Estimation and Inference of Linear Trend Slope Ratios with an Application to Global Temperature Data

forthcoming in

Journal of Time Series Analysis

Timothy J. Vogelsang*
Department of Economics
Michigan State University

Nasreen Nawaz
Faculty of Finance and Accounting
Shanghai Business School

November 5, 2015, Revised July 15, 2016

Abstract

We focus on estimation and inference of the ratio of trend slopes between two time series where the trending behavior of each series can be well approximated by a simple linear time trend. Our methodological results are motivated by a recent empirical climate literature that seeks to estimate and test hypotheses about the relative rate of warming in the lower-troposphere relative to surface warming - the so-called amplification ratio. We analyze the statistical properties of several estimators and test statistics that are configured to allow serial correlation in the data. The relative merits of the estimators and test statistics depend on the magnitude of the trend slopes relative to the noise in the data. Based on asymptotic theory and finite sample evidence, we make specific and concrete recommendations for practitioners. We apply the recommended estimator and confidence intervals to temperature data from the 1979-2014 period. We find that amplification ratios typically associated with climate models are rejected by the observed temperature data confirming and extending the empirical findings of Klotzbach *et al* (2009, 2010). Allowing for a structural change at the end of 1998 to account for the so-called "hiatus" in warming gives results similar to Klotzbach *et al* (2009, 2010).

Keywords: Trend Stationary, Bias Correction, HAC Estimator, Fixed-b Asymptotics, Amplification Ratio, Long Run Variance

*We thank three referees, a guest editor, Kyoo il Kim, Min Seong Kim, Phil Klotzbach, Aeimit Ladawala, Dick McNider, Jeff Wooldridge and participants at the 2nd International Workshop on Econometric Applications in Climatology (U. Guelph, April 2015) for helpful comments and suggestions. Correspondence: Tim Vogelsang, Department of Economics, 486 W. Circle Drive, 110 Marshall-Adams Hall, Michigan State University East Lansing, MI 48824-1038. Phone: 517-353-4582, Fax: 517–432-1068, email: tjv@msu.edu

1 Introduction

A time trend refers to systematic behavior of a time series that can be approximated by a function of time. Often when plotting macroeconomic or climate time series data, one notices a tendency for the series to increase (or decrease) over time. In some cases it is immediately apparent from the time series plot that the trend is approximately linear. In the econometrics literature there is a well developed literature on estimation and robust inference of deterministic trend functions with a focus on the case of the simple linear trend model. See for example, Canjels and Watson (1997), Vogelsang (1998), Bunzel and Vogelsang (2005), Harvey, Leybourne and Taylor (2009), and Perron and Yabu (2009).

When analyzing more than one time series with trending behavior, it may be interesting to compare the trending behavior across series as in Vogelsang and Franses (2005). Empirically, comparisons across trends are often made in the economic convergence literature where growth rates of gross domestic products (GDPs), i.e. trend slopes of GDPs, are compared across regions or countries. See for example, Fagerberg (1994), Evans (1997) and Tomljanovich and Vogelsang (2002). Often empirical work in the economic convergence literature seeks to determine whether countries or regions have growth rates that are consistent with convergence that has either occurred or is occurring. However, there is little, if any, focus on estimating and quantifying the relative speed by which convergence is occurring. An exception is the recent paper by Vogelsang and Nawaz (2015) where it is argued that the speed of convergence can be measured, in part, by the ratio of trend slopes.

Estimation and inference regarding trend slope ratios plays an important role in recent work in the empirical climate literature that documents the relative warming rates between surface temperatures and lower-troposphere temperatures. See Santer *et al* (2005), Thorne *et al* (2007), Klotzbach *et al* (2009, 2010), Christy *et al* (2010), Po-Chedley and Fu (2012) and the references cited therein. In this literature there is an explicit interest in estimating the ratio of trend slopes of lower-troposphere and surface temperature series - the so-called *amplification ratio.* What is missing from this empirical climate literature are sound statistical methods for computing confidence intervals for trend slope ratios. In fact, most papers in this literature report estimated trend slope ratios without reporting standard errors or confidence intervals.

This paper has two goals. First, we develop reliable statistical methods for estimation and inference of trend slope ratios, and we provide practitioners with concrete recommendations. Because temperature series are known to have serial correlation, the inference methods we propose are configured to be robust to serial correlation. Second, we revisit the empirical analysis of Klotzbach *et al* (2009, 2010) where trend slope ratios (amplification ratios) were estimated between surface and lower-troposphere temperatures. Klotzbach *et al* (2009, 2010) reported estimated amplification ratios but did not provide confidence intervals given the lack of statistical methodology in this area. The empirical contribution of this paper is to construct serial correlation robust confidence intervals for amplification ratios using the same temperature series as used by Klotzbach *et al* (2009, 2010) but extending the analysis to 2014 (Klotzbach *et al* (2009, 2010) used data ending in 2008). Our confidence intervals allow empirical researchers to statistically compare estimated amplification ratios from observed temperature series with amplification ratios of theoretical climate models used for projections of future climate scenarios. We also sketch methods for estimation and inference of amplification ratios when the trend functions are allowed to have a one-time structural change at a given date.

With regards to methodology, we are unaware of any papers in the statistics, econometrics or climate literatures focusing on estimation and inference of trend slope ratios. Our first goal fills that methodological hole in the literature. We focus on estimation of the ratio of trend slopes between two time series where is it reasonable to assume that the trending behavior of each series can be well approximated by a simple linear time trend. We obtain results under the assumption that the stochastic parts of the two time series comprise a zero mean time series vector that has sufficient stationarity and has dependence that is weak enough so that scaled partial sums of the vector satisfy a functional central limit theorem (FCLT)[1]. We compare two natural estimators of the trend slope ratio and propose a third bias-corrected estimator. We show how to use these three estimators to carry out inference about the trend slope ratio. When trend slopes are small in magnitude relative to the variation in the stochastic components (the trend slopes are small relative to the noise), we find that inference using any of the three estimators is compromised and could be potentially misleading. We propose an alternative inference procedure that remains valid when trend slopes are small or even zero.

We carry out an extensive theoretical analysis of the estimators and inference procedures. Our theoretical framework explicitly captures the impact of the magnitude of the trend slopes on the estimation and inference about the trend slope ratio. Our theoretical results are constructive in two important ways. First, the theory points to one of the three estimators as being preferred in terms of bias. Second, the theory strongly suggests that our alternative inference procedure is superior under the null and maintains reliable power under the alternative Finite sample simulations indicate that the predictions made by the asymptotic theory are relevant in practice. Therefore, we are able to give concrete and specific advice to practitioners on how to estimate and construct confidence intervals for a ratio of trend slopes.

The remainder of the paper is organized as follows: Section 2 describes the model and analyzes the asymptotic properties of the three estimators of the trend slope ratio. Section 3 provides some finite sample evidence on the relative performance of the three estimators. Section 4 investigates inference regarding the trend slope ratio. We show how to construct heteroskedasticity autocorrelation (HAC) robust tests using each of the three estimators. We propose an alternative testing approach and show how to compute confidence intervals for this approach. We derive asymptotic results of the tests under the null and under local alternatives. The asymptotic theory clearly shows that our alternative testing approach works well under both the null and local alternatives. Additional finite sample simulation results reported in Section 5 indicate that the predictions of the asymptotic theory are relevant in practice. In Section 6 we make some practical recommendations for empirical researchers. Section 7 shows how to extend the modelling to allow a one-time shift in intercept and trend at an unknown date. Section 8 revisits and extends the empirical analysis of Klotzbach *et al* (2009, 2010). We find that confidence intervals for amplification ratios for observed temperature series do not contain amplification ratios that are typical of climate models. Section 9 concludes. Proofs of theorems and plots of the data are provided in a Supplementary Information (SI)

[1] The case where the stochastic components have nonstationary unit root behavior is analyzed in Vogelsang and Nawaz (2015). There, the empirical application focuses on convergence of gross domestic product between regions of the United States.

document.

2 The Model and Estimation

2.1 Statistical Model and Assumptions

Suppose the univariate time series y_{1t} and y_{2t} are given by

$$y_{1t} = \mu_1 + \beta_1 t + u_{1t}, \tag{1}$$

$$y_{2t} = \mu_2 + \beta_2 t + u_{2t}, \tag{2}$$

where u_{1t} and u_{2t} are mean zero covariance stationary processes. Assume that

$$T^{-1/2}\sum_{t=1}^{[rT]} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \Rightarrow \mathbf{\Lambda W}(r) \equiv \begin{bmatrix} B_1(r) \\ B_2(r) \end{bmatrix}, \tag{3}$$

where $r \in [0,1]$, $[rT]$ is the integer part of rT and $W(r)$ is a 2×1 vector of independent standard Wiener processes. Λ is not necessarily diagonal allowing for correlation between u_{1t} and u_{2t}. In addition to (3), we assume that u_{1t} and u_{2t} are ergodic for the first and second moments.

Suppose that $\beta_2 \neq 0$ and we are interested in estimating the parameter

$$\theta = \frac{\beta_1}{\beta_2}$$

which is the ratio of trend slopes. Equation (1) can be rewritten so that y_{1t} depends on θ through y_{2t}. Rearranging (2) gives

$$t = \frac{1}{\beta_2}\left[y_{2t} - \mu_2 - u_{2t}\right], \tag{4}$$

and plugging this expression into Equation (1) and then rearranging, we obtain

$$y_{1t} = (\mu_1 - \frac{\beta_1}{\beta_2}\mu_2) + \frac{\beta_1}{\beta_2}y_{2t} + (u_{1t} - \frac{\beta_1}{\beta_2}u_{2t}) = (\mu_1 - \theta\mu_2) + \theta y_{2t} + (u_{1t} - \theta u_{2t}).$$

Defining $\delta = \mu_1 - \theta\mu_2$ and $\epsilon_t(\theta) = u_{1t} - \theta u_{2t}$ gives the regression model

$$y_{1t} = \delta + \theta y_{2t} + \epsilon_t(\theta). \tag{5}$$

Given the definition of $\epsilon_t(\theta)$, it immediately follows from (3) that

$$T^{-1/2}\sum_{t=1}^{[rT]} \epsilon_t(\theta) \Rightarrow \lambda_\theta w(r), \tag{6}$$

where $w(r)$ is a univariate standard Wiener process and $\lambda_\theta^2 = \begin{bmatrix} 1 & -\theta \end{bmatrix} \Lambda\Lambda' \begin{bmatrix} 1 & -\theta \end{bmatrix}'$ is the long run variance of $\epsilon_t(\theta)$.

2.2 Estimation of the Trend Slope Ratio

Using regression (5), the natural estimator of θ is ordinary least squares (OLS) which is defined as

$$\widetilde{\theta} = \left(\sum_{t=1}^{T} (y_{2t} - \overline{y}_2)^2 \right)^{-1} \sum_{t=1}^{T} (y_{2t} - \overline{y}_2)(y_{1t} - \overline{y}_1), \tag{7}$$

where $\overline{y}_1 = T^{-1} \sum_{t=1}^{T} y_{1t}$ and $\overline{y}_2 = T^{-1} \sum_{t=1}^{T} y_{2t}$. Standard algebra gives the relationship

$$\widetilde{\theta} - \theta = \left(\sum_{t=1}^{T} (y_{2t} - \overline{y}_2)^2 \right)^{-1} \sum_{t=1}^{T} (y_{2t} - \overline{y}_2) \epsilon_t(\theta). \tag{8}$$

Alternatively, one could estimate θ by the analogy principle by simply replacing β_1 and β_2 with estimators. Let $\widehat{\beta}_1$ and $\widehat{\beta}_2$ be the OLS estimators of β_1 and β_2 based on regressions (1) and (2):

$$\widehat{\beta}_i = \left(\sum_{t=1}^{T} (t - \overline{t})^2 \right)^{-1} \sum_{t=1}^{T} (t - \overline{t})(y_{it} - \overline{y}_i), \quad i = 1, 2, \tag{9}$$

where $\overline{t} = T^{-1} \sum_{t=1}^{T} t$ is the sample average of time and define $\widehat{\theta} = \widehat{\beta}_1 / \widehat{\beta}_2$. Simple algebra shows that

$$\widehat{\theta} = \frac{\widehat{\beta}_1}{\widehat{\beta}_2} = \left(\sum_{t=1}^{T} (t - \overline{t})(y_{2t} - \overline{y}_2) \right)^{-1} \sum_{t=1}^{T} (t - \overline{t})(y_{1t} - \overline{y}_1) \tag{10}$$

which is the instrumental variable (IV) estimator of θ in (5) where t has been used as an instrument for y_{2t}. Standard algebra gives the relationship

$$\widehat{\theta} - \theta = \left(\sum_{t=1}^{T} (t - \overline{t})(y_{2t} - \overline{y}_2) \right)^{-1} \sum_{t=1}^{T} (t - \overline{t}) \epsilon_t(\theta). \tag{11}$$

2.3 Asymptotic Properties of OLS and IV

We now explore the asymptotic properties of the OLS and IV estimators of θ. The asymptotic behavior of the estimators depends on the magnitude of the trend slope parameters relative to the variation in the random components, u_{1t} and u_{2t}, i.e. the noise. The following theorem summarizes the asymptotic behavior of the estimators for fixed βs and for βs that are modeled as local to zero at rate $T^{-1/2}$.

