \sim iid(0, \sigma^2_{\epsilon^*} + \sigma^2_{v2} + \beta^2\sigma^2_{v1})$. The presence of measurement error in y_t and x_t lead to an increase in the variance of the error term. In addition, since ϵ_t depends on v_{1t}, ϵ_t and x_t are correlated as long as β is not zero. It is easy to show that $\mathrm{E}(\epsilon_t|x_t) = -\beta v_{1t}$ and $\mathrm{Cov}(x_t, \epsilon_t) = -\beta\sigma^2_{v1}$. Since this covariance does not depend on the sample size, it does not vanish

[3] A natural usage area of our approach is financial time series. Only a select few papers in modern financial economics consider errors-in-variables problems. One of the first contributions that discussed measurement errors is Fama and MacBeth (1973). This paper used a portfolio grouping technique in a two factor portfolio model to minimize the impact of errors-in-variables biases. Other authors in the same vein include Shanken (1992), Ferson and Locke (1998), Pastor and Stambaugh (1999), and, more recently, Carmichael and Coën (2008).

asymptotically, and OLS estimator is downward biased and inconsistent.[4] A conventional way to deal with the inconsistency of the OLS estimator is to use IV estimation.

A new novel solution to the bias problem is to filter out the measurement noise and use the filtered regressand and regressors in OLS to obtain unbiased and consistent estimators. We apply the wavelet method to both y_t and x_t, and regress the scaling coefficients of $\tilde{y}^*_t$ onto $\tilde{x}^*_t$. Namely, we run the following regression

$$\tilde{y}^*_t = \alpha + \beta \tilde{x}^*_t + \tilde{\epsilon}_t, \quad \tilde{\epsilon}_t \sim iid(0, \sigma^2_{\tilde{\epsilon}}) \tag{4}$$

instead of Equation (3). Since noise terms are left behind in the wavelet coefficients, the scaling coefficients will provide unbiased and consistent coefficient estimators without any instruments added to the regression.

3. Wavelet Framework

Wavelet methods are rather newer ways of analyzing time series and can be seen as a natural extension of the Fourier analysis. The formal subject matter, in terms of their formal mathematical and statistical foundations go back only to the 1980s. In recent years, there have been several unique applications of wavelet methods to financial and econometric problems. Early applications of wavelets in economics and finance are by Ramsey and his coauthors (see Ramsey *et al.* (1995), Ramsey and Zhang (1997), Ramsey (1999), Ramsey (2002) for a review and references) who explore the use of DWT in decomposing various economic and financial data. Davidson *et al.* (1998) investigated U.S. commodity prices via wavelets. Gençay *et al.* (2003, 2005) propose a wavelet approach for estimating the systematic risk or the beta of an asset in a capital asset pricing model. The proposed method is based on a wavelet multiscaling approach where the wavelet variance of the market return and the wavelet covariance between the market return and a portfolio are calculated to obtain an estimate of the portfolio's systematic risk (beta) at each scale. In time series econometrics, one example of the successful application of wavelets is in the context of long memory processes where a number of estimation methods have been developed. These include wavelet-based OLS, the approximate wavelet-based maximum likelihood approach, and wavelet-based Bayesian approach. Fan (2003) and Fan and Whitcher (2003) provide an extensive list of references. The success of these methods relies on the so called 'approximate decorrelation' property of the DWT of a possibly nonstationary long memory process.[5] Fan and Whitcher (2003) propose overcoming the problem of spurious regression between fractionally differenced processes by applying the DWT to both processes and then estimating the regression in the wavelet domain. Other examples of applications of wavelets in econometrics include wavelet-based spectral density estimators and their

[4]Davidson and MacKinnon (2004)(pages 312-314) provide an excellent treatment of the errors-in-variables and instrumental variables approaches. Our notation closely follows Davidson and MacKinnon (2004).

[5]See Fan (2003) for a rigorous proof of this result for a nonstationary fractionally differenced process.

applications in testing for serial correlation/conditional heteroscedasticity, see e.g., Hong (2000), Hong and Lee (2001), Lee and Hong (2001), Duchesne (2006a), Duchesne (2006b), and Hong and Kao (2004).

A wavelet is a small wave which grows and decays in a limited time period.[6] To formalize the notion of a wavelet, let $\psi(.)$ be a real valued function such that its integral zero,

$$\int_{-\infty}^{\infty} \psi(t)\, dt = 0, \tag{5}$$

and its square integrates to unity,

$$\int_{-\infty}^{\infty} \psi(t)^2\, dt = 1. \tag{6}$$

Wavelets are, in particular, useful for the study of how weighted averages vary from one averaging period to the next. Let $x(t)$ be real-valued and consider the integral

$$\bar{x}(s, e) \equiv \frac{1}{s - e} \int_{s}^{e} x(u)\, du \tag{7}$$

where we assume that $e > s$. $\bar{x}(s, e)$ is the average value of $x(.)$ over the interval $[s, e]$. Instead of treating an average value $\bar{x}(s, e)$ as a function of end points of the interval $[s, e]$, it can be considered as a function of the length of the interval,

$$\lambda \equiv s - e$$

while centering the interval at

$$t = (s + e)/2.$$

λ is referred to as the scale associated with the average, and using λ and t, the average can be redefined such that

$$a(\lambda, t) \equiv \bar{x}(t - \frac{\lambda}{2}, t + \frac{\lambda}{2}) = \frac{1}{\lambda} \int_{t-\frac{\lambda}{2}}^{t+\frac{\lambda}{2}} x(u)\, du$$

where $a(\lambda, t)$ is the average value of $x(.)$ over a scale of λ centered at time t. The change in $a(\lambda, t)$ from one time period to another is measured by

[6] The contrasting notion is a big wave such as the sine function which keeps oscillating indefinitely.

$$w(\lambda,t) \equiv a(\lambda, t+\frac{\lambda}{2}) - a(\lambda, t-\frac{\lambda}{2}) = \frac{1}{\lambda}\int_{t}^{t+\lambda} x(u)\,du - \frac{1}{\lambda}\int_{t-\lambda}^{t} x(u)\,du. \tag{8}$$

Equation 8 measures how much the average changes between two adjacent nonoverlapping time intervals, from $t-\lambda$ to $t+\lambda$, each with a length of λ. Because the two integrals in Equation 8 involve adjacent nonoverlapping intervals, they can be combined into a single integral over the real axis to obtain

$$w(\lambda,t) = \int_{-\infty}^{\infty} \tilde{\psi}(t)x(u)\,du \tag{9}$$

where

$$\tilde{\psi}(t) = \begin{cases} -1/\lambda, & t-\lambda < u < t, \\ 1/\lambda, & t < u < t+\lambda, \\ 0, & \text{otherwise.} \end{cases}$$

$w(\lambda,t)$'s are the wavelet coefficients and they are essentially the changes in averages across adjacent (weighted) averages.