Theorem 1 *Suppose that (6) holds and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following hold as $T \to \infty$. Case 1 (large trend slopes): For $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,*

$$T^{3/2} \left(\widetilde{\theta} - \theta \right) \Rightarrow \left(\overline{\beta}_2^2 \int_0^1 \left(s - \frac{1}{2} \right)^2 ds \right)^{-1} \overline{\beta}_2 \lambda_\theta \int_0^1 \left(s - \frac{1}{2} \right) dw(s) \sim N \left(0, \frac{12 \lambda_\theta^2}{\overline{\beta}_2^2} \right),$$

$$T^{3/2} \left(\widehat{\theta} - \theta \right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 \left(s - \frac{1}{2} \right)^2 ds \right)^{-1} \lambda_\theta \int_0^1 \left(s - \frac{1}{2} \right) dw(s) \sim N \left(0, \frac{12 \lambda_\theta^2}{\overline{\beta}_2^2} \right).$$

Case 2 (medium trend slopes): For $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,

$$T\left(\widetilde{\theta}-\theta\right) \Rightarrow \left(\overline{\beta}_2^2 \int_0^1 \left(s-\frac{1}{2}\right)^2 ds\right)^{-1} \left[\overline{\beta}_2 \lambda_\theta \int_0^1 \left(s-\frac{1}{2}\right) dw(s) + E(u_{2t}\epsilon_t(\theta))\right] \sim N\left(\frac{12}{\overline{\beta}_2^2} E(u_{2t}\epsilon_t), \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right),$$

$$T\left(\widehat{\theta}-\theta\right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 \left(s-\frac{1}{2}\right)^2 ds\right)^{-1} \lambda_\theta \int_0^1 \left(s-\frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right).$$

Theorem 1 makes some interesting predictions about the sampling properties of OLS and IV. When the trend slopes are fixed, i.e. when the trend slopes are large relative to the noise, OLS and IV converge to the true value of θ at the rate $T^{3/2}$ and are asymptotically normal with equivalent asymptotic variances. The precision of both estimators improves when there is less noise (λ_θ^2 is smaller) or when the magnitude of the trend slope parameter for y_{2t} increases (β_2 is larger).

When the trend slopes are modeled as local to zero at rate $T^{-1/2}$, i.e. when trend slopes are medium sized relative to the noise, asymptotic equivalence of OLS and IV no longer holds. The IV estimator essentially has the same asymptotic behavior as in the fixed slopes case because the implied approximations are the same. In contrast, the result for OLS is markedly different in Case 2. While OLS consistently estimates θ and the asymptotic variance is the same as in Case 1, OLS now has an asymptotic bias that could matter when trend slopes are medium sized.

The fact that OLS is asymptotically biased in Case 2 is not that surprising because $\epsilon_t(\theta)$ is correlated with y_{2t} through the correlation[2] between $\epsilon_t(\theta)$ and u_{2t}. In Case 2, the trend slopes are small enough so that the covariance between u_{2t} and $\epsilon_t(\theta)$ asymptotically affects the OLS estimator. Because $E(u_{2t}\epsilon_t(\theta)) = E(u_{1t}u_{2t}) - \theta E(u_{2t}^2)$, the asymptotic bias will be non-zero unless $E(u_{1t}u_{2t}) = \theta E(u_{2t}^2)$ which only happens in very particular special cases. In general, OLS has an asymptotic bias when trend slopes are medium sized. Because Theorem 1 explicitly characterizes the bias of OLS, a feasible bias correction is possible.

2.4 Bias Corrected OLS

The approximate bias of OLS suggested by Theorem 1, Case 2 is given by the quantity

$$bias(\widetilde{\theta}) \approx T^{-1}\left(\overline{\beta}_2^2 \int_0^1 \left(s-\frac{1}{2}\right)^2 ds\right)^{-1} E(u_{2t}\epsilon_t(\theta)).$$

We can estimate $\overline{\beta}_2^2 \int_0^1 \left(s-\frac{1}{2}\right)^2 ds$ using $T^{-2}\sum_{t=1}^T (y_{2t}-\overline{y}_2)^2$, and we can estimate $E(u_{2t}\epsilon_t(\theta))$ using $T^{-1}\sum_{t=1}^T \widehat{u}_{2t}\widetilde{\epsilon}_t$ where $\widetilde{\epsilon}_t$ are the OLS residuals from regression (5) and $\widehat{u}_{2t}$ are the OLS residuals from regression (2). This leads to the bias corrected OLS estimator of θ given by

$$\widetilde{\theta}^c = \widetilde{\theta} - T^{-1}\left(\frac{T^{-1}\sum_{t=1}^T \widehat{u}_{2t}\widetilde{\epsilon}_t}{T^{-2}\sum_{t=1}^T (y_{2t}-\overline{y}_2)^2}\right) = \widetilde{\theta} - \frac{T^{-1}\sum_{t=1}^T \widehat{u}_{2t}\widetilde{\epsilon}_t}{T^{-1}\sum_{t=1}^T (y_{2t}-\overline{y}_2)^2}. \quad (12)$$

[2] The fact that IV is not asymptotically biased in Case 2 is also not suprising given that t, the instrument for y_{2t}, is nonrandom and cannot be correlated with $\epsilon_t(\theta)$.

The next theorem gives the asymptotic behavior of the bias corrected OLS estimator for the same cases covered by Theorem 1.

Theorem 2 *Suppose that (6) holds and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following hold as $T \to \infty$. Case 1 (large trend slopes): For $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,*

$$T^{3/2}\left(\widetilde{\theta}^c - \theta\right) \Rightarrow \left(\overline{\beta}_2^2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \overline{\beta}_2 \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right).$$

Case 2 (medium trend slopes): For $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,

$$T\left(\widetilde{\theta}^c - \theta\right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right).$$

As Theorem 2 shows, the bias corrected OLS estimator is asymptotically equivalent to the IV estimator for both large and medium trend slopes.

2.5 Asymptotic Properties of Estimators for Small Trend Slopes

As shown by Theorem 1, the magnitudes of the trend slopes relative to the noise can affect the behavior of estimators of the trend slope ratio, θ. Intuitively, we know that as the trend slopes become very small in magnitude, we approach the case where the trend slopes are zero in which case θ is not well defined. While it is clear that OLS and possibly bias-corrected OLS will have problems when trend slopes are very small, IV is also expected to have problems in this case. If the trend slopes are very small, then the sample correlation between t and y_{2t} also becomes very small and t becomes a weak instrument for y_{2t}. It is well known that weak instruments have important implications for IV estimation (see Staiger and Stock 1997) and estimation of θ is no exception.

The next two theorems provide asymptotic results for the estimators of θ for trend slopes that are local to zero at rates T^{-1} and $T^{-3/2}$ with the latter case corresponding to trend slopes that are very small relative to the noise.

Theorem 3 *Suppose that (3) and (6) hold and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following hold as $T \to \infty$. Case 3 (small trend slopes): For $\beta_1 = T^{-1}\overline{\beta}_1, \beta_2 = T^{-1}\overline{\beta}_2$,*

$$\widetilde{\theta} - \theta \xrightarrow{p} \Re, \qquad \widetilde{\theta}^c - \theta \xrightarrow{p} \Re_c,$$

$$T^{1/2}(\widehat{\theta} - \theta) \Rightarrow \left(\overline{\beta}_2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right),$$

where

$$\Re = \left(\overline{\beta}_2^2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds + E(u_{2t}^2)\right)^{-1} E(u_{2t}\epsilon_t(\theta)), \quad \Re_c = \frac{\Re^2 E(u_{2t}^2)}{E(u_{2t}\epsilon_t(\theta))}.$$

Case 4 (very small trend slopes): For $\beta_1 = T^{-3/2}\overline{\beta}_1, \beta_2 = T^{-3/2}\overline{\beta}_2$,

$$\widetilde{\theta} - \theta \xrightarrow{p} \frac{E(u_{2t}\epsilon_t(\theta))}{E(u_{2t}^2)}, \quad \widetilde{\theta}^c - \theta \xrightarrow{p} \frac{E(u_{2t}\epsilon_t(\theta))}{E(u_{2t}^2)},$$

$$\widehat{\theta} - \theta \Rightarrow \left(\overline{\beta}_2 \int_0^1 (s - \frac{1}{2})^2 ds + \int_0^1 (s - \frac{1}{2}) dB_2(s)\right)^{-1} \lambda_\theta \int_0^1 (s - \frac{1}{2}) dw(s).$$

Theorem 3 shows that for small trend slopes, OLS and bias-corrected OLS are biased and inconsistent. In contrast, IV has the same asymptotic behavior as for large and medium trend slopes. When trend slopes are very small, all three estimators are inconsistent and biased. The IV estimator converges to a ratio of normal random variables that are correlated with each other because $B_2(r)$ is correlated with $w(r)$ as long as u_{2t} is correlated with $\epsilon_t(\theta)$. Theorem 3 predicts that none of the estimators of θ will work well when trend slopes are very small. This is not surprising because it is very difficult to identify the ratio of trend slopes when the noise dominates the information in the sample regarding the trend slopes themselves.

2.6 Implications (Predictions) of Asymptotics for Finite Samples

Theorems 1-3 make clear predictions about the finite sample behavior of the three estimators of θ. For large βs the three estimators should have similar bias, variance and sampling properties given that they are asymptotically equivalent. When βs are of medium size, IV and bias-corrected OLS are asymptotically equivalent and should have similar behavior compared to the large βs case. In contrast, OLS should exhibit some finite sample bias in the medium βs case. With small or very small βs, OLS and bias-corrected OLS are inconsistent and biased. In contrast IV should continue to perform well even with small βs with the main implication of small βs being less precision given that the asymptotic variance of IV is inversely related to the magnitude of β_2. For very small βs IV is inconsistent and can exhibit substantial variability given that IV is approximately a ratio of two normal random variables.

For a given sample size and variability of the noise, as the trend slopes decrease from being large to becoming very small, we should see the performance of all three estimators deteriorating with OLS deteriorating quickest followed by bias-corrected OLS followed by IV.

It is interesting to note that none of the asymptotic results in Theorems 1-3 require $\overline{\beta}_1$ to be non-zero and the results hold for $\overline{\beta}_1 = 0$ in which case $\theta = 0$ is allowed.

3 Finite Sample Means and Standard Deviations of Estimators

In this section we illustrate the finite sample performance of the estimators via a Monte Carlo simulation study. For the data generating process (DGP) that we consider, the finite sample behavior of the three estimators closely follows the predictions suggested by Theorems 1-3.

The following DGP was used. The y_{1t} and y_{2t} variables were generated by models (1) and (2) where the noise is given by

$$u_{1t} = 0.4u_{2t} + 0.3u_{1t-1} + \varepsilon_{1t},$$
$$u_{2t} = 0.5u_{2t-1} + \varepsilon_{2t},$$
$$[\varepsilon_{1t}, \varepsilon_{2t}]' \sim i.i.d.\ N(0, I_2), \quad u_{10} = u_{20} = 0.$$

Given that all three estimators are exactly invariant to the values of $\mu_1$and μ_2, we set $\mu_1 = 0, \mu_2 = 0$ without loss of generality. We report results for various magnitudes of β_1 and β_2 where it is almost always the case that $\theta = \beta_1/\beta_2 = 2$. The exception is when $\beta_1 = 0$, $\beta_2 = 0$ in which case θ is not defined. We report results for $T = 50, 100, 200$ with 10,000 replications used in all cases.

Given that the bias-corrected OLS estimator uses a bias correction based on the OLS residuals from (5), we experimented with an iterative procedure for the bias-corrected OLS estimator that improved its finite sample performance. We first compute $\widetilde{\theta}^c$ as given by (12). Then we updated the OLS residuals using $\widetilde{\theta}^c$ in place of $\widetilde{\theta}$ and recalculated $\widetilde{\theta}^c$. We iterated between updated residuals and bias-correction 100 times.

Table 1 reports estimated means and standard deviations of the three estimators across the 10,000 replications. Focusing on the $T = 50$ case we see that when the trend slopes are large ($\beta_2 = 10, 5$), the means and standard deviations of the three estimators are the same and none of the estimators shows any bias. For medium sized trend slopes ($\beta_2 = 2, .2, .15, .1$), bias-corrected OLS and IV have the same means and standard deviations with little bias being present. In contrast, OLS shows bias that increases substantially as β_2 decreases. For all three estimators we see that the standard deviations increase as β_2 decreases as expected. For small and very small trend slopes ($\beta_2 = .05, .02, .002$), OLS and bias-corrected OLS show substantial bias. It is difficult to determine whether IV is biased given the very large standard deviation of IV in this case. Overall, IV has the least bias but IV becomes very imprecise as the trend slopes approach zero.

Results for the cases of $T = 100, 200$ are similar to the $T = 50$ case. The only difference is that the bias of OLS and bias-corrected OLS increases more slowly as β_2 decreases. With $T = 200$, bias-corrected OLS and IV have the same means and standard deviations for β_2 as small as 0.02.

The results for $\beta_1 = 0$, $\beta_2 = 0$ at first may look surprising but make sense upon deeper inspection. The OLS estimator is no longer estimating θ which is not defined. Instead, OLS is estimating the population quantity $E(u_{2t}\epsilon_t(\theta))/E(u_{2t}^2)$ which is very close to 0.469 in our DGP. Bias-corrected OLS is attempting to correct the wrong bias and the IV estimator is based on an instrument that has zero correlation with y_{2t}. In fact, the estimators are behaving as expected when the trend slopes are zero.

Overall, the finite sample means and variances exhibit patterns as predicted by the asymptotic theory. IV and bias-corrected OLS work equally well for large, medium and somewhat small trend slopes. As trend slopes become very small, none of the estimators are very good and this is to be expected given that the data has relatively little information about the trend slope ratio when trend slopes are small relative to the noise and/or the sample size is small.

4 Inference

In this section we analyze test statistics for testing simple hypotheses about θ. Suppose we are interested in testing the null hypothesis

$$H_0 : \theta = \theta_0, \tag{13}$$

against the alternative hypothesis

$$H_1 : \theta = \theta_1 \neq \theta_0.$$

It is straightforward to construct HAC robust statistics using the three estimators of θ as

$$t_{OLS} = \frac{(\widetilde{\theta} - \theta_0)}{\sqrt{\widetilde{\lambda}^2_{\theta} \left[\sum_{t=1}^{T} (y_{2t} - \overline{y}_2)^2 \right]^{-1}}}, \tag{14}$$

$$t_{BC} = \frac{(\widetilde{\theta}^c - \theta_0)}{\sqrt{\widetilde{\lambda}^2_{\theta c} \left[\sum_{t=1}^{T} (y_{2t} - \overline{y}_2)^2 \right]^{-1}}}, \tag{15}$$

$$t_{IV} = \frac{(\widehat{\theta} - \theta_0)}{\sqrt{\widehat{\lambda}^2_{\theta} \left[\sum_{t=1}^{T} (t - \overline{t})(y_{2t} - \overline{y}_2) \right]^{-2} \sum_{t=1}^{T} (t - \overline{t})^2}}, \tag{16}$$

where the estimated long run variance estimators are given by

$$\widetilde{\lambda}^2_{\theta} = \widetilde{\gamma}_0 + 2 \sum_{j=1}^{T-1} k\left(\frac{j}{M}\right) \widetilde{\gamma}_j, \quad \widetilde{\gamma}_j = T^{-1} \sum_{t=j+1}^{T} \widetilde{\epsilon}_t \widetilde{\epsilon}_{t-j},$$

$$\widetilde{\lambda}^2_{\theta c} = \widetilde{\gamma}^c_0 + 2 \sum_{j=1}^{T-1} k\left(\frac{j}{M}\right) \widetilde{\gamma}^c_j, \quad \widetilde{\gamma}^c_j = T^{-1} \sum_{t=j+1}^{T} \left(\widetilde{\epsilon}^c_t - \overline{\widetilde{\epsilon}^c}\right)\left(\widetilde{\epsilon}^c_{t-j} - \overline{\widetilde{\epsilon}^c}\right),$$

$$\widehat{\lambda}^2_{\theta} = \widehat{\gamma}_0 + 2 \sum_{j=1}^{T-1} k\left(\frac{j}{M}\right) \widehat{\gamma}_j, \quad \widehat{\gamma}_j = T^{-1} \sum_{t=j+1}^{T} \widehat{\epsilon}_t \widehat{\epsilon}_{t-j},$$

with $\widetilde{\epsilon}_t = y_{1t} - \widetilde{\delta} - \widetilde{\theta} y_{2t}$ being the OLS residuals from (5), $\widetilde{\epsilon}^c_t = y_{1t} - \widetilde{\delta} - \widetilde{\theta}^c y_{2t}$ being the bias-corrected OLS residuals, and $\widehat{\epsilon}_t = y_{1t} - \widehat{\delta} - \widehat{\theta} y_{2t}$ being the IV residuals from (5). Because

$$\widetilde{\epsilon}^c_t = y_{1t} - \widetilde{\delta} - \widetilde{\theta} y_{2t} - \left(\widetilde{\theta}^c - \widetilde{\theta}\right) y_{2t} = \widetilde{\epsilon}_t - \left(\widetilde{\theta}^c - \widetilde{\theta}\right) y_{2t},$$

the bias-corrected OLS residuals do not sum to zero. Therefore, we construct $\widetilde{\lambda}^2_{\theta c}$ using $\widetilde{\epsilon}^c_t - \overline{\widetilde{\epsilon}^c}$ where $\overline{\widetilde{\epsilon}^c} = T^{-1} \sum_{t=1}^{T} \widetilde{\epsilon}^c_t$, i.e. we demean $\widetilde{\epsilon}^c_t$ before computing $\widetilde{\lambda}^2_{\theta c}$. The long run variance estimators are constructed using the kernel weighting function $k(x)$ and M is the bandwidth tuning parameter.

4.1 Linear in Slopes Approach

Because all three estimators of θ deteriorate as the trend slopes approach zero, we consider a fourth test statistic that is exactly invariant to the true values of the slope parameters under H_0. Given the null value of θ_0, H_0 and H_1 can be written in terms of the trend slopes as

$$H_0 : \frac{\beta_1}{\beta_2} = \theta_0, \quad H_1 : \frac{\beta_1}{\beta_2} = \theta_1$$

which is a nonlinear restriction on the trend slopes. Obviously, the restrictions implied by these hypotheses can be written as linear functions of β_1 and β_2 as

$$H_0 : \beta_1 - \beta_2\theta_0 = 0,$$

$$H_1 : \beta_1 - \beta_2\theta_0 = \beta_2\theta_1 - \beta_2\theta_0 = \beta_2 (\theta_1 - \theta_0) \neq 0.$$

Given θ_0, define the univariate time series

$$z_t (\theta_0) = y_{1t} - \theta_0 y_{2t},$$

where it follows from (1) and (2) that

$$z_t (\theta_0) = \pi_0(\theta_0) + \pi_1(\theta_0)t + \epsilon_t (\theta_0), \tag{17}$$

where $\pi_0(\theta_0) = \mu_1 - \theta_0\mu_2$, $\pi_1(\theta_0) = \beta_1 - \theta_0\beta_2$ and $\epsilon_t (\theta_0) = u_{1t} - \theta_0 u_{2t}$.

Under H_0 it follows that $\pi_1(\theta_0) = 0$ whereas under H_1 it follows that $\pi_1(\theta_0) = \beta_2 (\theta_1 - \theta_0) \neq 0$. We can test the original null hypothesis given by (13) by testing $H_0 : \pi_1(\theta_0) = 0$ in (17) against the alternative $H_1 : \pi_1(\theta_0) \neq 0$ using the following t-statistic:

$$t_{\theta_0} = \frac{\widehat{\pi}_1(\theta_0)}{\sqrt{\widehat{\lambda}^2_{\theta_0} \left(\sum_{t=1}^{T} (t - \bar{t})^2 \right)^{-1}}}, \tag{18}$$

where

$$\widehat{\pi}_1(\theta_0) = \left(\sum_{t=1}^{T} (t - \bar{t})^2 \right)^{-1} \sum_{t=1}^{T} (t - \bar{t}) \left(z_t (\theta_0) - \bar{z} (\theta_0) \right),$$

$$\widehat{\lambda}^2_{\theta_0} = \widehat{\gamma}^{\theta_0}_0 + 2 \sum_{j=1}^{T-1} k \left(\frac{j}{M} \right) \widehat{\gamma}^{\theta_0}_j, \quad \widehat{\gamma}^{\theta_0}_j = T^{-1} \sum_{t=j+1}^{T} \widehat{\epsilon}_t (\theta_0) \widehat{\epsilon}_{t-j} (\theta_0),$$

$$\widehat{\epsilon}_t (\theta_0) = z_t (\theta_0) - \bar{z} (\theta_0) - \widehat{\pi}_1 (\theta_0) \left(t - \bar{t} \right).$$

Note that $\widehat{\pi}_1(\theta_0)$ is simply the OLS estimator of $\pi_1(\theta_0)$ from (17) and $\widehat{\epsilon}_t (\theta_0)$ are the corresponding OLS residuals. Note that when $\widehat{\lambda}^2_{\theta_0}$ is constructed using the Bartlett kernel with $M = T$, $t^2_{\theta_0}$ is identical to one of the F-statistics proposed by Vogelsang and Franses (2005).

4.2 Confidence Intervals Using t_{θ_0}

Confidence intervals for θ can be constructed by finding the values of θ_0 such that

$$|t_{\theta_0}| \leq cv_{\alpha/2} \tag{19}$$

where $cv_{\alpha/2}$ is the two-tail critical value for significance level α. Because both $\widehat{\pi}_1$ and $\widehat{\lambda}^2_{\theta_0}$ are functions of θ_0, finding the values of θ_0 that result in a non-rejection, i.e. satisfy (19), is equivalent to finding the roots of a particular second-order polynomial. Depending on whether the roots are real or complex, the confidence interval for θ_0 can be a closed interval on the real line, the complement of an open interval on the real line, or the entire real itself.

The form of the confidence interval depends on the magnitudes of the trend slopes relative to the noise as we now explain. Let

$$\widehat{\mathbf{\Omega}} = \widehat{\mathbf{\Gamma}}_0 + \sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)(\widehat{\mathbf{\Gamma}}_j + \widehat{\mathbf{\Gamma}}'_j)$$

where

$$\widehat{\mathbf{\Gamma}}_j = T^{-1} \sum_{t=j+1}^{T} \widehat{\boldsymbol{u}}_t \widehat{\boldsymbol{u}}'_{t-j}, \quad \widehat{\boldsymbol{u}}_t = [u_{1t},\ u_{2t}]'.$$

Partition $\widehat{\mathbf{\Omega}}$ as

$$\widehat{\mathbf{\Omega}} \equiv \begin{bmatrix} \widehat{\Omega}_{11} & \widehat{\Omega}_{12} \\ \widehat{\Omega}_{21} & \widehat{\Omega}_{22} \end{bmatrix}.$$

It is easy to show that

$$\widehat{\pi}_1(\theta_0) = \widehat{\beta}_1 - \theta_0 \widehat{\beta}_2,$$

and

$$\widehat{\lambda}^2_{\theta_0} = \widehat{\Omega}_{11} - 2\theta_0 \widehat{\Omega}_{12} + \theta_0^2 \widehat{\Omega}_{22},$$

allowing us to write (19) as

$$\frac{\left|\widehat{\beta}_1 - \theta_0 \widehat{\beta}_2\right|}{\sqrt{\left(\widehat{\Omega}_{11} - 2\theta_0 \widehat{\Omega}_{12} + \theta_0^2 \widehat{\Omega}_{22}\right)\left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}}} \leq cv_{\alpha/2},$$

or equivalently

$$\frac{\left(\widehat{\beta}_1 - \theta_0 \widehat{\beta}_2\right)^2}{\left(\widehat{\Omega}_{11} - 2\theta_0 \widehat{\Omega}_{12} + \theta_0^2 \widehat{\Omega}_{22}\right)\left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}} \leq cv^2_{\alpha/2}. \tag{20}$$

The inequality given by (20) can be rewritten as

$$c_2 \theta_0^2 + c_1 \theta_0 + c_0 \leq 0, \tag{21}$$

where

$$c_2 = \widehat{\beta}_2^2 - \Psi\widehat{\Omega}_{22}, \quad c_1 = -2(\widehat{\beta}_1\widehat{\beta}_2 - \Psi\widehat{\Omega}_{12}), \quad c_0 = \widehat{\beta}_1^2 - \Psi\widehat{\Omega}_{11}, \quad \Psi = cv_{\alpha/2}^2\left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}.$$

The values of θ_0 that solve (21) depend on the roots of the polynomial $p(\theta_0) = c_2\theta_0^2 + c_1\theta_0 + c_0$. Let r_1 and r_2 denote the roots of $p(\theta_0)$ and when r_1 and r_2 are real, order the roots so that $r_1 \leq r_2$. There are three potential shapes for the confidence interval for θ_0:

Case 1: Suppose that $c_2 > 0$ and $c_1^2 - 4c_2c_0 \geq 0$. In this case, the roots are real and $p(\theta_0)$ opens upwards. The confidence interval is the values of θ_0 between the two roots, i.e. $\theta_0 \in [r_1, r_2]$. The inequality $c_2 > 0$ is equivalent to the inequality

$$\frac{\widehat{\beta}_2^2}{\widehat{\Omega}_{22}\left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}} > cv_{\alpha/2}^2. \tag{22}$$

Notice that the left hand side of (22) is simply the square of the HAC robust t-statistic for testing that the trend slope of y_{2t} is zero. Inequality (22) holds if the trend slope of y_{2t} is statistically different from zero at the α level. This occurs when the trend slope for y_2 is large relative to the variation in u_{2t}. Mechanically, c_2 will be positive when the t-statistic for testing that $\beta_2 = 0$ is large in magnitude. Although not obvious, if $c_2 > 0$, it is impossible for $c_1^2 - 4c_2c_0 < 0$ to hold. In this case $p(\theta_0)$ opens upward and has its vertex above zero and the roots are complex; therefore, there are no solutions to (21) and the confidence interval would be empty. This case is impossible because the confidence interval cannot be empty because $\theta_0 = \widehat{\theta} = \widehat{\beta}_1/\widehat{\beta}_2$ is always contained in the interval given that $\widehat{\beta}_1 - \widehat{\theta}\widehat{\beta}_2 = 0$ in which case (20) must hold (equivalently (21) must hold).

Case 2: Suppose that $c_2 < 0$ and $c_1^2 - 4c_2c_0 > 0$. In this case, $p(\theta_0)$ has two real roots and opens downwards and the confidence interval is the values of θ_0 **not** between the two roots. In this case the confidence interval is the union of two disjoint sets and is given by $\theta_0 \in (-\infty, r_1] \cup [r_2, \infty)$. With $c_2 < 0$ a sufficient condition for $c_1^2 - 4c_2c_0 > 0$ is $c_0 > 0$ which occurs if the trend slope of y_{1t} is statistically different from zero at the α level.

Case 3: Suppose that $c_2 < 0$ and $c_1^2 - 4c_2c_0 \leq 0$. In this case, $p(\theta_0)$ opens downward and has a vertex at or below zero. The entire $p(\theta_0)$ function lies at or below zero and the confidence interval is the entire real line: $\theta_0 \in (-\infty, \infty)$. A necessary condition for Case 3 is $c_0 \leq 0$ in which case neither y_{1t} nor y_{2t} have statistically significant trend slopes.

Although it is a zero probability event, should $c_2 = 0$ then the confidence interval will take the form of either $\theta_0 \in (-\infty, -c_0/c_1]$ when $c_1 > 0$ or $\theta_0 \in [-c_0/c_1, \infty)$ when $c_1 < 0$.