3.1. Discrete wavelet transformation

In principle, wavelet analysis can be carried out in all arbitrary time scales. This may not be necessary if only key features of the data are in question, and if so, discrete wavelet transformation (DWT) is an efficient and parsimonious route as compared to the continuous wavelet transformation (CWT). The DWT is a subsampling of $w(\lambda,t)$ with only dyadic scales, i.e., λ is of the form $2^{j-1},\ j=1,2,3,\ldots$ and, within a given dyadic scale 2^{j-1}, t's are separated by multiples of 2^j.

Let $\mathbf{x}$ be a dyadic length vector ($N=2^J$) of observations. The length N vector of discrete wavelet coefficients $\mathbf{w}$ is obtained by

$$\mathbf{w} = \mathcal{W}\mathbf{x},$$

where $\mathcal{W}$ is an $N\times N$ real-valued orthonormal matrix defining the DWT which satisfies $\mathcal{W}^T\mathcal{W} = I_N$ ($n\times n$ identity matrix).[7] The nth wavelet coefficient w_n is associated with a particular scale and with a particular set of times. The vector of wavelet coefficients may be organized into $J+1$ vectors,

[7]Since DWT is an orthonormal transform, orthonormality implies that $\mathbf{x} = \mathcal{W}^T\mathbf{w}$ and $||w||^2 = ||x||^2$.

$$\mathbf{w} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_J, \mathbf{v}_J]^T,$$

where $\mathbf{w}_j$ is a length $N/2^j$ vector of wavelet coefficients associated with changes on a scale of length $\lambda_j = 2^{j-1}$ and $\mathbf{v}_J$ is a length $N/2^J$ vector of scaling coefficients associated with averages on a scale of length $2^J = 2\lambda_J$.

Using the DWT, we may formulate an additive decomposition of $\mathbf{x}$ by reconstructing the wavelet coefficients at each scale independently. Let $\mathbf{d}_j = \mathcal{W}_j^T \mathbf{w}_j$ define the jth level *wavelet detail* associated with changes in $\mathbf{x}$ at the scale λ_j (for $j = 1, \ldots, J$). The wavelet coefficients $\mathbf{w}_j = \mathcal{W}_j \mathbf{x}$ represent the portion of the wavelet analysis (decomposition) attributable to scale λ_j, while $\mathcal{W}_j^T \mathbf{w}_j$ is the portion of the wavelet synthesis (reconstruction) attributable to scale λ_j. For a length $N = 2^J$ vector of observations, the vector $\mathbf{d}_{J+1}$ is equal to the sample mean of the observations.

A multiresolution analysis (MRA) may now be defined via

$$x_t = \sum_{j=1}^{J+1} \mathbf{d}_{j,t} \quad t = 1, \ldots, N. \tag{10}$$

That is, each observation x_t is a linear combination of wavelet detail coefficients at time t. Let $\mathbf{s}_j = \sum_{k=j+1}^{J+1} \mathbf{d}_k$ define the jth level *wavelet smooth*. Whereas the wavelet detail $\mathbf{d}_j$ is associated with variations at a particular scale, $\mathbf{s}_j$ is a cumulative sum of these variations and will be smoother and smoother as j increases. In fact, $\mathbf{x} - \mathbf{s}_j = \sum_{k=1}^{j} \mathbf{d}_k$ so that only lower-scale details (high-frequency features) from the original series remain. The jth level *wavelet rough* characterizes the remaining lower-scale details through

$$\mathbf{r}_j = \sum_{k=1}^{j} \mathbf{d}_k, \qquad 1 \leq j \leq J+1.$$

The wavelet rough $\mathbf{r}_j$ is what remains after removing the wavelet smooth from the vector of observations. A vector of observations may thus be decomposed through a wavelet smooth and rough via

$$\mathbf{x} = \mathbf{s}_j + \mathbf{r}_j,$$

for all j.

A variation of the DWT is called the maximum overlap DWT (MODWT). Similar to the DWT, the MODWT is a subsampling at dyadic scales, but in contrast to the DWT, the analysis involves all times t rather than the multiples of 2^j. Retainment of all possible times eliminates alignment effects of DWT and leads to more efficient time series representation at multiple time scales. In this paper, we use the MODWT in our disentangling of the intraday from the interday dynamics.

3.2. Analysis of variance

The orthonormality of the matrix $\mathcal{W}$ implies that the DWT is a variance preserving transformation where

$$\|\mathbf{w}\|^2 = \sum_{j=1}^{J} \sum_{t=0}^{N/2^j-1} w_{j,t}^2 + v_{J,0}^2 = \sum_{t=0}^{N-1} x_t^2 = \|\mathbf{x}\|^2 \,.$$

This can be easily proven through basic matrix manipulation via

$$\begin{aligned} \|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x} &= (\mathcal{W}\mathbf{w})^T\mathcal{W}\mathbf{w} \\ &= \mathbf{w}^T\mathcal{W}^T\mathcal{W}\mathbf{w} = \mathbf{w}^T\mathbf{w} = \|\mathbf{w}\|^2 \,. \end{aligned}$$

Given the structure of the wavelet coefficients, $\|\mathbf{x}\|^2$ is decomposed on a scale-by-scale basis via

$$\|\mathbf{x}\|^2 = \sum_{j=1}^{J} \|\mathbf{w}_j\|^2 + \|\mathbf{v}_J\|^2 \,, \tag{11}$$

where $\|\mathbf{w}_j\|^2$ is the sum of squared variation of $\mathbf{x}$ due to changes at scale λ_j and $\|\mathbf{v}_J\|^2$ is the information due to changes at scales λ_J and higher. An alternative decomposition of $\|\mathbf{x}\|^2$ to Equation 11 is

$$\|\mathbf{x}\|^2 = \sum_{j=1}^{J} \|\mathbf{d}_j\|^2 + \|\mathbf{s}_J\|^2$$

which decomposes the variations in $\mathbf{x}$ across the variations in details and the smooth.