While the confidence intervals constructed using t_{θ_0} can be wide when the trend slopes are small and there is no guarantee that these confidence intervals will contain the OLS or bias-corrected OLS estimators of θ, t_{θ_0} has a major advantage over the other t-statistics. Recall that we can write t_{θ_0} as

$$t_{\theta_0} = \frac{\widehat{\beta}_1 - \theta_0 \widehat{\beta}_2}{\sqrt{\left(\widehat{\Omega}_{11} - 2\theta_0 \widehat{\Omega}_{12} + \theta_0^2 \widehat{\Omega}_{22}\right) \left(\sum_{t=1}^{T} (t - \bar{t})^2\right)^{-1}}}.$$

The denominator is a function $\widehat{u}_{1t}$ and $\widehat{u}_{2t}$ each of which are exactly invariant to the true values of β_1 and β_2. When H_0 is true, it follows that $\beta_1 - \beta_2\theta_0 = 0$ and we can write the numerator of t_{θ_0} as

$$\widehat{\beta}_1 - \theta_0 \widehat{\beta}_2 = \widehat{\beta}_1 - \theta_0 \widehat{\beta}_2 - (\beta_1 - \beta_2 \theta_0) = \left(\widehat{\beta}_1 - \beta_1\right) - \theta_0 \left(\widehat{\beta}_2 - \beta_2\right).$$

Because $\widehat{\beta}_1 - \beta_1$ and $\widehat{\beta}_2 - \beta_2$ are only functions of t and u_{1t}, u_{2t}, the numerator of t_{θ_0} is also exactly invariant to the true values of β_1 and β_2. Therefore, the null distribution of t_{θ_0} is exactly invariant to the true values of β_1 and β_2 including the case where both trend slopes are zero. In contrast, the other t-statistics have null distributions that depend on the magnitudes of β_1 and β_2. Because of its exact invariance to β_1 and β_2 under the null, t_{θ_0} will deliver much more robust inference (with respect to the magnitudes of β_1 and β_2) than the other t-statistics.

4.3 Asymptotic Results for t-statistics

In this section we provide asymptotic limits of the four t-statistics described in the previous sub-section. We derive asymptotic limits under alternatives that are local to the null given by (13). Suppose that $\beta_2 = T^{-\kappa}\overline{\beta}_2$. Then the alternative value of θ_1 is modeled local to θ_0 as

$$\theta_1 = \theta_0 + T^{-3/2+\kappa}\overline{\theta}_\Delta. \tag{23}$$

The parameter $\overline{\theta}_\Delta$ measures the magnitude of the departure from the null under the local alternative. Obviously, asymptotic null distributions are obtained by setting $\overline{\theta}_\Delta = 0$.

Recall that t_{θ_0} is constructed using $\widehat{\pi}_1(\theta_0)$ from (17). Under the local alternative (23), it follows that

$$\pi_1(\theta_0) = \beta_2 (\theta_1 - \theta_0) = \beta_2 T^{-3/2+\kappa}\overline{\theta}_\Delta = T^{-\kappa}\overline{\beta}_2 T^{-3/2+\kappa}\overline{\theta}_\Delta = T^{-3/2}\overline{\beta}_2\overline{\theta}_\Delta, \tag{24}$$

regardless of the value of κ. Therefore, the limit of t_{θ_0} is invariant to the asymptotic nesting used for β_2 under both the null and local alternative for θ.

We derive the limits of the various HAC estimators using fixed-b theory following Bunzel and Vogelsang (2005). Fixed-b theory obtains asymptotic results for the long run variance estimators by treating the bandwidth M, as a fixed proportion of the sample size. In other words, asymptotic results are obtained for $M = bT$ where $b \in (0, 1]$. Recent theoretical work in econometrics and statistics has shown that the fixed-b approach, or more generally the fixed-smoothing approach, leads to asymptotic results for HAC robust test statistics that are more accurate than what has been obtained using more traditional asymptotic theory. See Jansson (2004), Kiefer and Vogelsang (2005), Sun, Phillips and Jin (2008), Zhang and Shao (2013) and Sun (2014).

The form of the fixed-b limits depends on the type of kernel function used to compute the HAC estimator. We follow Bunzel and Vogelsang (2005) and use the following definitions.

Definition 1 *A kernel is labelled Type 1 if $k(x)$ is twice continuously differentiable everywhere and as a Type 2 kernel if $k(x)$ is continuous, $k(x) = 0$ for $|x| \geq 1$ and $k(x)$ is twice continuously differentiable everywhere except at $|x| = 1$.*

We also consider the Bartlett kernel which is neither Type 1 or 2. The fixed-b limiting distributions are expressed in terms of the following stochastic functions.

Definition 2 *Let $Q(r)$ be a generic stochastic process. Define the random variable $P_b(Q(r))$ as*

$$P_b(Q(r)) = \begin{cases} \int_0^1 \int_0^1 -k^{*\prime\prime}(r-s)\, Q(r)Q(s)drds & \text{if } k(x) \text{ is Type 1} \\ \int\int_{|r-s|<b} -k^{*\prime\prime}(r-s)\, Q(r)\, Q(s)\, drds & \\ \quad +2k_{-}^{*\prime}(b) \int_0^{1-b} Q(r+b)\, Q(r)\, dr & \text{if } k(x) \text{ is Type 2} \\ \frac{2}{b}\int_0^1 Q(r)^2\, dr - \frac{2}{b}\int_0^{1-b} Q(r+b)\, Q(r)\, dr & \text{if } k(x) \text{ is Bartlett} \end{cases}$$

where $k^*(x) = k\left(\frac{x}{b}\right)$ and $k_{-}^{*\prime}$ is the first derivative of k^* from below.

The following theorems summarize the asymptotic limits of the t-statistics for testing (13) when the alternative is given by (23).

Theorem 4 *(Large Trend Slopes) Suppose that (6) holds. Let $M = bT$ where $b \in (0, 1]$ is fixed. Let $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$ where $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T, and let $\theta_1 = \theta_0 + T^{-3/2}\overline{\theta}_\Delta$. Then as $T \to \infty$,*

$$t_{OLS}, t_{BC}, t_{IV} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b(Q(r))}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_0}^2 P_b(Q(r))}},$$

where $Z \sim N(0, 1)$, $Q(r) = \widetilde{w}(r) - 12L(r)\int_0^1 \left(s - \frac{1}{2}\right) dw(s)$, $\widetilde{w}(r) = w(r) - rw(1)$, $L(r) = \int_0^r \left(s - \frac{1}{2}\right) ds$ and Z and $Q(r)$ are independent.

Theorem 5 *(Medium Trend Slopes): Suppose that (6) holds. Let $M = bT$ where $b \in (0, 1]$ is fixed. Let $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$ where $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T, and let $\theta_1 = \theta_0 + T^{-1}\overline{\theta}_\Delta$. Then as $T \to \infty$,*

$$t_{OLS} \Rightarrow \frac{Z}{\sqrt{P_b(H_1(r))}} + \frac{12\overline{\beta}_2^{-1} E(u_{2t}\epsilon_t(\theta))}{\sqrt{12\lambda_{\theta_1}^2 P_b(H_1(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b(H_1(r))}},$$

$$t_{BC}, t_{IV} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b(Q(r))}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_0}^2 P_b(Q(r))}},$$

where $H_1(r) = Q(r) - 12\left(\lambda_{\theta_1}\overline{\beta}_2\right)^{-1} L(r) \cdot E(u_{2t}\epsilon_t(\theta))$.

Theorem 6 *(Small Trend Slopes): Suppose that (6) holds. Let $M = bT$ where $b \in (0,1]$ is fixed. Let $\beta_1 = T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$ where $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T, and let $\theta_1 = \theta_0 + T^{-1/2}\overline{\theta}_\Delta$. Then as $T \to \infty$,*

$$t_{OLS}, t_{BC} \xrightarrow{d} \frac{1}{\sqrt{\overline{\beta}_2^2 P_b\left(L(r)\right)\left(\overline{\beta}_2^2 \int_0^1 (s-\frac{1}{2})^2 ds + E\left(u_{2t}^2\right)\right)^{-1}}},$$

$$t_{IV} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2 \overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b(Q(r))}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2 \overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_0}^2 P_b(Q(r))}}.$$

Theorem 7 *(Very Small Trend Slopes): Suppose that (3) and (6) hold. Let $M = bT$ where $b \in (0,1]$ is fixed. Let $\beta_1 = T^{-3/2}\overline{\beta}_1$, $\beta_2 = T^{-3/2}\overline{\beta}_2$ where $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T, and let $\theta_1 = \theta_0 + \overline{\theta}_\Delta$. Then as $T \to \infty$,*

$$T^{-1/2}t_{OLS}, T^{-1/2}t_{BC} = \frac{\left(E(u_{2t}^2)\right)^{-1} E(u_{2t}\epsilon_t(\theta)) + \overline{\theta}_\Delta}{\sqrt{P_b\left(H_2(r)\right)\left[E\left(u_{2t}^2\right)\right]^{-1}}},$$

$$t_{IV} \Rightarrow \frac{\int_0^1 (s-\frac{1}{2})dw(s) + \left(\overline{\beta}_2 \int_0^1 (s-\frac{1}{2})^2 ds + \int_0^1 (s-\frac{1}{2})dB_2(s)\right)\lambda_\theta^{-1}\overline{\theta}_\Delta}{\sqrt{P_b\left(H_3(r)\right)\int_0^1 (s-\frac{1}{2})^2 ds}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2 \overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_0}^2 P_b(Q(r))}},$$

where

$$H_2(r) = \widetilde{w}(r) - \left(\overline{\beta}_2 \int_0^1 (s-\frac{1}{2})^2 ds + \int_0^1 (s-\frac{1}{2})dB_2(s)\right)^{-1} \int_0^1 (s-\frac{1}{2})dw(s)\left(\overline{\beta}_2 L(r) + \widetilde{B}_2(r)\right),$$

$$H_3(r) = \widetilde{w}(r) - 12\left(\overline{\beta}_2 L(r) + \widetilde{B}_2(r)\right)\left(\overline{\beta}_2 + 12\int_0^1 \left(s-\frac{1}{2}\right)dB_2(s)\right)^{-1} \int_0^1 \left(s-\frac{1}{2}\right)dw(s),$$

and $\widetilde{B}_2(r) = B_2(r) - rB_2(1)$.

Some interesting results and predictions about the finite sample behavior of the t-statistics are given by the Theorems 4-7. First consider the limiting null distributions that are obtained when $\overline{\theta}_\Delta = 0$. For large trend slopes, all four t-statistics have the same asymptotic null limit and the limiting random variable is the same fixed-b limit obtained by Bunzel and Vogelsang (2005) for inference regarding the trend slope in a simple linear trend model with stationary errors. Therefore, fixed-b critical values are available from Bunzel and Vogelsang (2005). As the trend slopes become smaller relative to the noise, differences among the t-statistics emerge. As anticipated, t_{θ_0} has the same limiting null distribution regardless of the magnitudes of the trend slopes. Except for very small trend slopes, t_{IV} has the same limiting null distribution as t_{θ_0}. The bias in OLS affects the null limit of t_{OLS} for medium, small and very small trend slopes. The bias correction helps for medium trend slopes in which case t_{BC} has the same null limit as t_{IV} and t_{θ_0}. For small and very small trend slopes the bias correction no longer works effectively and t_{BC} has the same limit as t_{OLS}. Both tests will tend to over-reject under the null when trend slopes are very small given that they diverge with the sample size.

In terms of finite sample null behavior, the asymptotic theory predicts that t_{OLS} and t_{BC} will only work well when trend slopes are relatively large whereas t_{IV} should work well except when trend slopes are very small. The most reliable test in terms of robustness to magnitudes of trend slopes under the null should be t_{θ_0}.

Under the alternative $\overline{\theta}_\Delta \neq 0$ and the t-statistics have additional terms in their limits that push the distributions away from the null distributions giving the tests power. When trend slopes are large, all four t-statistics have the same limiting distributions with the only difference being that t_{θ_0} depends on $\lambda^2_{\theta_0}$ rather than $\lambda^2_{\theta_1}$ as for the other t-statistics. In general we cannot rank $\lambda^2_{\theta_0}$ and $\lambda^2_{\theta_1}$ as any difference depends on the joint serial correlation structure of u_{1t} and u_{2t}. Unless $\lambda^2_{\theta_0}$ and $\lambda^2_{\theta_1}$ are nontrivially different, we would expect power of the tests to be similar in the large trend slope cases. As the trend slopes become smaller, power of t_{OLS} and t_{BC} becomes meaningless given that the statistics have poor behavior under the null. In Theorem 6, the limits of t_{OLS} and t_{BC} do not depend on $\overline{\theta}_\Delta$ which suggests that power will very low when trend slopes are small. In Theorem 7, t_{OLS} and t_{BC} diverge with the sample size which suggests large rejections with very small trend slopes. In contrast power of t_{IV} should be similar to t_{θ_0} except when trend slopes are very small. As in the case of null behavior, the asymptotic theory predicts that t_{θ_0} should have power that is most robust to the magnitudes of the trend slopes.

One thing to keep in mind regarding the limit of t_{θ_0} in Theorems 4-7 is that while the limit under the alternative is the same in each case, the relevant values of $\overline{\theta}_\Delta$ are farther away from the null in the case of smaller trend slopes compared to the case of larger trend slopes. Therefore, θ_1 needs to be much farther away from θ_0 in the case of small trend slopes than for the case of large trend slopes for power of t_{θ_0} to be the same in both cases. In other words, for a given value of θ_1, power of t_{θ_0} decreases as the trend slopes become smaller. This relationship between power and magnitudes of the trend slopes can be seen clearly in Theorem 4 where we can see that as $\beta_2 \to 0$ the limiting distribution under the local alternative for θ approaches the null distribution and power decreases.