Percival and Mofjeld (1997) proved that the MODWT is an energy (variance) preserving transform such that the variance of the original time series is perfectly captured by the variance of the coefficients from the MODWT. Specifically, the total variance of a time series can be partitioned using the MODWT wavelet and scaling coefficient vectors by

$$\|\mathbf{x}\|^2 = \sum_{j=1}^{J} \|\widetilde{\mathbf{w}}_j\|^2 + \|\widetilde{\mathbf{v}}_J\|^2 \tag{12}$$

where $\widetilde{\mathbf{w}}_j$ is a length $N/2^j$ vector of MODWT wavelet coefficients associated with changes on a scale of length $\lambda_j = 2^{j-1}$ and $\widetilde{\mathbf{v}}_J$ is a length $N/2^J$ vector of MODWT scaling coefficients associated with averages on a scale of length $2^J = 2\lambda_J$.[8] This will allow us to construct MODWT versions of the wavelet variance.

[8] More information on the MODWT transformation can be found in Percival and T. (2000).

4. Monte Carlo Simulations

Consider a linear regression model

$$y_t^* = \alpha + \beta x_t^* + \epsilon_t, \quad \epsilon_t \sim iid(0, \sigma_\epsilon^2) \tag{13}$$

where y_t^* and $x_t^* = \gamma x_{t-1}^* + w_t$ are unobserved, $w_t \sim iid(0, \sigma_w^2)$ and $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$. The observables are $y_t = y_t^* + v_{1t}$ and $x_t = x_t^* + v_{2t}$ where $v_{1t} \sim iid(0, \sigma_{v1}^2)$ and $v_{2t} \sim iid(0, \sigma_{v2}^2)$ are measurement errors. The simulation set up is such that x_t^* is a persistent process ($\gamma \in \{0.99, 0.97, 0.95\}$).

Figure 1 illustrates the OLS, IV and MODWT (level of decomposition = 4, least asymmetric - LA(8) wavelet) performance for this model where $\alpha = 1$, $\beta = 1$, $\gamma = 0.99$, and all error terms are $\mathcal{N}(0, 1)$.[9] The sample size is $T = 2000$ observations, the number of replications is $B = 1000$, boundary-free MODWT scaling coefficients are used in the wavelet-OLS regression and the standard errors of the estimated OLS regressions are calculated with the Newey-West heteroscedasticity and autocorrelation correction (Newey and West, 1987). The IV estimation is performed with the lagged value of the regressor used as an instrument. The IV and MODWT approaches exhibit similar distributional properties. They are both superior to the OLS distributions that are clearly distorted with the slope coefficient estimates biased towards zero. Decreasing the persistence (γ) from 0.99 to 0.95 exacerbates these biases for the OLS estimator (see also Table 1, explained below).

Next, we introduce serial correlation in the measurement error for the regressor, i.e., $v_{2t} = \rho v_{2(t-1)} + e_{2t}$ where $e_{2t} \sim iid(0, \sigma_{e2}^2)$. As before, we observe the performance of the OLS, IV and MODWT (level 4, LA(8)) models where the true values $\alpha = 1$ and $\beta = 1$, while all error terms are $\mathcal{N}(0, 1)$. Table 1 presents the results of the Monte Carlo simulations ($T = 2000$ and $B = 1000$) for a variety of model specifications: $\gamma \in \{0.99, 0.97, 0.95\}$, $\rho \in \{0, 0.2, 0.6\}$. Our framework thus sheds light on the effects of decreasing the persistence of the regressor (γ) and of increasing the serial correlation in the measurement error (ρ). We report the mean of the simulated distribution of $\hat{\beta}$ (Mean) along with its standard deviation (SD), median (Med) and mean absolute deviation (MAD). The results show that the OLS and IV estimators in general do not work well in the presence of the serial correlation in measurement error. The poor performance is an increasing function of the serial correlation and a decreasing function of the persistence level. For instance, when $\gamma = 0.95$ and $\rho = 0.6$, Mean$(\hat{\beta})_{MODWT} = 0.9496$ while Mean$(\hat{\beta})_{IV} = 0.9093$. It is noteworthy that the OLS estimator is the most biased of the three approaches with the corresponding Mean$(\hat{\beta})_{OLS} = 0.7957$. Figure 2 confirms these findings and reveals that the MODWT method is the only one that to a certain extent preserves the shape of the distributions. Not surprisingly, the OLS estimator distorts the t-statistic of $\hat{\beta}$ and pushes it to the left, being just slightly tangent to the true t-distribution with its right tail.

[9] The results in this section are in general robust to the wavelet family choice (extremal phase, least asymmetric, best localized and coiflets) and the level of decomposition up to Level 7.

Finally, we allow for serially correlated measurement errors in both the regressor and the regressand: $v_{1t} = \rho v_{1(t-1)} + e_{1t}$, $v_{2t} = \rho v_{2(t-1)} + e_{2t}$ where $e_{1t} \sim iid(0, \sigma_{e1}^2)$ and $e_{2t} \sim iid(0, \sigma_{e2}^2)$. Table 2 shows findings similar to the ones in Table 1: serial correlation affects the IV and OLS estimators more than the MODWT model, especially when the persistence (γ) decreases. In addition, we observe higher standard deviation and MAD for all methods.