5 Finite Sample Null Rejection Probabilities and Power

Using the same DGP as used in Section 3 we simulated finite sample null rejection probabilities and power of the four t-statistics. Table 2 reports null rejection probabilities for 5% nominal level tests for testing $H_0 : \theta = \theta_0 = 2$ against the two-sided alternative $H_1 : \theta \neq 2$. Results are reported for the same values of β_1, β_2 as used in Table 1 for $T = 50, 100, 200$ and $10,000$ replications are used in all cases. The HAC estimators are implemented using the Daniell kernel. Results for three bandwidth sample size ratios are provided: $b = 0.1, 0.5, 1.0$. For a given sample size, T, we use the bandwidth $M = bT$ for each of the three values of b. We compute empirical rejections using fixed-b asymptotic critical values using the critical value function

$$cv_{0.025}(b) = 1.9659 + 4.0603b + 11.6626b^2 + 34.8269b^3 - 13.9506b^4 + 3.2669b^5, \quad (25)$$

as given by Bunzel and Vogelsang (2005) for the Daniell kernel.

The patterns in the empirical null rejections closely match the predictions of the asymptotic results. When the trend slopes are large, $\beta_1 \geq 4$, $\beta_2 \geq 2$, null rejections are the essentially the same for all t-statistics and are close to 0.05 even when $T = 50$. This is true for nearly all three bandwidth choices which illustrates the effectiveness of the fixed-b critical values. The exception is $b = 0.1$ where there are some mild over-rejections. It is well known that small bandwidths can lead to over-rejection problems when serial correlation is positive and the sample size is relatively small. For medium sized trend slopes, $0.1 \leq \beta_1 \leq 0.4$, $0.05 \leq \beta_2 \leq 0.2$, t_{OLS} begins to show over-rejection problems that become very severe as the trend slopes decrease in magnitude. The bias-corrected OLS t-statistic, t_{BC}, is less subject to over-rejection problems especially when T is not small, although for $T = 50$, t_{BC} shows nontrivial over-rejection problems. In contrast both t_{IV} and t_{θ_0} have null rejections close to 0.05 for medium sized trend slopes. When the trend slopes are small or very small, $\beta_1 \leq 0.04$, $\beta_2 \leq 0.02$, the t_{OLS} and t_{BC} statistics have severe over-rejection problems and can reject 100% of the time. While t_{IV} has less over-rejection problems in this case, the over-rejections are nontrivial and are problematic. In contrast, t_{θ_0} has null rejections that are close to 0.05 regardless of the magnitudes of β_1, β_2 including the case of $\beta_1 = \beta_2 = 0$. In fact, the rejections are identical for t_{θ_0} across values of β_1, β_2. This is because t_{θ_0} is exactly invariant to the values of β_1, β_2. It is clear in terms of null rejection probabilities that t_{θ_0} is the preferred test statistic.

Given that t_{θ_0} is the preferred statistic in terms of size, we computed for each of the parameter configurations in Table 2 the proportions of replications that lead to the three possible shapes of confidence intervals obtained by inverting t_{θ_0}. Table 3 gives these results. For large trend slopes Case 1 occurs 100% of the time. As the trend slopes decrease in magnitude, Case 2 occurs some of the time and as the trend slopes decrease further, Case 3 can occur frequently if the trend slopes are very small. As T increases, the likelihood of Case 1 increases for all trend slope magnitudes. The relative frequencies of the three cases also depends on the bandwidth but the relationship appears complicated. This is not surprising given the complex manner in which the bandwidth affects the null distribution of t_{θ_0}. Overall, unless trend slopes are small or very small, Case 1 is the most likely confidence interval shape.

While t_{θ_0} is the preferred test in terms of size, how do the t-statistics compare in terms of power? Table 4 reports power results for a subset of the grid of β_2 as used in Tables 2,3. For a given value of β_2, we specify a grid of nine equally spaced values for θ in the range $\theta_0 \pm 0.04/\beta_2$ where $\theta_0 = 2$ is the null value. By construction $\beta_1 = \theta\beta_2$ in all cases. Given the way we define the grid for θ, we ensure that $\overline{\beta}_2\overline{\theta}_\Delta$ is the same for all values of β_2. Results are reported for $T = 100$. Results for other values of T are qualitatively similar and are omitted.

As in Tables 2 and 3, we report results for $b = 0.5, 1.0$. Rather than reporting results for $b = 0.1$, we report results using the data dependent bandwidth formula for M using the approach of Andrews (1991) which seeks a bandwidth that minimizes the approximate mean square error of the long run variance estimator[3]. We label the corresponding value of b as b_{A91} and we continue to use fixed-b critical values. Looking at null

[3]The approach of Andrews (1991) requires the user to choose an approximate model of serial correlation for the purposes of bandwidth choice. Following common practice in the econometrics literature, we used an autoregressive model with one lag (AR(1)) as the approximating model.

rejections in Table 4 (in bold), we see that rejections are about 0.070 for large and medium trend slopes when using $b = b_{A91}$. In contrast, Table 2 has null rejections equal to 0.054 when using $b = 0.1$. The reason that over-rejections emerge when using $b = b_{A91}$ is because b_{A91} tends to be smaller than 0.1; b_{A91} is approximately equal to 0.05 on average across the replications. If one were to appeal to consistency of the long run variance estimators when using the Andrews (1991) bandwidth rule and use $N(0, 1)$ critical values for the t-statistics, null rejections would increase to 0.098. Using fixed-b critical values reduces over-rejection problems under the null because serial correlation is positive. Notice that as $b \to 0$, the critical value function approaches the $N(0, 1)$ critical value. For very small bandwidths the fixed-b approach and the traditional approach to asymptotics coincide.

The non-bold entries in Table 4 are empirical power. The patterns in power, are what one would expect given the local asymptotic limiting distributions. For large trend slopes ($\beta_2 = 10, 2$), power of the four tests is essentially the same as predicted by Theorem 4. As the bandwidth increases, power of all the tests decreases. This inverse relationship between power and bandwidth is well known in the fixed-b literature (see Kiefer and Vogelsang 2005).

For medium sized trend slopes ($\beta_2 = 0.2$) noticeable differences in power begin to emerge with t_{OLS} having substantially lower power than the other tests for $\theta > 2$. This lower power occurs even though t_{OLS} substantially over-rejects under the null when a small bandwidth is used ($b = b_{A91}$). With larger bandwidths, t_{OLS} severely under-rejects under the null and has no power. Both t_{BC} and t_{IV} have good power that is somewhat lower than t_{θ_0} for $\theta > 2$ but higher for $\theta < 2$. These power patterns are predicted by Theorem 5 because of differences between $\lambda^2_{\theta_0}$ and $\lambda^2_{\theta_1}$. For $\beta_2 = 0.2$, one can show that $\lambda^2_{\theta_1}$ is an increasing function in θ_1 over the range $\theta \in [1.8, 2.2]$ in our DGP. Therefore, when $\theta_1 > 2$, it follow that $\lambda^2_{\theta_0} < \lambda^2_{\theta_1}$ leading to t_{θ_0} having higher power. The opposite is true when $\theta_1 < 2$. In general the statistics cannot be ranked in terms of power because of this dependence on θ_1.

For small and very small trend slopes ($\beta_2 = 0.1, 0.005$), both t_{OLS} and t_{BC} are distorted under the null with severe over-rejections although t_{OLS} severely under-rejects with large bandwidths ($b = 0.5, 1.0$) when $\beta_2 = 0.1$. While t_{IV} is less size distorted than t_{OLS} and t_{BC}, it can have very low power for values of θ_1 that are far from the null value of 2. In contrast t_{θ_0} continues to have excellent size and good power[4].

In summary, the patterns in the finite sample simulations are consistent with the predictions of Theorems 4-7. Overall t_{θ_0} is the recommended statistic given its superior behavior under the null and its reliable power under the alternative.

6 Practical Recommendations

For point estimation, we recommend the IV estimator, i.e. the ratio of OLS trend slope estimators, given its relative robustness to the magnitude of the trend slopes. OLS and bias-corrected OLS are not recommended given that they can become severely biased for small to very small trend slopes. For inference, we strongly recommend the t_{θ_0} statistic given its superior behavior under the null and reliable power that is robust to

[4]Some readers may notice that for a given bandwidth, power of t_{θ_0} is the same regardless of the value of β_2. This occurs because we have configured the grids for θ so that $\beta_2(\theta_1 - \theta_0) = \beta_2(\theta_1 - 2)$ is the same across values of β_2.

the magnitude of the trend slopes. Good empirical practice is to report the IV estimator, $\widehat{\theta}$, along with the confidence interval constructed by inverting t_{θ_0}. Because this confidence interval must contain $\widehat{\theta}$, situations are avoided where the recommended point estimator lies outside the recommended confidence interval.

For confidence interval construction, there is also the practical need to choose a kernel and bandwidth and the Andrews (1991) approach is widely used. We do not explore this choice here from a theoretical perspective but encourage empirical researchers to use the fixed-b critical values provided by Bunzel and Vogelsang (2005) once a kernel and bandwidth have been chosen. Some practical suggestions on bandwidth choice are given in the empirical application.

7 Allowing for Structural Change

In the empirical application we report results that allow for a one-time structural change in the trend functions in each of the two time series. While a detailed analysis of trend ratio estimation and inference in the presence of structural change in the trend function is beyond the scope of this paper, we can sufficiently sketch the theory for the t_{θ_0} statistic for use in the empirical application.

Suppose we extend (1) and (2) to allow a shift in the intercept and trend slope at some time, T_B, as

$$y_{1t} = \mu_1 + \beta_1^{(1)} t + \varphi_1 DU_t + \phi_1 DT_t^* + u_{1t},$$

$$y_{2t} = \mu_2 + \beta_2^{(1)} t + \varphi_2 DU_t + \phi_2 DT_t^* + u_{2t},$$

where $DU_t = 1$ if $t > T_B$ and 0 otherwise and $DT_t^* = (t - T_B) DU_t$. For $t \leq T_B$, the trend slopes are given by $\beta_1^{(1)}$ and $\beta_2^{(1)}$ whereas for $t > T_B$, the trend slopes are given by $\beta_1^{(1)} + \phi_1$ and $\beta_2^{(1)} + \phi_2$. Letting $\beta_1^{(2)} = \beta_1^{(1)} + \phi_1$ and $\beta_2^{(2)} = \beta_2^{(1)} + \phi_2$, we can reparameterize the trend functions as

$$y_{1t} = \mu_1 + \beta_1^{(1)} (t - DT_T^*) + \varphi_1 DU_t + \beta_1^{(2)} DT_t^* + u_{1t}, \tag{26}$$

$$y_{2t} = \mu_2 + \beta_2^{(1)} (t - DT_T^*) + \varphi_2 DU_t + \beta_2^{(2)} DT_t^* + u_{2t}, \tag{27}$$

where $\beta_1^{(1)}$ and $\beta_2^{(1)}$ are the trend slopes up to date T_B and $\beta_1^{(2)}$ and $\beta_2^{(2)}$ are trend slopes after date T_B. The trend ratios are defined as $\theta^{(1)}$ and $\theta^{(2)}$ and can be estimated using ratios of the OLS estimators: $\widehat{\theta}^{(1)} = \widehat{\beta}_1^{(1)} / \widehat{\beta}_2^{(1)}$ and $\widehat{\theta}^{(2)} = \widehat{\beta}_1^{(2)} / \widehat{\beta}_2^{(2)}$.

Inference about $\theta^{(1)}$ and $\theta^{(2)}$ can be conducted using $\widehat{\theta}^{(1)}$and $\widehat{\theta}^{(2)}$ using the linear in slopes approach. Confidence intervals can be computed as the values of $\theta_0^{(i)}$, $i = 1, 2$, such that

$$\frac{\left(\widehat{\beta}_1^{(i)} - \theta_0^{(i)} \widehat{\beta}_2^{(i)} \right)^2}{\left(\widehat{\Omega}_{11} - 2\theta_0^{(i)} \widehat{\Omega}_{12} + \left(\theta_0^{(i)} \right)^2 \widehat{\Omega}_{22} \right) \left(\sum_{t=1}^{T} \widetilde{g}_t^2 \right)^{-1}} \leq cv_{\alpha/2}^2,$$

where $\widehat{\boldsymbol{\Omega}}$ is now computed using the OLS residuals from (26) and (27) and $\widetilde{g}_t$ are the OLS residuals from the regression of **i)** $(t - DT_t^*)$ on $1, DU_t, DT_t^*$ for $\theta_0^{(1)}$ and **ii)** DT_t^* on $1, t - DT_t^*, DU_t$ for $\theta_0^{(2)}$. From results in Bunzel and Vogelsang (2005) it follows that the fixed-b critical values depend on the functions of time included in the model which in turn depend on the date of the structural change. When we add the intercept and slope shift regressors to the models for y_{1t} and y_{2t}, fixed-b critical values must be adjusted to reflect these additional regressors. Details of the fixed-b critical values used for the case of structural change are given in the next section.

An alternative to using (26) and (27) to allow for different trend slopes (and intercepts) before and after the date T_B, would be to carry out analysis using subsamples of the time series. The advantage of (26) and (27) is that the full time span of the data is used to estimate the variance parameters needed to compute confidence intervals. As long as the correlation structure of the random components, u_{1t} and u_{2t}, does not have structural change, using (26) and (27) should lead to more powerful inference (tighter confidence intervals) than pure subsample analysis.

8 Empirical Application

Recent papers in the empirical climate literature have investigated apparent differences in temperature trends in surface temperature data and temperature trends in lower-troposphere data. See for example Santer *et al* (2005), Thorne *et al* (2007), Klotzbach *et al* (2009, 2010), Christy *et al* (2010), Po-Chedley and Fu (2012). An important quantity in these investigations is the so-called amplification ratio which is simply the ratio of temperature trends in the lower-troposphere to temperature trends at the surface. In terms of the statistical model in this paper, if we let y_{1t} denote a lower-troposphere temperature series and y_{2t} denote a surface temperature series, then $\theta = \beta_1/\beta_2$ is the amplification ratio. While estimated amplification ratios are routinely reported, confidence intervals are not usually provided. For example, Table 1 of Klotzbach *et al* (2009) reports $\widehat{\theta}$ values but does not provide confidence intervals.