Next, to illustrate the robustness of the MODWT method, we present the results for the LA(8) filter and the level of decomposition is set to 6 (Tables 3 and 4). We observe that increasing the level of decomposition improves upon the results in Tables 1 and 2, but the evidence in principal confirms the dominance of the MODWT approach. In Tables 5-8, we also report the results for the Haar wavelet with the levels of decomposition 4 and 6. These results are essentially the same as for the LA(8) filter. It is worth noting that the higher level of decomposition again increases the precision of the MODWT (Haar) estimates.[10]

5. Conclusions

This paper introduces a new approach to tackling the problem of errors-in-variables in a regression with highly persistent regressors. By employing extensive Monte Carlo simulations, we demonstrate that the IV estimation dominates the OLS approach which consistently produces biased estimates. However, the IV approach becomes unreliable when the persistence of regressors decreases and the serial correlation in the measurement error increases. The Monte Carlo results further show that the MODWT approach dominates both the OLS and IV estimation methods and is particularly effective for high levels of serial correlation in the measurement error. The MODWT method requires no instruments and yields unbiased and consistent parameter estimates in a single or multiple regression models. These findings are robust to various specifications of the wavelet filters.

The current version of this paper constitutes an initial exploration of the application of wavelets to the problem of errors-in-variables; much remains to be done. In future work, we would like to adapt our methodology to the case of less persistent regressors. In addition, there are more practical aspects to consider, such as testing the MODWT approach on real data. For example, the market microstructure literature routinely ignores the fact that order flows are subject to measurement errors. Therefore, our approach may shed more light and complement existing literature in this area. The application of wavelets to this particular problem as well as problems that might arise when the model specification is non-linear, remains to be explored.

[10]In unreported results, we also compared the performance of the OLS, IV, and MODWT methods for regressions with two persistent and independent regressors contaminated with measurement errors. We did not find any major differences between these results and the results for a single regressor.

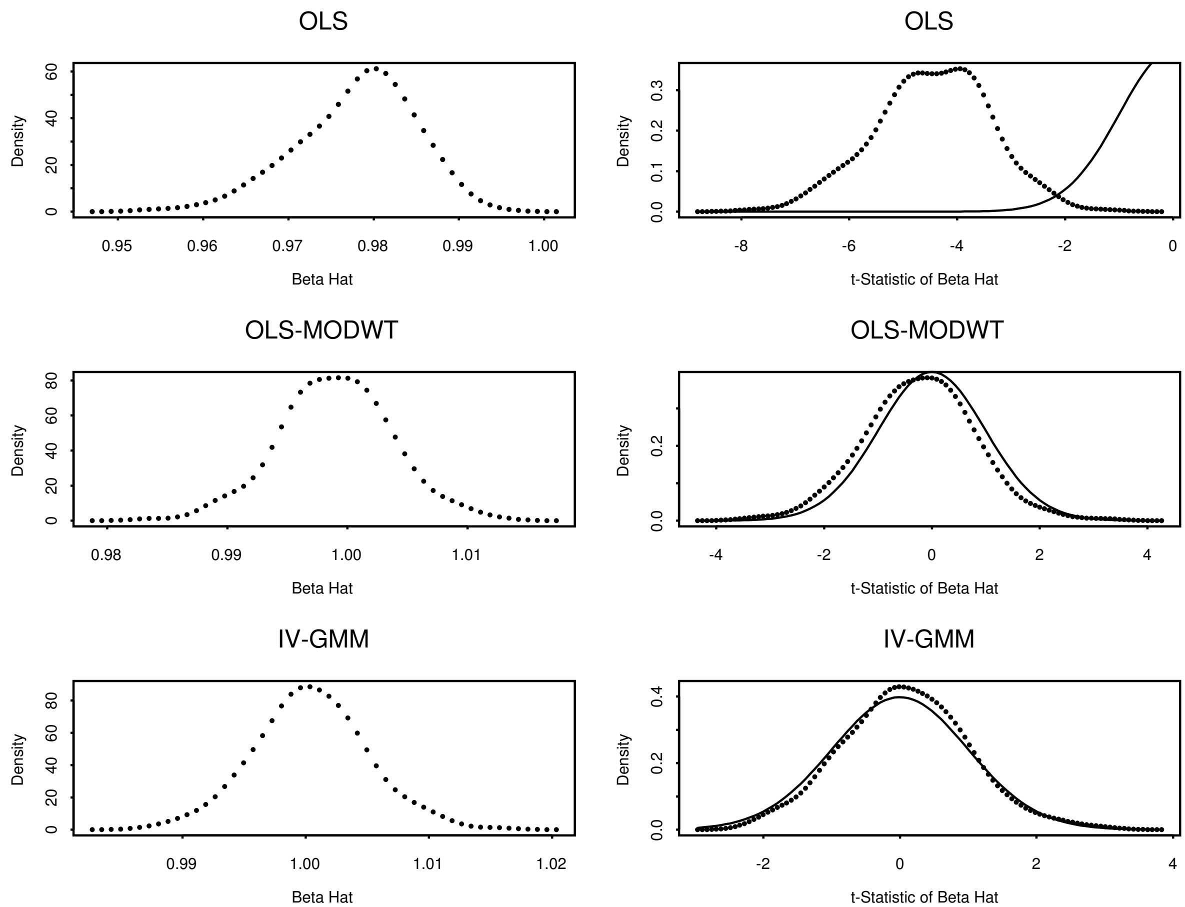

Figure 1: Distributions of $\hat{\beta}$ and $t_{\hat{\beta}}$ from OLS, IV and Wavelet (MODWT) regressions when the regressor is measured with error (γ=0.99). The estimated distributions are marked by dotted lines. The true $\beta = 1$. The OLS estimator is biased and centered incorrectly. The MODWT and IV estimators are unbiased and centered correctly. The corresponding t-distribution is marked by a solid line.

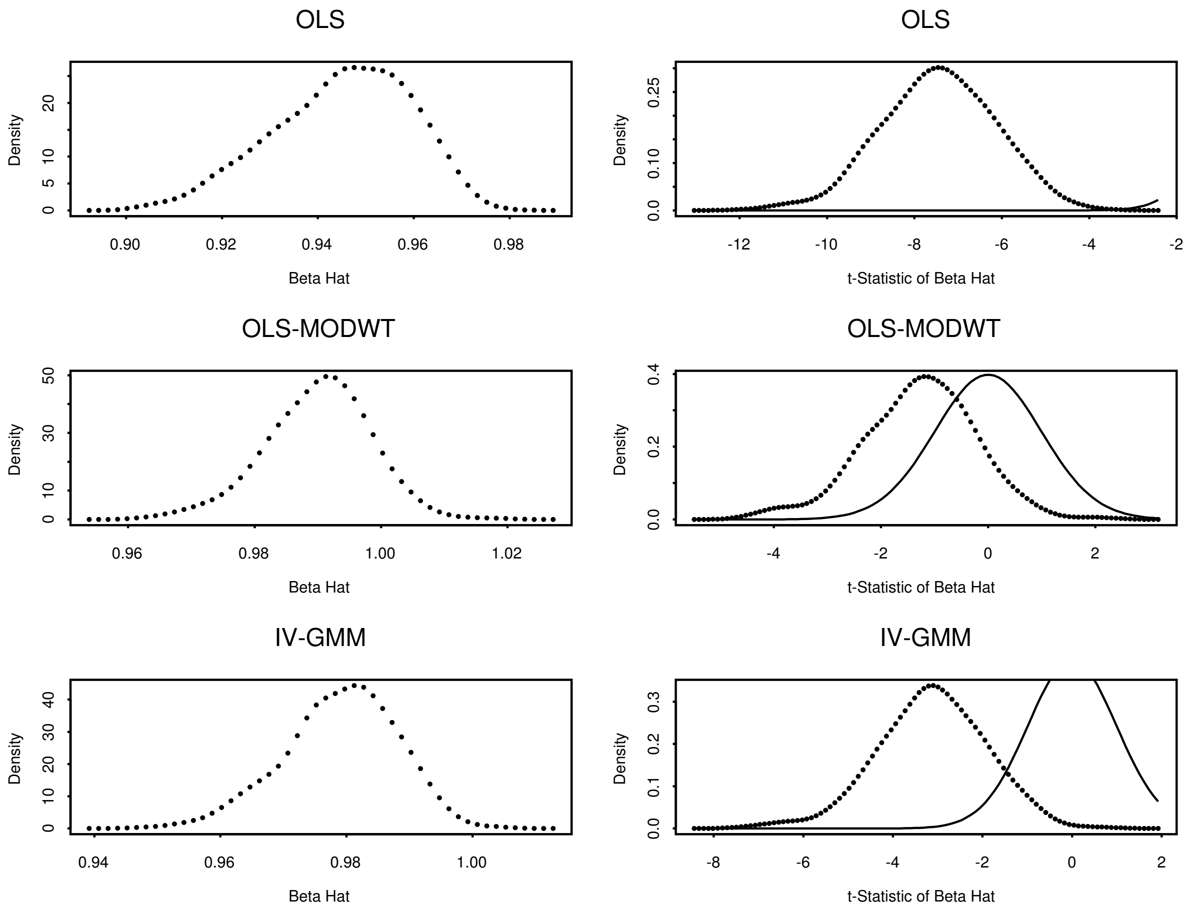

Figure 2: Distributions of $\hat{\beta}$ and $t_{\hat{\beta}}$ from OLS, IV and Wavelet (MODWT) regressions when the regressor is measured with a serially correlated error ($\gamma = 0.99$, $\rho = 0.6$). The estimated distributions are marked by dotted lines. The true $\beta = 1$. The OLS and IV estimators are biased and centered incorrectly. The MODWT estimator is more robust to serially correlated measurement errors. The corresponding t-distribution is marked by a solid line.

γ	ρ	MODWT				IV				OLS			
		Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)
0.99	0	0.998	0.004	0.998	0.003	0.999	0.004	0.999	0.003	0.976	0.007	0.977	0.005
	0.2	0.996	0.006	0.997	0.004	0.995	0.006	0.996	0.004	0.955	0.013	0.956	0.010
	0.6	0.989	0.009	0.989	0.007	0.978	0.009	0.979	0.007	0.943	0.016	0.946	0.012
0.97	0	0.995	0.008	0.996	0.006	0.999	0.008	1.000	0.006	0.940	0.011	0.941	0.009
	0.2	0.988	0.009	0.988	0.007	0.986	0.010	0.986	0.008	0.886	0.017	0.887	0.013
	0.6	0.970	0.015	0.970	0.012	0.943	0.016	0.944	0.013	0.863	0.021	0.863	0.017
0.95	0	0.992	0.010	0.991	0.008	0.999	0.010	0.999	0.008	0.908	0.013	0.909	0.011
	0.2	0.982	0.013	0.982	0.010	0.979	0.013	0.979	0.010	0.831	0.018	0.831	0.014
	0.6	0.949	0.019	0.950	0.015	0.909	0.018	0.910	0.015	0.795	0.022	0.795	0.017

Table 1: Comparison of OLS, IV and MODWT (LA-8, level of decomposition 4) regressions when the regressor is measured with error.
This table reports the results of the Monte Carlo simulations for OLS, IV and MODWT regressions. The regressor is measured with error. We list the mean value of $\hat{\beta}$ (Mean), the standard deviation of $\hat{\beta}$ (SD), the median of $\hat{\beta}$ (Med) and the mean-absolute deviation of $\hat{\beta}$ (MAD) after $B = 1000$ replications for $T = 2000$. $\alpha = 1$ and $\beta = 1$.

γ	ρ	MODWT				IV				OLS			
		Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)
0.99	0	0.998	0.005	0.998	0.004	0.999	0.005	0.999	0.004	0.977	0.007	0.978	0.006
	0.2	0.995	0.007	0.995	0.006	0.994	0.007	0.995	0.005	0.954	0.013	0.954	0.010
	0.6	0.989	0.012	0.989	0.009	0.978	0.013	0.979	0.010	0.944	0.017	0.945	0.014
0.97	0	0.995	0.009	0.995	0.007	0.999	0.009	0.999	0.007	0.941	0.011	0.942	0.009
	0.2	0.989	0.013	0.989	0.010	0.986	0.013	0.986	0.011	0.887	0.018	0.887	0.014
	0.6	0.969	0.020	0.969	0.016	0.942	0.020	0.943	0.016	0.862	0.025	0.863	0.019
0.95	0	0.992	0.011	0.992	0.009	0.999	0.011	0.999	0.009	0.908	0.014	0.908	0.011
	0.2	0.980	0.018	0.981	0.014	0.977	0.017	0.978	0.013	0.829	0.023	0.831	0.018
	0.6	0.951	0.026	0.951	0.021	0.910	0.026	0.912	0.020	0.795	0.027	0.796	0.021

Table 2: Comparison of OLS, IV and MODWT (LA-8, level of decomposition 4) regressions when the regressor and the regressand are measured with error.
This table reports the results of the Monte Carlo simulations for OLS, IV and MODWT regressions. The regressor and the regressand are measured with error. We list the mean value of $\hat{\beta}$ (Mean), the standard deviation of $\hat{\beta}$ (SD), the median of $\hat{\beta}$ (Med) and the mean-absolute deviation of $\hat{\beta}$ (MAD) after $B = 1000$ replications for $T = 2000$. $\alpha = 1$ and $\beta = 1$.