We revisit the Klotzbach *et al* (2009) analysis and report confidence intervals for estimated amplification ratios using our recommended linear in slopes approach by inverting the t_{θ_0} statistic. We examine the same sample period of 1979-2008 as in Klotzbach *et al* (2009) and we also include results using more recent and updated data from the 1979-2014 period[5]. In addition, for the 1979-2014 sample we allow for structural change in the trend functions at the end of 1998 to accommodate the possibility that trend slopes are lower after 1998, i.e. the "pause" or "hiatus" in warming that has been vigorously debated in recent years. We report estimated amplification ratios and confidence intervals for the subperiods 1979-1998 and 1999-2014 using (26) and (27).

The data set used in our analysis includes surface temperatures from two sources and satellite lower-troposphere temperatures also from two sources. The sources for the surface data are the National Climate Data Center (NCDC) and the Hadley Centre (HADC). The sources for the satellite series are the University of Alabama Huntsville (UAH) and Remote Sensing Systems (RSS). For each of the four sources three temperature series are examined: **i)** a land series, **ii)** an ocean series and **iii)** a land+ocean (global) series. All temperature series are monthly and additional details on the data and their sources can be found in Klotzbach *et al* (2009).

A referee asked whether it is reasonable to assume that the temperature data used here is stationary around trends. The literature on amplification ratio estimation treats the temperature data as trend sta-

[5] We are grateful to Phil Klotzbach for providing us with the data.

tionary based on the view that unit roots in temperature data does not make sense on physical grounds. Regardless, it is useful to provide some formal statistical evidence regarding unit roots in the temperature data used here. The augmented Dickey-Fuller generalized least squares (ADF-GLS) unit root tests of Elliott, Rothenberg and Stock (1996) and Perron and Rodriguez (2003) were applied to the individual temperature series for the 1979-2014 case. In the case where the trend function is modeled as linear with no breaks, the ADF-GLS statistics range from -3.44 to -6.89 with 5% and 1% critical values of -2.89 and -3.48. When a single break in trend at an unknown date is permitted, the infimum ADF-GLS statistics range from -4.58 to –8.61 with 5% and 1% critical values of -3.91 and -4.49. Unit roots are rejected at the 1% level in nearly all cases so there is strong evidence supporting trend stationarity of the temperature series[6].

Plots of the data for the 1979-2014 time period are provided in the SI document. While temperatures appear to be increasing, it is not easy to see the relative rates of warming between the troposphere and surface series. Formal statistical analysis is clearly needed to learn about amplification ratios.

When constructing confidence intervals for either trend slope parameters or trend ratios (amplification ratios), we implemented the relevant long run variance estimator using the Daniell kernel as was done in the Monte Carlo simulations. We used the data dependent bandwidth formula of Andrews (1991), $b = b_{A91}$, in conjunction with fixed-b critical values computed using (25). The values of b_{A91} ranged from approximately 0.02 to 0.1 across the various series. As a robustness check, confidence intervals for amplification ratios are also reported using the Bartlett kernel with $M = T$ ($b = 1$) following Vogelsang and Franses (2005) and McKitrick and Vogelsang (2014). We use the label VF-MV for the Bartlett case and the relevant right tail 2.5% critical value is 6.482 (see the entry for t_2^* in Table 1 of Vogelsang and Franses 2005).

For the case where we allow for a possible intercept shift and slope shift at the end of 1998, (25) is no longer valid because the critical values depend on the presence of the intercept and trend slope regressors included in the model. Furthermore, the critical values also depend on the timing of the structural change through the ratio of the number of observations before the structural change relative to the sample size: $\tau_B = T_B/T$. In the application for the 1979-2014 sample we have 20 years of monthly data up to the end of 1998 with 36 years of monthly data in total. Therefore, $\tau_B = 240/432 = 0.556$. In this case the fixed-b critical values for the Daniell kernel are given by the function

$$cv_{0.025}(b) = 1.9600 - 1.2196b + 57.0925b^2 + 251.8196b^3 - 391.8571b^4 + 190.8516b^5, \tag{28}$$

where the coefficients implicitly depend on the dummy variables being included in the model. Similarly, the 2.5% right tail critical value for VF-MV depends on the presence of the dummy variables and is equal to 8.422 when $\tau_B = 0.556$.

Table 5 reports OLS estimates of linear trend slopes for the individual temperature series. Estimated trend slopes are scaled to be in degrees Celsius per decade. Confidence intervals are calculated using the formula

$$\widehat{\beta} \pm se(\widehat{\beta}) \cdot cv_{0.025}(b).$$

[6] These findings are consistent with Estrada, Perron and Martínez-López (2013) where unit roots were rejected in global temperature series using annual data from 1850-2010.

In the cases without structural change, $\widehat{\beta}$ is the OLS trend slope estimator from a regression with an intercept and linear trend, $cv_{0.025}(b)$ is given by (25) and

$$se(\widehat{\beta}) = \sqrt{\left(\widehat{\gamma}_0 + 2\sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\widehat{\gamma}_j\right)\left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}}$$

where $\widehat{\gamma}_j = T^{-1}\sum_{t=j+1}^{T}\widehat{u}_t\widehat{u}_{t-j}$ and $\widehat{u}_t$ are the OLS residuals from the regression of the temperature series on an intercept and time. For the case of structural change, the relevant trend slopes are estimated using (26) and (27) with $T_B = 240$ and $cv_{0.025}(b)$ is given by (28). The standard error formula changes to

$$se(\widehat{\beta}) = \sqrt{\left(\widehat{\gamma}_0 + 2\sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\widehat{\gamma}_j\right)\left(\sum_{t=1}^{T}\widetilde{g}_t^2\right)^{-1}}$$

where $\widehat{\gamma}_j$ is computed using the OLS residuals from regressions (26) and (27) and the computation of $\widetilde{g}_t$ is described previously.

For the 1979-2008 time span the point estimates and confidence intervals in Table 5 are similar to the results reported by Klotzbach *et al* (2009) in their Table 1. The results are not identical because the data over the 1979-2008 period has changed due to adjustments. Using the extended data to 2014 gives similar results. All temperature series have positive trend slopes that are statistically significant at the 5% level. Most series exhibit less warming over the 1979-2014 period relative to the 1979-2008 period.

When we allow for structural change at the end of 1998, we see that estimated trend slopes tend to be larger in the 1979-1998 subperiod and smaller in the 1998-2014 subperiod. The RSS series has very small estimated trend slopes for the 1999-2014 subperiod. Confidence intervals are generally wider in the subsamples and often contain zero. Wider confidence intervals reflect reduced precision in the estimation of trend slopes because of the smaller number of observations within subperiods.

For illustrative purposes Table 6 reports estimated amplification ratios (ratios of trend slopes) using the OLS estimator which is not recommended by our analysis because of the potential for bias when trend slopes are not large. The top panel of Table 6 reports estimated amplification ratios with troposphere series in the numerator. The bottom panels flips things and places the troposphere series in the denominator. For the case of the combined land-ocean series, the estimated amplification ratios are less than one regardless as to whether the troposphere series are in the numerator or denominator. The troposphere series have smaller trend slopes and when the troposphere series is y_{2t}, $\widetilde{\theta}$ is coming from a regression with a relatively small value of β_2. Recall from Theorem 1, Case 2, that $\widetilde{\theta}$ is biased when β_2 is not large and the magnitude of the bias is inversely related to β_2. Therefore, $\widetilde{\theta}$ can be less than one for both ratios because in the case where the troposphere variable is on the right hand side, there is substantial downward bias in $\widetilde{\theta}$. Table 6 nicely illustrates the pitfalls of using $\widetilde{\theta}$ to estimate trend ratios.

Results using our recommended estimator $\widehat{\theta}$, the ratio of the trends slope estimators, are given in Table 7. This was the estimator used by Klotzbach *et al* (2009). We only report troposphere/surface amplification ratios because surface/troposphere ratios are exactly the reciprocals of the numbers reported in Table 7. For the full samples, we see that, except for the UAH/NCDC ratio for the ocean data from the 1979-2014 time span, the values of $\widehat{\theta}$ are less than one indicating less warming in the lower-troposphere than at the surface. When we allow for structural change after 1998, we see similar amplification ratios from 1979-1998. For the 1999-2014 period, the amplification ratios are very different and depend on whether the UAH or RSS series is used for the troposphere.

While the $\widehat{\theta}$ values generally suggest slower warming in the troposphere than at the surface, it is important to calculate confidence intervals for $\widehat{\theta}$ so that sampling variability of $\widehat{\theta}$ can be taken into account when comparing $\widehat{\theta}$ values with amplification ratios suggested by climate models. Table 8 reports 95% confidence intervals using our recommended linear in slopes approach. Over land, the confidence intervals are below one for both the 1979-2008 and 1979-2014 samples indicating less warming in the troposphere relative to the surface. Over the entire globe (land+ocean), the confidence intervals are also below one in three of four cases for the 1979-2014 data. Over the oceans, the confidence intervals include values greater than one. Results are similar using the Daniell kernel and VF-MV although VF-MV tends to give slightly wider confidence intervals. Wider confidence intervals are to be expected given the more conservative nature of the VF-MV approach because of the large bandwidth being used.

Confidence intervals that allow structural change after 1998 are, for the period 1979-1998, qualitatively similar but wider than the confidence intervals obtained with the full samples when structural change is not included in the model. For the 1999-2014 period, confidence intervals are very different and we see examples of cases 2 and 3 described in Section 4.2. The possibility of disjoint confidence intervals (case 2) or the entire real line (case 3) depends on whether the series in the denominator has a trend slope that is statistically significant. In Table 5 we saw that the surface series has cases where the null of a zero trend slope cannot be rejected. When that happens, cases 2 and 3 can be obtained for confidence intervals for the amplification ratios. In general, the confidence intervals for 1999-2014 are very wide because estimated trend slopes are smaller after 1998. The 1999-2014 results nicely illustrate the difficulty in precisely estimating a ratio of trend slopes when trend slopes are small.

The confidence intervals in Table 8 can be used to test whether an amplification ratio of a climate model is consistent with the data. For example, Klotzbach *et al* (2009) tested the null hypothesis, $H_0 : \theta = 1.2$. Except for the UAH/NCDC ratio, this null value falls outside of the confidence intervals and the hypothesis is rejected for both the 1979-2008 and 1979-2014 time periods. Klotzbach *et al* (2009) were criticized for testing the same null value of 1.2 over both land and oceans. In response to this criticism, Klotzbach *et al* (2010) tested a null amplification ratio of 1.1 over land and 1.6 over the oceans. In all cases for observed land temperature series for 1979-2008 and 1979-2014, 1.1 falls outside of the confidence intervals. Likewise, for all ocean temperature series for 1979-2008 and 1979-2014, 1.6 falls outside of the confidence intervals. Table 8 confirms the original findings of Klotzbach *et al* (2010) and shows that their findings continues to hold with data updated to 2014.

When we allow for structural change after 1998, results similar to Klotzbach *et al* (2009,2010) generally hold for 1979-1998. For 1999-2014, the confidence intervals are usually either not informative (the entire real) or are disjoint. Interestingly, for the RSS ratios the null hypothesis of 1.1 is rejected for the land series and the null hypothesis of 1.6 is rejected in three of four cases for the ocean series. For the RSS ratios, the results are similar to Klotzbach *et al* (2009,2010).

It is important to understand how Klotzbach *et al* (2009,2010) carried out their tests and how their approach is related to the confidence intervals being used here. Klotzbach *et al* (2009,2010) used the linear in slopes approach, i.e. the t_{θ_0} statistic. However, they used a parametric estimator of the long run variance based on an AR(1) model for serial correlation in the temperature series. In contrast, we are using a nonparametric kernel based estimator of the long run variance and an associated critical value that depends on the kernel and the bandwidth. Our approach allows for more general forms of serial correlation and uses critical values that are known to reduce finite sample type 1 error distortions. In addition, by reporting confidence intervals for amplification ratios of observed temperature series, we allow the reader to immediately and easily test whether a particular null hypothesis about an amplification ratio can or cannot be rejected by the data. There is no need for us to take a stand on what are the interesting, relevant or important null hypotheses to test regarding amplification ratios.

9 Conclusion

In this paper we analyze estimation and inference of the ratio of trend slopes of two time series with linear deterministic trend functions. We consider three estimators of the trend slope ratio: OLS, bias-corrected OLS, and IV. Asymptotic theory indicates that when the magnitude of the trend slopes are large relative to the noise in the series, the three estimators are approximately unbiased and have essentially equivalent sampling distributions. For small trend slopes, the IV estimator tends to remain unbiased whereas OLS and bias-corrected OLS can have substantial bias. For very small trend slopes all three estimators become poor estimators of the trend slopes ratio.

We analyze four t-statistics for testing hypotheses about the trend slopes ratio. We consider t-statistics based on each of the three estimators of the trend slopes ratio and we propose a fourth t-statistic based on an alternative testing approach. Asymptotic theory indicates that the alternative test dominates the other three tests in terms of size and has reliable power that is robust to the magnitudes of the trend slopes.

Finite sample simulations show that the predictions of the asymptotic theory tend to hold in practice. Based on the asymptotic theory and finite sample evidence we recommend that the IV estimator be used to estimate the trend slopes ratio and that confidence intervals be computed using our alternative test statistic. A nice property of our recommendation is that the IV estimator is always contained in the confidence interval even though the confidence interval is not constructed using the IV estimator itself.

We carried out an empirical analysis of amplification ratios in warming trends in the lower-troposphere relative to the surface. Using the same temperature series as Klotzbach *et al* (2009,2010) but extended to 2014, we find that estimated amplification ratios tend be less than one. Using our recommended linear in slopes confidence intervals for the temperature based amplification ratios, we confirm the findings of Klotzbach *et al* (2010) that amplification ratios of 1.1 and 1.6 are rejected at the 5% level over land and oceans respectively using data from 1979-2008. Our results also show that these findings continue to hold using data from 1979-2014. If structural change is allowed at the end of 1998 to accommodate the "pause" or "hiatus" in warming since 1999, inference regarding amplification ratios within sub-periods is less sharp but is generally consistent with the findings of Klotzbach *et al* (2009,2010).

References

Andrews, D. W. K.: (1991), Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **59**, 817–854.

Bunzel, H. and Vogelsang, T. J.: (2005), Powerful trend function tests that are robust to strong serial correlation with an application to the prebisch-singer hypothesis, *Journal of Business and Economic Statistics* **23**, 381–394.

Canjels, E. and Watson, M. W.: (1997), Estimating deterministic trends in the presence of serially correlated errors, *Review of Economics and Statistics* **May**, 184–200.

Christy, J. R., Herman, B., Pielke Sr., R. A., Klotzbach, P., McNider, R. T., Hnilo, J. J., Spencer, R. W., Chase, T. and Douglass, D.: (2010), What do observational datasets say about modeled tropospheric temperature trends since 1979?, *Remote Sensing* **2**(9), 2148–2169.

Elliott, G., Rothenberg, T. J. and Stock, J. H.: (1996), Efficient tests for an autoregressive unit root, *Econometrica* **64**, 813–836.

Estrada, F., Perron, P. and Martínez-López, B.: (2013), Statistically derived contributions of diverse human influences to twentieth-century temperature changes, *Nature Geoscience* **6**(12), 1050–1055.

Evans, P.: (1997), How fast do economies converge?, *Review of Economics and Statistics* **79**(2), 219–225.

Fagerberg, J.: (1994), Technology and international differences in growth rates, *Journal of Economic Literature* **32**, 1147–75.

Harvey, D., Leybourne, S. J. and Taylor, A. M. R.: (2009), Simple, robust and powerful tests of the breaking trend hypothesis, *Econometric Theory* **25**, 995–1029.

Jansson, M.: (2004), The error rejection probability of simple autocorrelation robust tests, *Econometrica* **72**, 937–946.

Kiefer, N. M. and Vogelsang, T. J.: (2005), A new asymptotic theory for heteroskedasticity-autocorrelation robust tests, *Econometric Theory* **21**, 1130–1164.

Klotzbach, P. J., Pielke Sr., R. A., Pielke Jr., R. A., Christy, J. R. and McNider, R. T.: (2009), An alternative explanation for differential temperature trends at the surface and in the lower troposphere, *J. Geophys. Res.* **114**, D21102.

Klotzbach, P. J., Pielke Sr., R. A., Pielke Jr., R. A., Christy, J. R. and McNider, R. T.: (2010), Correction to "An alternative explanation for differential temperature trends at the surface and in the lower troposphere", *J. Geophys. Res.* **115**, D01107.

McKitrick, R. R. and Vogelsang, T. J.: (2014), HAC robust trend comparisons among climate series with possible level shifts, *Environmetrics* **25**(7), 528–547.

Perron, P. and Rodriguez, G.: (2003), GLS detrending, efficient unit root tests and structural change, *Journal of Econometrics* **115**, 1–27.

Perron, P. and Yabu, T.: (2009), Estimating deterministic trends with an integrated or stationary noise component, *Journal of Econometrics* **151**(1), 56–69.

Po-Chedley, S. and Fu, Q.: (2012), Discrepancies in tropical upper tropospheric warming between atmospheric circulation models and satellites, *Environmental Research Letters* **7**(4), 044018.

Santer, B. D., Wigley, T. M. L., Mears, C., Wentz, F. J., Klein, S. A., Seidel, D. J., Taylor, K. E., Thorne, P. W., Wehner, M. F. and Gleckler, P. J.: (2005), Amplification of surface temperature trends and variability in the tropical atmosphere, *Science* **309**(5740), 1551–1556.

Staiger, D. and Stock, J. H.: (1997), Instrumental variables regression with weak instruments, *Econometrica* **65**, 557–586.

Sun, Y.: (2014), Fixed-smoothing asymptotics in a two-step generalized method of moments framework, *Econometrica* **82**(6), 2327–2370.

Sun, Y., Phillips, P. C. B. and Jin, S.: (2008), Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing, *Econometrica* **76**, 175–194.

Thorne, P. W., Parker, D. E., Santer, B. D., McCarthy, M. P., Sexton, D. M. H., Webb, M. J., Murphy, J. M., Collins, M., Titchner, H. A. and Jones, G. S.: (2007), Tropical vertical temperature trends: A real discrepancy?, *Geophys. Res. Lett.* **34**(16).

Tomljanovich, M. and Vogelsang, T. J.: (2002), Are U.S. regions converging? using new econometric methods to examine old issues, *Empirical Economics* **27**, 49–62.

Vogelsang, T. J.: (1998), Trend function hypothesis testing in the presence of serial correlation correlation parameters, *Econometrica* **66**, 123–148.

Vogelsang, T. J. and Franses, P. H.: (2005), Testing for common deterministic trend slopes, *Journal of Econometrics* **126**, 1–24.

Vogelsang, T. J. and Nawaz, N.: (2015), Estimation and inference of linear trend slope ratios with I(1) errors. Working paper, Department of Economics, Michigan State University.

Zhang, X. and Shao, X.: (2013), Fixed-smoothing asymptotics for time series, *The Annals of Statistics* **41**(3), 1329–1349.

Table 1: Finite Sample Means and Standard Deviations
10,000 Replications, $\theta = 2$.

T	β_1	β_2	Mean			Standard Deviation		
			OLS	OLS_{BC}	IV	OLS	OLS_{BC}	IV
50	20	10	2.000	2.000	2.000	0.003	0.003	0.003
	10	5	2.000	2.000	2.000	0.006	0.006	0.006
	4	2	1.998	2.000	2.000	0.015	0.015	0.015
	.4	.2	1.816	2.014	2.014	0.117	0.156	0.156
	.3	.15	1.697	2.024	2.024	0.132	0.214	0.214
	.2	.1	1.441	2.058	2.058	0.145	0.353	0.353
	.1	.05	0.911	2.242	2.288	0.185	1.051	14.117
	.04	.02	0.553	1.409	1.897	0.175	1.454	75.360
	.004	.002	0.463	0.544	1.931	0.155	0.860	191.451
	0	0	0.462	0.534	2.114	0.155	0.840	160.777
100	20	10	2.000	2.000	2.000	0.001	0.001	0.001
	10	5	2.000	2.000	2.000	0.002	0.002	0.002
	4	2	1.999	2.000	2.000	0.005	0.005	0.005
	.4	.2	1.946	2.002	2.002	0.050	0.054	0.054
	.3	.15	1.906	2.003	2.003	0.063	0.073	0.073
	.2	.1	1.802	2.007	2.007	0.082	0.110	0.110
	.1	.05	1.422	2.028	2.028	0.102	0.232	0.232
	.04	.02	0.778	2.192	2.205	0.130	0.799	3.965
	.004	.002	0.469	0.602	0.263	0.110	0.719	73.150
	0	0	0.465	0.528	0.631	0.109	0.619	89.149
200	20	10	2.000	2.000	2.000	0.000	0.000	0.000
	10	5	2.000	2.000	2.000	0.001	0.001	0.001
	4	2	2.000	2.000	2.000	0.002	0.002	0.002
	.4	.2	1.985	2.000	2.000	0.019	0.019	0.019
	.3	.15	1.974	2.000	2.000	0.025	0.025	0.025
	.2	.1	1.944	2.001	2.001	0.035	0.038	0.038
	.1	.05	1.796	2.003	2.003	0.058	0.077	0.077
	.04	.02	1.244	2.021	2.021	0.074	0.201	0.201
	.004	.002	0.484	0.968	1.644	0.078	0.731	88.359
	0	0	0.469	0.517	0.862	0.076	0.437	50.018

Note: OLS, OLS_{BC} and IV denote the estimators given by (7), (10), and (12) respectively.

Table 2: Empirical Null Rejection Probabilities, 5% Nominal Level, 10,000 Replications.
$H_0 : \theta = \theta_0 = 2$, $H_1 : \theta \neq 2$.

			$b = 0.1$				$b = 0.5$				$b = 1.0$			
T	β_1	β_2	t_{OLS}	t_{BC}	t_{IV}	t_{θ_0}	t_{OLS}	t_{BC}	t_{IV}	t_{θ_0}	t_{OLS}	t_{BC}	t_{IV}	t_{θ_0}
50	20	10	.065	.065	.065	.065	.050	.051	.051	.051	.049	.053	.053	.053
	10	5	.065	.065	.065	.065	.049	.051	.051	.051	.038	.053	.053	.053
	4	2	.066	.065	.065	.065	.043	.051	.051	.051	.010	.053	.053	.053
	.4	.2	.236	.075	.059	.065	.001	.056	.050	.051	.000	.055	.050	.053
	.3	.15	.356	.087	.059	.065	.001	.062	.050	.051	.000	.059	.049	.053
	.2	.1	.610	.119	.061	.065	.001	.075	.049	.051	.000	.071	.049	.053
	.1	.05	.981	.261	.082	.065	.004	.131	.052	.051	.000	.122	.051	.053
	.04	.02	1.00	.538	.138	.065	.203	.340	.079	.051	.031	.250	.074	.053
	.004	.002	1.00	.874	.232	.065	.708	.719	.129	.051	.258	.546	.116	.053
	0	0	1.00	.882	.231	.065	.721	.730	.132	.051	.262	.552	.118	.053
100	20	10	.054	.054	.054	.054	.054	.053	.053	.053	.049	.052	.052	.052
	10	5	.054	.054	.054	.054	.054	.053	.053	.053	.042	.052	.052	.052
	4	2	.054	.054	.054	.054	.047	.053	.053	.053	.018	.052	.052	.052
	.4	.2	.140	.057	.054	.054	.001	.055	.054	.053	.000	.054	.052	.052
	.3	.15	.198	.059	.054	.054	.000	.056	.054	.053	.000	.055	.053	.052
	.2	.1	.341	.067	.053	.054	.000	.060	.054	.053	.000	.059	.052	.052
	.1	.05	.811	.106	.055	.054	.000	.080	.053	.053	.000	.076	.054	.052
	.04	.02	1.00	.334	.072	.054	.001	.183	.054	.053	.000	.161	.056	.052
	.004	.002	1.00	.908	.211	.054	.855	.807	.121	.053	.356	.627	.109	.052
	0	0	1.00	.947	.226	.054	.908	.856	.124	.053	.412	.679	.114	.052
200	20	10	.047	.047	.047	.047	.046	.045	.045	.045	.050	.049	.049	.049
	10	5	.047	.047	.047	.047	.045	.045	.045	.045	.046	.049	.049	.049
	4	2	.046	.047	.047	.047	.044	.045	.045	.045	.027	.049	.049	.049
	.4	.2	.090	.048	.046	.047	.001	.046	.045	.045	.000	.049	.049	.049
	.3	.15	.121	.048	.047	.047	.000	.046	.046	.045	.000	.049	.049	.049
	.2	.1	.197	.050	.046	.047	.000	.047	.046	.045	.000	.050	.049	.049
	.1	.05	.504	.061	.047	.047	.000	.053	.046	.045	.000	.056	.049	.049
	.04	.02	.995	.135	.049	.047	.000	.087	.047	.045	.000	.087	.050	.049
	.004	.002	1.00	.846	.149	.047	.818	.680	.092	.045	.183	.453	.080	.049
	0	0	1.00	.991	.216	.047	.989	.953	.122	.045	.605	.789	.113	.049

Note: The formulas t_{OLS}, t_{BC}, t_{IV}, and t_{θ_0} are given by (14), (15), (16) and (18) respectively.

Table 3: Finite Sample Proportions of Confidence Interval Shapes Based on t_{θ_0}.
5% Nominal Level, 10,000 Replications, $H_0 : \theta = \theta_0 = 2$, $H_1 : \theta \neq 2$.

			$b = 0.1$			$b = 0.5$			$b = 1.0$		
T	β_1	β_2	Case1	Case2	Case3	Case1	Case2	Case3	Case1	Case2	Case3
50	20	10	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	10	5	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	4	2	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	.4	.2	1.000	.000	.000	.987	.013	.000	.922	.078	.000
	.3	.15	1.000	.000	.000	.928	.072	.000	.812	.187	.000
	.2	.1	.999	.001	.000	.744	.255	.001	.612	.384	.004
	.1	.05	.723	.277	.000	.339	.605	.056	.293	.650	.057
	.02	.01	.202	.514	.284	.109	.517	.374	.104	.610	.287
	.002	.001	.071	.131	.799	.053	.270	.677	.053	.378	.569
	0	0	.066	.127	.806	.051	.271	.678	.052	.376	.572
100	20	10	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	10	5	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	4	2	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	.4	.2	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	.3	.15	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	.2	.1	1.000	.000	.000	.999	.001	.000	.986	.014	.000
	.1	.05	1.000	.000	.000	.897	.103	.000	.772	.227	.001
	.02	.01	.761	.239	.000	.386	.574	.040	.332	.623	.045
	.002	.001	.063	.165	.772	.053	.304	.643	.055	.404	.541
	0	0	.053	.122	.825	.050	.279	.671	.051	.382	.567
200	20	10	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	10	5	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	4	2	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	.4	.2	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	.3	.15	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	.2	.1	1.000	.000	.000	1.000	.000	.000	1.000	.000	.000
	.1	.05	1.000	.000	.000	1.000	.000	.000	.999	.001	.000
	.02	.01	1.000	.000	.000	.931	.069	.000	.819	.181	.000
	.002	.001	.113	.412	.475	.085	.449	.467	.081	.547	.372
	0	0	.043	.120	.837	.046	.274	.680	.051	.376	.573

Notes: Case 1 is $\theta_0 \in [r_1, r_2]$, Case 2 is $\theta_0 \in (-\infty, r_1] \cup [r_2, \infty)$ and Case 3 is $\theta_0 \in (-\infty, \infty)$.