γ	ρ	MODWT				IV				OLS			
		Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)
0.99	0	0.999	0.005	0.999	0.004	1.000	0.004	1.000	0.003	0.976	0.007	0.977	0.006
	0.2	0.998	0.006	0.998	0.005	0.994	0.005	0.994	0.004	0.955	0.012	0.956	0.009
	0.6	0.996	0.010	0.996	0.008	0.979	0.010	0.979	0.008	0.944	0.016	0.945	0.012
0.97	0	0.997	0.009	0.996	0.007	0.999	0.007	0.999	0.006	0.940	0.010	0.940	0.008
	0.2	0.995	0.013	0.995	0.010	0.986	0.009	0.987	0.007	0.887	0.017	0.888	0.013
	0.6	0.988	0.019	0.988	0.015	0.943	0.015	0.944	0.012	0.862	0.020	0.862	0.016
0.95	0	0.997	0.014	0.997	0.011	1.000	0.010	1.000	0.008	0.908	0.012	0.908	0.010
	0.2	0.991	0.018	0.991	0.014	0.978	0.013	0.978	0.010	0.830	0.019	0.830	0.015
	0.6	0.975	0.028	0.975	0.022	0.908	0.020	0.909	0.016	0.794	0.024	0.794	0.019

Table 3: Comparison of OLS, IV and MODWT (LA-8, level of decomposition 6) regressions when the regressor is measured with error.

This table reports the results of the Monte Carlo simulations for OLS, IV and MODWT regressions. The regressor is measured with error. We list the mean value of $\hat{\beta}$ (Mean), the standard deviation of $\hat{\beta}$ (SD), the median of $\hat{\beta}$ (Med) and the mean-absolute deviation of $\hat{\beta}$ (MAD) after $B = 1000$ replications for $T = 2000$. $\alpha = 1$ and $\beta = 1$.

γ	ρ	MODWT				IV				OLS			
		Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)
0.99	0	0.999	0.006	0.999	0.004	0.999	0.005	0.999	0.003	0.977	0.007	0.978	0.005
	0.2	0.999	0.008	0.999	0.006	0.994	0.007	0.995	0.006	0.954	0.013	0.956	0.011
	0.6	0.997	0.013	0.998	0.010	0.978	0.012	0.978	0.010	0.943	0.017	0.943	0.014
0.97	0	0.999	0.011	0.999	0.008	1.000	0.008	1.000	0.007	0.942	0.011	0.942	0.009
	0.2	0.996	0.018	0.996	0.014	0.986	0.013	0.986	0.010	0.886	0.019	0.887	0.015
	0.6	0.988	0.028	0.987	0.022	0.943	0.020	0.944	0.016	0.862	0.024	0.865	0.019
0.95	0	0.995	0.017	0.996	0.014	0.999	0.012	0.999	0.010	0.908	0.014	0.908	0.011
	0.2	0.992	0.023	0.990	0.018	0.979	0.017	0.979	0.013	0.830	0.021	0.831	0.017
	0.6	0.973	0.042	0.973	0.033	0.908	0.027	0.910	0.022	0.795	0.028	0.794	0.022

Table 4: Comparison of OLS, IV and MODWT (LA-8, level of decomposition 6) regressions when the regressor and the regressand are measured with error.
This table reports the results of the Monte Carlo simulations for OLS, IV and MODWT regressions. The regressor and the regressand are measured with error. We list the mean value of $\hat{\beta}$ (Mean), the standard deviation of $\hat{\beta}$ (SD), the median of $\hat{\beta}$ (Med) and the mean-absolute deviation of $\hat{\beta}$ (MAD) after $B = 1000$ replications for $T = 2000$. $\alpha = 1$ and $\beta = 1$.

γ	ρ	MODWT				IV				OLS			
		Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)
0.99	0	0.999	0.004	0.999	0.003	1.000	0.004	1.000	0.003	0.977	0.007	0.978	0.005
	0.2	0.998	0.006	0.998	0.004	0.994	0.005	0.995	0.004	0.955	0.012	0.956	0.010
	0.6	0.996	0.010	0.996	0.007	0.978	0.010	0.979	0.008	0.944	0.017	0.946	0.013
0.97	0	0.998	0.009	0.997	0.007	1.000	0.007	1.000	0.006	0.941	0.011	0.942	0.009
	0.2	0.995	0.013	0.995	0.010	0.987	0.010	0.987	0.008	0.887	0.017	0.887	0.014
	0.6	0.989	0.018	0.988	0.014	0.944	0.016	0.944	0.013	0.863	0.022	0.864	0.018
0.95	0	0.996	0.014	0.995	0.011	0.999	0.010	0.999	0.008	0.909	0.013	0.909	0.010
	0.2	0.991	0.018	0.991	0.014	0.978	0.013	0.979	0.010	0.829	0.019	0.831	0.015
	0.6	0.976	0.027	0.975	0.022	0.910	0.019	0.910	0.015	0.796	0.023	0.797	0.018

Table 5: Comparison of OLS, IV and MODWT (Haar, level of decomposition 6) regressions when the regressor is measured with error.

This table reports the results of the Monte Carlo simulations for OLS, IV and MODWT regressions. The regressor is measured with error. We list the mean value of $\hat{\beta}$ (Mean), the standard deviation of $\hat{\beta}$ (SD), the median of $\hat{\beta}$ (Med) and the mean-absolute deviation of $\hat{\beta}$ (MAD) after $B = 1000$ replications for $T = 2000$. $\alpha = 1$ and $\beta = 1$.