Table 4: Finite Sample Power, 5% Nominal Level, $T = 100$, Two-sided Tests.
10,000 Replications, $H_0 : \theta = \theta_0 = 2$, $H_1 : \theta = \theta_1$, $\beta_1 = \theta_1\beta_2$.

		$b = b_{A91}$				$b = 0.5$				$b = 1.0$			
β_2	θ_1	t_{OLS}	t_{BC}	t_{IV}	t_{θ_0}	t_{OLS}	t_{BC}	t_{IV}	t_{θ_0}	t_{OLS}	t_{BC}	t_{IV}	t_{θ_0}
10	1.996	.953	.951	.951	.951	.509	.506	.506	.505	.420	.422	.422	.421
	1.997	.796	.788	.788	.787	.357	.354	.354	.353	.300	.301	.301	.300
	1.998	.487	.477	.477	.477	.201	.198	.198	.198	.182	.181	.181	.180
	1.999	.183	.179	.179	.179	.093	.092	.092	.092	.087	.089	.089	.089
	2.000	**.072**	**.070**	**.070**	**.070**	**.054**	**.053**	**.053**	**.053**	**.049**	**.052**	**.052**	**.052**
	2.001	.176	.180	.180	.181	.090	.093	.093	.093	.082	.085	.085	.085
	2.002	.468	.478	.478	.478	.200	.205	.205	.205	.173	.180	.180	.180
	2.003	.780	.787	.787	.787	.346	.350	.350	.351	.295	.300	.300	.300
	2.004	.949	.951	.951	.951	.500	.503	.503	.504	.413	.419	.419	.419
2	1.980	.960	.952	.952	.951	.516	.511	.511	.505	.360	.426	.426	.421
	1.985	.818	.791	.791	.787	.364	.357	.357	.353	.234	.304	.304	.300
	1.990	.522	.481	.481	.477	.209	.199	.199	.198	.121	.183	.183	.180
	1.995	.202	.181	.181	.179	.094	.092	.092	.092	.046	.090	.090	.089
	2.000	**.075**	**.071**	**.071**	**.070**	**.047**	**.053**	**.053**	**.053**	**.018**	**.052**	**.052**	**.052**
	2.005	.157	.179	.179	.181	.072	.092	.092	.093	.034	.085	.085	.085
	2.010	.433	.474	.474	.478	.176	.203	.203	.205	.092	.179	.179	.180
	2.015	.752	.784	.784	.787	.321	.345	.345	.351	.195	.297	.297	.300
	2.020	.938	.950	.950	.951	.470	.499	.499	.504	.320	.416	.416	.419
.2	1.80	.997	.964	.962	.951	.202	.584	.573	.505	.000	.482	.475	.421
	1.85	.971	.828	.819	.787	.082	.407	.399	.353	.000	.341	.334	.300
	1.90	.836	.530	.519	.477	.023	.229	.223	.198	.000	.204	.200	.180
	1.95	.533	.205	.198	.179	.005	.102	.098	.092	.000	.098	.095	.089
	2.00	**.198**	**.076**	**.073**	**.070**	**.001**	**.055**	**.054**	**.053**	**.000**	**.054**	**.052**	**.052**
	2.05	.058	.165	.156	.181	.000	.086	.084	.093	.000	.081	.079	.085
	2.10	.117	.446	.432	.478	.000	.188	.181	.205	.000	.171	.166	.180
	2.15	.346	.763	.749	.787	.000	.317	.310	.351	.000	.278	.271	.300
	2.20	.658	.939	.935	.951	.003	.451	.442	.504	.000	.381	.372	.419
.1	1.6	1.00	.979	.970	.951	.064	.682	.641	.505	.000	.563	.532	.421
	1.7	.999	.872	.847	.787	.014	.479	.446	.353	.000	.400	.371	.300
	1.8	.977	.596	.553	.477	.003	.271	.248	.198	.000	.238	.220	.180
	1.9	.842	.248	.218	.179	.000	.120	.109	.092	.000	.112	.102	.089
	2.0	**.513**	**.089**	**.071**	**.070**	**.000**	**.060**	**.054**	**.053**	**.000**	**.059**	**.052**	**.052**
	2.1	.173	.167	.132	.181	.000	.087	.077	.093	.000	.083	.072	.085
	2.2	.033	.442	.382	.478	.000	.183	.162	.205	.000	.165	.149	.180
	2.3	.044	.755	.702	.787	.000	.302	.274	.351	.000	.265	.241	.300
	2.4	.173	.936	.913	.951	.000	.426	.386	.504	.000	.353	.330	.419
.005	-6	1.00	.965	.256	.951	.000	.703	.132	.505	.000	.490	.119	.421
	-4	1.00	.938	.289	.787	.020	.713	.154	.353	.000	.512	.138	.300
	-2	1.00	.901	.354	.477	.328	.759	.189	.198	.038	.573	.170	.180
	0	1.00	.965	.394	.179	.891	.891	.206	.092	.387	.736	.180	.089
	2	**1.00**	**.782**	**.198**	**.070**	**.614**	**.595**	**.098**	**.053**	**.150**	**.426**	**.085**	**.052**
	4	1.00	.629	.044	.181	.058	.312	.029	.093	.002	.181	.029	.085
	6	1.00	.812	.019	.478	.000	.364	.022	.205	.000	.211	.026	.180
	8	.997	.926	.015	.787	.000	.438	.028	.351	.000	.267	.031	.300
	10	.889	.962	.015	.951	.000	.490	.035	.504	.000	.304	.038	.419

Note: The formulas t_{OLS}, t_{BC}, t_{IV}, and t_{θ_0} are given by (14), (15), (16) and (18) respectively. Bold entries are null rejection probabilities.

Table 5: Estimated Trend Slopes of Individual Temperature Series
Degrees Celsius per Decade, Klotzbach *et al* Data

	Land+Ocean	Land	Ocean
1979-2008	$\widehat{\beta}$	$\widehat{\beta}$	$\widehat{\beta}$
NCDC	.164 [.128, .201]	.289 [.224, .355]	.117 [.073, .161]
HADC	.176 [.135, .217]	.290 [.231, .349]	.134 [.095, .174]
UAH	.134 [.060, .208]	.178 [.102, .254]	.109 [.042, .175]
RSS	.142 [.072, .211]	.205 [.131, .278]	.111 [.048, .174]
1979-2014	$\widehat{\beta}$	$\widehat{\beta}$	$\widehat{\beta}$
NCDC	.148 [.120, .175]	.259 [.212, .306]	.106 [.073, .138]
HADC	.158 [.127, .188]	.255 [.212, .298]	.125 [.095, .154]
UAH	.133 [.081, .186]	.180 [.126, .234]	.107 [.058, .155]
RSS	.122 [.071, .174]	.166 [.110, .223]	.101 [.054, .147]
1979-1998	$\widehat{\beta}$	$\widehat{\beta}$	$\widehat{\beta}$
NCDC	.157 [.094, .221]	.247 [.137, .358]	.123 [.046, .200]
HADC	.169 [.097, .251]	.249 [.148, .349]	.139 [.070, .207]
UAH	.112 [-.017, .241]	.139 [.008, .269]	.098 [-.021, .217]
RSS	.153 [.031, .275]	.207 [.079, .336]	.127 [.016, .238]
1999-2014	$\widehat{\beta}$	$\widehat{\beta}$	$\widehat{\beta}$
NCDC	.084 [-.005, .172]	.144 [-.011, .299]	.061 [-.047, .168]
HADC	.091 [-.009, .192]	.136 [-.005, .276]	.083 [-.013, .180]
UAH	.122 [-.059, .302]	.180 [-.002, .362]	.088 [-.078, .255]
RSS	.028 [-.142, .198]	-.001 [-.181, .179]	.041 [-.114, .196]

Note: 95% confidence intervals in brackets using Daniell $k(x)$ function with Andrews (1991) data dependent bandwidth. Results for 1979-1998 and 1999-2014 are obtained using the full 1979-2014 sample and allowing for a shift in intercept and trend slope at the end of 1998.

Table 6: Estimated Amplification Ratios using $\widetilde{\theta}$: Klotzbach *et al* Data

	Land+Ocean		Land		Ocean	
Troposphere/Surface	1979-2008	1979-2014	1979-2008	1979-2014	1979-2008	1979-2014
UAH/NCDC	.910	.934	.563	.583	1.126	1.133
RSS/NCDC	.945	.912	.644	.623	1.106	1.067
UAH/HADC	.854	.883	.682	.708	.946	.941
RSS/HADC	.887	.862	.769	.745	.935	.889
Surface/Troposphere	1979-2008	1979-2014	1979-2008	1979-2014	1979-2008	1979-2014
NCDC/UAH	.707	.713	1.082	1.061	.522	.545
NCDC/RSS	.735	.752	1.127	1.127	.527	.552
HADC/UAH	.755	.762	.991	.980	.587	.626
HADC/RSS	.785	.804	1.018	1.024	.596	.636

Note: $\widetilde{\theta}$ is the OLS estimator of θ from regression (5).

Table 7: Estimated Amplification Ratios using $\widehat{\theta}$: Klotzbach *et al* Data

Land+Ocean	Troposphere/Surface	1979-2008	1979-2014	1979-1998	1999-2014
	UAH/NCDC	.818	.904	.713	1.454
	RSS/NCDC	.862	.829	.974	.336
	UAH/HADC	.763	.847	.665	1.331
	RSS/HADC	.805	.777	.908	.308
Land	**Troposphere/Surface**	**1979-2008**	**1979-2014**	**1979-1998**	**1999-2014**
	UAH/NCDC	.615	.694	.561	1.250
	RSS/NCDC	.707	.643	.838	-.006
	UAH/HADC	.614	.704	.559	1.326
	RSS/HADC	.705	.652	.834	-.006
Ocean	**Troposphere/Surface**	**1979-2008**	**1979-2014**	**1979-1998**	**1999-2014**
	UAH/NCDC	.930	1.010	.795	1.455
	RSS/NCDC	.948	.952	1.031	.676
	UAH/HADC	.810	.857	.706	1.059
	RSS/HADC	.826	.808	.915	.492

Note: $\widehat{\theta}$ is the ratio of OLS trend slopes (the IV estimator (10)). Results for 1979-1998 and 1999-2014 use the full sample allowing for a shift in intercept and slope at the end of 1998.

Table 8: Troposphere/Surface Amplification Ratio 95% Confidence Intervals, Klotzbach *et al* Data

	Land+Ocean		Land		Ocean	
1979-2008	Daniell	VF-MV	Daniell	VF-MV	Daniell	VF-MV
UAH/NCDC	[.498, 1.012]	[.429, 1.060]	[.446, .748]	[.371, .766]	[.398, 1.211]	[.372, 1.266]
RSS/NCDC	[.588, 1.036]	[.525, 1.076]	[.565, .823]	[.452, .914]	[.527, 1.153]	[.532, 1.172]
UAH/HADC	[.468, .942]	[.402, .984]	[.444, .738]	[.377, .756]	[.454, 1.045]	[.323, 1.099]
RSS/HADC	[.548, .969]	[.483, 1.017]	[.563, .814]	[.449, .921]	[.544, 1.008]	[.454, 1.034]
1979-2014	**Daniell**	**VF-MV**	**Daniell**	**VF-MV**	**Daniell**	**VF-MV**
UAH/NCDC	[.678, 1.070]	[.629, 1.171]	[.562, .811]	[.491, .895]	[.688, 1.242]	[.661, 1.328]
RSS/NCDC	[.599, .990]	[.529, 1.011]	[.510, .750]	[.356, .842]	[.678, 1.126]	[.662, 1.123]
UAH/HADC	[.640, .998]	[.594, 1.092]	[.574, .815]	[.498, .919]	[.621, 1.042]	[.562, 1.094]
RSS/HADC	[.563, .926]	[.497, .948]	[.518, .756]	[.371, .842]	[.601, .960]	[.535, .973]
1979-1998	**Daniell**	**VF-MV**	**Daniell**	**VF-MV**	**Daniell**	**VF-MV**
UAH/NCDC	[-.060, 1.047]	[-.467, 1.128]	[.117, .828]	[-.256, .904]	[-.327, 1.196]	[-.993, 1.276]
RSS/NCDC	[.441, 1.246]	[.072, 1.329]	[.552, 1.060]	[.137, 1.199]	[.331, 1.331]	[-.121, 1.397]
UAH/HADC	[-.055, .968]	[-.459, 1.030]	[.104, .802]	[-.281, .872]	[-.025, 1.063]	[-.859, 1.142]
RSS/HADC	[.417, 1.159]	[.069, 1.239]	[.551, 1.020]	[.147, 1.187]	[.456, 1.189]	[-.105, 1.249]
1999-2014	**Daniell**	**VF-MV**	**Daniell**	**VF-MV**	**Daniell**	**VF-MV**
UAH/NCDC	[$-\infty$, ∞]	[$-\infty$, ∞]	-4.150], [.416	[$-\infty$, ∞]	[$-\infty$, ∞]	[$-\infty$, ∞]
RSS/NCDC	1.126], [10.02	1.295], [4.645	.632], [40.26	.959], [7.370	1.487], [2.686	1.610], [2.646
UAH/HADC	[$-\infty$, ∞]	[$-\infty$, ∞]	-1.374], [.553	[$-\infty$, ∞]	1.865], [19.17	1.791], [2.692
RSS/HADC	1.032], [9.208	1.193], [3.915	.645], [19.55	.964], [3.964	[-85.97, 1.140]	1.259], [3.698

Note: 95% confidence intervals in brackets. Daniell uses the Daniell $k(x)$ function with Andrews (1991) data dependent bandwidth. VF-MV uses Bartlett $k(x)$ function with bandwidth equal to sample size ($b = 1$). Results for 1979-1998 and 1999-2014 use the full sample allowing for a shift in intercept and slope at the end of 1998.