γ	ρ	MODWT				IV				OLS			
		Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)
0.99	0	0.999	0.005	0.999	0.004	1.000	0.005	1.000	0.004	0.977	0.007	0.978	0.006
	0.2	0.998	0.008	0.998	0.006	0.995	0.007	0.995	0.005	0.955	0.013	0.957	0.010
	0.6	0.996	0.012	0.997	0.009	0.978	0.012	0.978	0.010	0.943	0.017	0.944	0.013
0.97	0	0.999	0.012	1.000	0.009	1.000	0.009	1.000	0.007	0.942	0.011	0.943	0.009
	0.2	0.994	0.016	0.995	0.012	0.985	0.013	0.985	0.010	0.884	0.019	0.885	0.015
	0.6	0.988	0.027	0.988	0.021	0.943	0.020	0.943	0.016	0.863	0.024	0.863	0.019
0.95	0	0.996	0.017	0.997	0.013	0.999	0.012	0.999	0.009	0.908	0.014	0.909	0.012
	0.2	0.991	0.024	0.991	0.019	0.978	0.017	0.978	0.013	0.829	0.022	0.830	0.018
	0.6	0.974	0.039	0.975	0.031	0.911	0.024	0.911	0.019	0.796	0.026	0.797	0.021

Table 6: Comparison of OLS, IV and MODWT (Haar, level of decomposition 6) regressions when the regressor and the regressand are measured with error.
This table reports the results of the Monte Carlo simulations for OLS, IV and MODWT regressions. The regressor and the regressand are measured with error. We list the mean value of $\hat{\beta}$ (Mean), the standard deviation of $\hat{\beta}$ (SD), the median of $\hat{\beta}$ (Med) and the mean-absolute deviation of $\hat{\beta}$ (MAD) after $B = 1000$ replications for $T = 2000$. $\alpha = 1$ and $\beta = 1$.

γ	ρ	MODWT				IV				OLS			
		Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)
0.99	0	0.998	0.004	0.998	0.003	1.000	0.004	1.000	0.003	0.976	0.007	0.978	0.006
	0.2	0.996	0.005	0.996	0.004	0.995	0.005	0.995	0.004	0.955	0.012	0.956	0.009
	0.6	0.990	0.009	0.990	0.007	0.978	0.011	0.978	0.008	0.944	0.017	0.945	0.013
0.97	0	0.995	0.007	0.995	0.005	0.999	0.007	0.999	0.005	0.941	0.010	0.941	0.008
	0.2	0.988	0.010	0.988	0.007	0.986	0.010	0.986	0.008	0.887	0.017	0.888	0.013
	0.6	0.972	0.015	0.972	0.012	0.943	0.016	0.944	0.013	0.863	0.022	0.865	0.017
0.95	0	0.992	0.010	0.992	0.007	1.000	0.009	1.000	0.007	0.908	0.013	0.908	0.011
	0.2	0.979	0.013	0.979	0.011	0.977	0.013	0.977	0.011	0.828	0.019	0.829	0.015
	0.6	0.950	0.019	0.949	0.015	0.910	0.019	0.910	0.015	0.796	0.023	0.796	0.019

Table 7: Comparison of OLS, IV and MODWT (Haar, level of decomposition 4) regressions when the regressor is measured with error.
This table reports the results of the Monte Carlo simulations for OLS, IV and MODWT regressions. The regressor is measured with error. We list the mean value of $\hat{\beta}$ (Mean), the standard deviation of $\hat{\beta}$ (SD), the median of $\hat{\beta}$ (Med) and the mean-absolute deviation of $\hat{\beta}$ (MAD) after $B = 1000$ replications for $T = 2000$. $\alpha = 1$ and $\beta = 1$.

γ	ρ	MODWT				IV				OLS			
		Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)	Mean($\hat{\beta}$)	SD($\hat{\beta}$)	Med($\hat{\beta}$)	MAD($\hat{\beta}$)
0.99	0	0.998	0.005	0.998	0.004	1.000	0.005	1.000	0.004	0.977	0.008	0.978	0.006
	0.2	0.996	0.007	0.996	0.006	0.995	0.007	0.995	0.006	0.954	0.014	0.955	0.011
	0.6	0.990	0.012	0.990	0.009	0.978	0.013	0.979	0.010	0.944	0.018	0.946	0.014
0.97	0	0.995	0.009	0.994	0.007	0.999	0.009	0.999	0.007	0.941	0.011	0.941	0.009
	0.2	0.988	0.013	0.987	0.010	0.986	0.012	0.985	0.010	0.887	0.018	0.887	0.014
	0.6	0.973	0.020	0.973	0.016	0.944	0.020	0.944	0.016	0.863	0.024	0.863	0.019
0.95	0	0.992	0.012	0.992	0.009	1.000	0.011	1.000	0.009	0.908	0.014	0.908	0.011
	0.2	0.980	0.017	0.980	0.013	0.979	0.017	0.979	0.013	0.830	0.022	0.831	0.017
	0.6	0.952	0.027	0.952	0.021	0.911	0.025	0.910	0.019	0.797	0.026	0.795	0.021

Table 8: Comparison of OLS, IV and MODWT (Haar, level of decomposition 4) regressions when the regressor and the regressand are measured with error.
This table reports the results of the Monte Carlo simulations for OLS, IV and MODWT regressions. The regressor and the regressand are measured with error. We list the mean value of $\hat{\beta}$ (Mean), the standard deviation of $\hat{\beta}$ (SD), the median of $\hat{\beta}$ (Med) and the mean-absolute deviation of $\hat{\beta}$ (MAD) after $B = 1000$ replications for $T = 2000$. $\alpha = 1$ and $\beta = 1$.

References

Amemiya, Y. (1985). Instrumental variable estimator for the nonlinear errors-in-variables model. *Journal of Econometrics*, **28**, 273–289.

Bound, J., Brown, C., and Mathiowetz, N. (2001). *Measurement Error in Survey Data*, pages 3705–3843. Handbook of Econometrics. New York: North Holland Publishing.

Carmichael, B. and Coën, A. (2008). Asset pricing models with errors-in-variables. *Journal of Empirical Finance*, **15**(4), 778–788.

Cashin, P., McDermott, C., and Pattillo, C. (2004). Terms of trade shocks in africa: Are they short-lived of long-lived? *Journal of Development Economics*, **73**(3), 727–744.

Chen, X., Hong, H., and Tamer, E. (2005). Measurement error models with auxiliary data. *Review of Economic Studies*, **72**, 343366.

Cragg, J. (1994). Making good inferences from bad data. *Canadian Journal of Economics*, **27**(4), 776800.

Cragg, J. (1997). Using higher moments to estimate the simple errors-in-variables model. *RAND Journal of Economics*, **28**, 7191.

Dagenais, M. and Dagenais, D. (1997). Higher moment estimators for linear regression models with errors in the variables. *Journal of Econometrics*, **76**(1-2), 193221.

Davidson, R. and MacKinnon, J. (2004). *Econometric Theory and Methods*. Oxford University Press, New York.

Davidson, R., Labys, W. C., and Lesourd, J.-B. (1998). Walvelet analysis of commodity price behavior. *Computational Economics*, **11**, 103–128.

Duchesne, P. (2006a). On testing for serial correlation with a wavelet-based spectral density estimator in multivariate time series. *Econometric Theory*, **22**, 633–676.

Duchesne, P. (2006b). Testing for multivariate autoregressive conditional heteroskedasticity using wavelets. *Computational Statistics & Data Analysis*, **51**, 2142–2163.

Elliott, G. and Stock, J. (1994). Inference in time series regression when the order of integration of a regressor is unknown. *Econometric Theory*, **10**, 672700.

Fama, E. and MacBeth, J. (1973). Risk, return, and equilibrium: empirical tests. *Journal of Political Economy*, **81**(3), 607636.

Fan, Y. (2003). On the approximate decorrelation property of the discrete wavelet transform for fractionally differenced processes. *IEEE Transactions on Information Theory*, **49**, 516–521.

Fan, Y. and Whitcher, B. (2003). A wavelet solution to the spurious regression of fractionally differenced processes. *Applied Stochastic Models in Business and Industry*, **19**, 171–183.

Ferson, W. and Locke, D. (1998). Estimating the cost of capital through time: An analysis of the sources of error. *Management Science*, **44**, 485–500.

Gençay, R. and Fan, Y. (2009). Unit root tests with wavelets. *Econometric Theory*. Forthcoming.

Gençay, R., Selçuk, F., and Whitcher, B. (2003). Systematic risk and time scales. *Quantitative Finance*, **3**, 108–116.

Gençay, R., Selçuk, F., and Whitcher, B. (2005). Multiscale systematic risk. *Journal of International Money and Finance*, **24**, 55–70.

Hausman, J., Ichimura, H., Newey, W., and Powell, J. (1991). Measurement errors in polynomial regression models. *Journal of Econometrics*, **50**, 271295.

Hong, Y. (2000). Wavelet-based estimation for heteroskedasticity and autocorrelation consistent variance-covariance matrices. Technical report, Department of Economics and Department of Statistical Science, Cornell University.

Hong, Y. and Kao, C. (2004). Wavelet-based testing for serial correlation of unknown form in panel models. *Econometrica*, **72**, 1519–1563.

Hong, Y. and Lee, J. (2001). One-sided testing for ARCH effects using wavelets. *Econometric Theory*, **17**, 1051–1081.

Horowitz, J. and Manski, C. (1995). Identification and robustness with contaminated and corrupted data. *Econometrica*, **63**, 281–302.

Lee, J. and Hong, Y. (2001). Testing for serial correlation of unknown form using wavelet methods. *Econometric Theory*, **17**, 386–423.

Lee, L.-F. and Sepanski, J. (1985). Estimation of linear and nonlinear error-in-variable models using validation data. *Journal of American Statistical Association*, **90**, 130–140.

Levi, M. (1973). Errors in the variables bias in the presence of correctly measured variables. *Econometrica*, **41**, 985–986.

Molinari, F. (2008). Partial identification of probability distributions with misclassified data. *Journal of Econometrics*, **144**(1), 81–117.

Newey, W. (2001). Flexible simulated moment estimation of nonlinear errors-in-variables models. *Review of Economics and Statistics*, **83**, 616627.

Newey, W. and West, K. (1987). A simple positive semidefinite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55**, 703–707.

Pastor, L. and Stambaugh, R. (1999). Costs of equity capital and model mispricing. *Journal of Finance*, **54**(1), 67–121.

Percival, D. B. and Mofjeld, H. O. (1997). Analysis of subtidal coastal sea level fluctuations using wavelets. *Journal of the American Statistical Association*, **92**, 868–880.

Percival, D. B. and T., W. A. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge Press, Cambridge.

Phillips, P. (1987). Towards a unified asymptotic theory of autoregression. *Biometrika*, **74**, 535–547.

Ramsey, J. B. (1999). The contribution of wavelets to the anlaysis of economic and financial data. *Philosophical Transactions of the Royal Society of London A*, **357**, 2593–2606.

Ramsey, J. B. (2002). Wavelets in economics and finance: Past and future. *Studies in Nonlinear Dynamics & Econometrics*, **3**, 1–29.

Ramsey, J. B. and Zhang, Z. (1997). The analysis of foreign exchange data using waveform dictionaries. *Journal of Empirical Finance*, **4**, 341–372.

Ramsey, J. B., Zaslavsky, G., and Usikov, D. (1995). An analysis of U. S. stock price behavior using wavelets. *Fractals*, **3**, 377–389.

Schennach, S. (2004). Exponential specifications and measurement error. *Economics Letters*, **85**, 8591.

Sepanski, J. and Carroll, R. (1993). Semiparametric quasilikelihood and variance function estimation in measurement error models. *Journal of Econometrics*, **58**, 223–256.

Shanken, J. (1992). On the estimation of beta-pricing models. *Review of Financial Studies*, **5**(1), 1–33.

Wang, L. and Hsiao, C. (2003). Identification and estimation of semiparametric nonlinear errors-in-variables models. Working Paper.